Driver’s Age and Automated Vehicle Explanations

Qiaoning Zhang 1, X. Jessie Yang 1 and Lionel P. Robert Jr. 1

1 University of Michigan Ann Arbor, MI 48109; qiaoning@umich.edu (Q. Zhang); xijyang@umich.edu (X. J. Yang); lprobert@umich.edu (L. P. Robert)
* Correspondence: qiaoning@umich.edu

Abstract: Automated Vehicles (AV) have the potential to benefit our society. However, a lack of trust is a major barrier to the adoption of AVs. Providing explanations is one approach to facilitating AV trust by decreasing uncertainty about AV’s decision-making and action. However, explanations might increase drivers’ cognitive effort and anxiety. Because of differences in cognitive ability across age groups, it is not clear whether explanations are equally beneficial for drivers across age groups in terms of trust, effort, and anxiety. To examine this, we conducted a mixed-design experiment with 40 participants divided into three age groups (i.e., younger, middle-age, and older). Participants were presented with: (1) no explanation, or (2) explanation given before or (3) after the AV took action, or (4) explanation along with a request for permission to take action. Results suggest that the explanations provided before AV take actions produced the highest trust and lowest effort for all drivers regardless of age group. The request-for-permission condition led to the highest trust and lowest effort only for older drivers. Younger drivers had the lowest anxiety and effort under the AV-explanation-after-action condition; however, this condition produced the highest level of anxiety and effort in middle-age and older drivers, respectively. These results have important implications in designing AV explanations and promoting trust.

Keywords: Trust in AVs; Human-Machine Interface; Artificial Intelligence Transparency; Older Drivers; Automated Driving; AV Explanation; Artificial Intelligence Explainability; Driver’s Age

1. Introduction

Automated Vehicles (AV) have the potential to benefit our society in part because Americans outlive their ability to drive safely by an average of 7–10 years [1-3]. For many, aging can correspond with greater difficulty in driving and according to the U.S. Census Bureau, approximately 70 million individuals in the United States will be over the age of 64 by 2030 (Figure 1) [4]. This explains why AVs are suggested as one potential solution, but a lack of trust hinders that adoption across all age groups [5,6]. The Society of Automotive Engineers (SAE) classifies driving automation into six levels ranging from 0 to 5 as shown in Table 1. As the levels increase from 0 to 5, the need for driver involvement decreases. At SAE level 3, the human driver still has to intervene when asked to do so by the automated driving system, whereas at SAE levels 4 and 5, the automated driving system takes full responsibility for all of the driving tasks in some and all circumstances, respectively [7]. In this study, AV refers to SAE level 4 and higher vehicles.

AV explanations are one approach to promoting trust in the AV but driver age might undermine its impact. Explanations—reasoning or logic behind actions—have been shown to facilitate trust in automation. Explanations—reasoning or logic behind actions—have been shown to facilitate trust in automation [8]. Explanations reduce anxiety about the actions taken by automation [9-11]. Despite receiving little attention, age is likely to be important to determining the effectiveness of AV explanations. Aging often corresponds with greater difficulty in driving. Older drivers (55 and older) have been shown to be slower to respond at signal lights, have more difficulty in judging visuospatial relations, and be more prone to accidents at moderate to high speeds compared to those in...
Figure 1. Projections of the Older Adult Population: 2020 to 2060[4].

Table 1. SAE (J3016) Automation Levels[7]

<table>
<thead>
<tr>
<th>SAE Level</th>
<th>Name</th>
<th>Narrative definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No Automation</td>
<td>The full-time performance by the human driver of all aspects of the dynamic driving task, even when “enhanced by warning or intervention systems”</td>
</tr>
<tr>
<td>1</td>
<td>Driver Assistance</td>
<td>The driving mode-specific execution by a driver assistance system of “either steering or acceleration/deceleration” using information about the driving environment and with the expectation that the human driver performs all remaining aspects of the dynamic driving task</td>
</tr>
<tr>
<td>2</td>
<td>Partial Automation</td>
<td>The driving mode-specific execution by one or more driver assistance systems of both steering and acceleration/deceleration</td>
</tr>
<tr>
<td>3</td>
<td>Conditional Automation</td>
<td>Automated driving system monitors the driving environment with the expectation that the human driver will respond appropriately to a request to intervene even if a human driver does not respond appropriately to a request to intervene the car can pull over safely by guiding system</td>
</tr>
<tr>
<td>4</td>
<td>High Automation</td>
<td>The driving mode-specific performance by an automated driving system of all aspects of the dynamic driving task</td>
</tr>
<tr>
<td>5</td>
<td>Full Automation</td>
<td>Under all roadway and environmental conditions that can be managed by a human driver</td>
</tr>
</tbody>
</table>

younger (18–24 years) and middle-age groups (25-54 years) [12-14]. This has been attributed, at least in part, to their decrements in cognitive (e.g., cognitive processing speed, sustained attention), psychomotor (e.g., manual dexterity), and perceptual abilities [15]. However, little work has been done to understand the relationship between age and AV explanations.
To address this issue, we sought to understand the influence of the driver’s age on the impacts of AV explanations on the driver’s trust, effort and anxiety. We conducted a mixed-design experiment with 40 adults in three age groups (i.e., younger, middle-age, and older). Participants were presented with an AV that (1) gave no explanation, (2) gave an explanation before or (3) after the AV took action, or (4) gave an explanation along with a request for permission to take action. Results reveal that the driver’s age is indeed vital to understanding when AV explanations promote trust and reduce effort and anxiety.

This study provides several contributions to the literature. First, we demonstrated the importance of the driver’s age on the ability of AV explanations to promote trust and reduce effort and anxiety. Second, in doing so, we answered numerous calls for the development of more inclusive Artificial Intelligence (AI) systems [16,17]. These calls highlighted the problems of AI bias. In this paper we define AI bias as the underlying assumption that an AI system built for one subgroup is good for all groups. Finally, we provide design recommendations that are likely to help reduce age-based biases in AVs.

The rest of this paper is organized as follows. Section 2 presents the background for the work, and Section 3 illustrates the present study and hypothesis development. Section 4 describes the method, and the results are presented in Section 5 and discussed in Section 6. The conclusion of this paper is presented in Section 7.

2. Background

Driving automation includes Advanced Driver Assistance Systems (ADAS) and Automated Driving Systems [18]. In this paper the distinction between ADAS and ADS is based on SAE’s taxonomy. According to SAE, ADAS are represented by levels 1 and 2. ADAS are systems that assist humans with driving by performing some aspect of the driving (i.e., steering and braking/accelerating). ADS are represented by levels 3, 4 and 5. ADS are capable of driving with various degrees of human supervisor/intervention [7].

Explanations are reasons that underlie why an action was or should be taken [8,16]. Explanations have been used to support a range of automation such as: automated decision aid, driving automation, and recommender systems. Explanations reduce surprise and concerns about the actions taken by the automation, and facilitate trust in that automation [19-21]. For example, explanations in the human–automation interaction interface design promote users’ trust and acceptance [22].

Research examining AVs has also demonstrated that explanations can promote AV trust and reduce negative emotional reactions [8-11,23]. For example, researchers found that providing explanations about AV actions led to the highest level of positive emotional valence and AV acceptance compared to a no-explanation condition [9]. Also, people gave a higher rating to an interface that had speech output to explain the AV action in terms of its usability, anthropomorphism, and trust compared to interfaces that did not explain the AV actions [11].

Prior research also looked deeper to examine the timing of the AV explanation and the degree of AV autonomy, which could help us understand why or when AV explanations are likely to be effective at promoting trust and reducing anxiety. AV that provides an explanation before rather than after it takes action reduces the uncertainty associated with its action [8]. This reduction in uncertainty increases trust and decreases anxiety. Although providing an explanation after an AV takes action allows drivers to know the reasons behind the action and increases their understanding of the system, it cannot necessarily increase trust in the AV because of the lack of an alert and sense of control [10,24,25].

Other scholars suggest that the degree of autonomy by the AV might also influence the effectiveness of its explanations [26,27]. AV explanations might be more effective when the AV provides them and seeks approval from the driver to take action. Handing over driving control is one of the barriers to human drivers trusting and accepting advanced driving systems including AVs. A loss of driving control is always associated with a sense of worry [28]. With a lower degree of AV autonomy, which asks drivers for
permission to act, a higher level of control can be endowed. In one study, providing the driver with an explanation along with the option to approve or disapprove the AV action did not promote more trust and lower anxiety any more than just providing the explanation [8]. As such, there is little evidence to support the potential benefits of lower autonomy.

In summary, previous literature provides some guidance on how AV explanations can influence drivers. First, AV explanations are most effective when provided before an AV acts. At the same time, AV explanations are the least effective when provided after an AV acts. Finally, the AV’s level of autonomy has little impact on the effectiveness of its explanations. However, the literature offers little insight into the role of the driver’s age in the effectiveness of AV explanations.

3. Hypotheses Development

The literature on driver’s age and driving automation has found differences among age groups in several areas. First, older drivers vary greatly when it comes to being more or less comfortable with giving the AV control over the driving. Generally, younger drivers feel more comfortable giving up driving control to the driving automation. This was highlighted in a recent survey with 2,954 participants [29]. The survey found that younger drivers were more comfortable with letting the vehicle drive itself when compared to older drivers [26]. One reason often given is that younger drivers are more likely to have been exposed to driving automation, which is a strong predictor of whether a driver will be comfortable relinquishing control [30]. At the other end of the spectrum, older drivers have been shown to be less comfortable with giving up driving control [29,31]. For example, a recent study using a driving simulator found that older drivers prefer to retain some degree of driving control instead of giving it all up [31].

Second, the degree of trust afforded to driving automation has also shown to vary by age groups. Trust has been repeatedly shown to be one of the most important factors that influence people’s willingness to use driving automation [32-35]. Trust refers to “the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party.” [36]. Based on the literature, the degree of trust in driving automation varies greatly by the driver’s age. Younger drivers have shown higher trust in the ADS when compared to other age groups [37]. Middle-age drivers tend to be hesitant to trust an ADS [38]. Older drivers tend to distrust ADSs, especially if they do not understand how the systems operate [30]. This is even more problematic when the driving automation (i.e. ADS or ADAS) seems to fail and drivers do not comprehend why [39-41].

Third, the level of anxiety associated with the use of driving automation has also shown to vary by age groups. Anxiety has also been identified as an important factor in understanding the adoption of driving automation among different age groups [8,42]. Defined as a feeling of fear, worry, apprehension, or concern, anxiety can reduce cue utilization, shrink the perspective field, or reduce an individual’s environment scan [43]. Reimer, Mehler, Coughlin, Godfrey and Tan [28] in 2009 designed a field study using a real vehicle and found that older drivers tend to experience more anxiety when employing an ADAS. Anxiety has been shown to be negatively correlated with trust and ADS adoption [8]. High levels of anxiety can discourage drivers from trusting and further adopting ADS.

Finally, to be clear, there is literature that has shown that younger drivers do not always have positive attitudes toward driving automation. A study employing an automated driving simulator found that younger drivers’ use of an ADS decreased their driving enjoyment [44]. Another study employing an automated driving simulator found that middle-age drivers thought that ADAS were less useful than older drivers did [38].

Building on and integrating the literature on driver’s age and driving automation along with the literature on AV explanation we expect to see age differences on the impact
of AV explanation on outcomes such as trust and anxiety. There are several reasons to expect age differences. First, age differences regarding trust and anxiety have been found with regards to advance driving automation. Younger drivers are more inclined to trust driving automation, followed by middle-age drivers, while older drivers appear to be the most reluctant to trust driving automation [30,37,38]. Also, older drivers appear to experience more anxiety and stress than drivers in other age groups when employing driving automation [28].

That being said, we also know that AV explanations can promote trust and lower anxiety [8-10,45]. In particular, the timing of the explanation can be important for promoting trust. The explanation before rather than after the AV has taken action reduces the driver’s uncertainty about the driving situation, further increasing AV trust [8,11] and acceptance of the AV [11]. Similarly, Koo, Shin, Steinert and Leifer [10] found that when the AV explained what it was going to do before it acted, it significantly decreased drivers’ anxiety associated with driving. In addition, previous studies also illustrated the importance of automation autonomy on drivers’ perceptions of the automation (e.g., computer control, driving automation). The degree of autonomy refers to how much independence the automation has with regard to making decisions and taking actions without human intervention. Research has shown that automation with a high degree of autonomy is often less trusted [46,47].

Based on this literature, we derived the following two hypotheses to answer this research question: Does the driver’s age influence the relationship between AV explanations and the driver’s anxiety and trust?

**H1:** We expect to see mean differences in drivers’ AV trust both within and between age groups across AV explanation conditions.

**H2:** We expect to see mean differences in drivers’ anxiety both within and between age groups across AV explanation conditions.

### 4. Materials and Methods

This research complied with the American Psychological Association code of ethics and was approved by the university’s institutional review board. All participants provided informed consents.

#### 4.1. Participants

A total of 40 drivers (mean age = 34.9 years, standard deviation [SD] = 17.3 years) participated in this study. Before conducting the study, a power analysis was performed to determine the sample size. The effect size (ES) in this study was 0.5, considered to be medium using Cohen’s (1988) criteria. With an alpha = .05 and power = 0.8, the total sample size needed with this effect size (GPower3.1) is 12 for this “ANOVA: repeated measures, within-between interaction” group comparison. The results indicate that 40 participants in this study is sufficient to produce statistically significant results.

Participants were divided into three age groups: younger, middle- age and older adults. Twelve younger drivers (mean age = 21.5 years, SD = 0.26 years, 6 women) and 20 middle-age drivers (mean age = 30.1 years, SD = 0.64 years, 5 women) were recruited from email groups, and 8 older drivers (mean age = 66.9 years, SD = 0.75 years, 4 women) were recruited by advertisements on University of Michigan Health Research website. All participants were screened for inclusion criteria including driver’s license status, visual and hearing impairments, and susceptibility to simulator sickness. Each subject was paid $20 for participating.

#### 4.2. Study Design

We conducted a mixed-design experiment with a 4 (AV explanation conditions) x 3 (age groups) design in a controlled lab setting to examine the hypotheses. The human-
subject experiment involved 40 participants using a high-fidelity driving simulator. The sequence of the four AV explanation conditions was counterbalanced via a Latin square design. In each AV explanation condition, there were three unexpected and unique events (i.e., events by other drivers, events by police vehicles, and events of unexpected re-routes) in the environments of urban, highway, and rural. The simulated environment consists of urban, rural, and highway environments that are typical in the US. The urban and rural roads are four lanes, two for each traffic direction separated by lane markings. The highways comprise two two-lane roadways separated by grass median strips. All participants were exposed to the four exact same conditions with the three exact same events in each condition. The driving weather was sunny with clear visibility and the road conditions were good. Both the weather and road conditions remained the same across all conditions.

4.3. Independent Variables

The independent variables in this study include the driver’s age and the AV explanation conditions. Driver’s age consisted of three age groups which were based on the age categories used in other studies [50-52]. The three age groups were younger drivers (18–24 years), middle-age drivers (25–54 years), and older drivers (55 years and older). There were four AV explanation conditions. The first condition was the no AV explanation condition. In this condition, the AV provided no explanation about its actions to the driver. The second was the AV explanation before action condition. In this condition, the AV provided an explanation to the driver prior to it taking the action. The third explanation condition was the AV explanation after action condition. In this condition, the AV provided an explanation after it took an action. In the fourth condition, the AV provided an explanation then asked for the driver’s approval before taking or not taking any action. For example, before taking action the AV would explain “Unclear lane lines, reroute?” If the participant responded with a “Yes,” the AV would reroute; otherwise, the AV would continue with its original route.

4.4. Control Variables

The study includes participants’ trust propensity and physical workload as the control variables to reduce the possibility of alternative explanations. These variables have been found to influence people’s trust in AVs in previous studies [48,49].

4.5. Dependent Variables

The dependent variables in this study include trust and anxiety. We measured trust by adapting Muir’s (1987) 7-point Likert scale (1: strongly disagree; 7: strongly agree [50], which is a highly validated automation trust scale comparable with Jian’s trust scale [51]. The Muir scale consisted of six dimensions: competence, predictability, dependability, responsibility, reliability, and faith. We measured anxiety using a questionnaire adapted from Nass et al. [52] that is used to measure driver attitude. Anxiety comprised the averaged responses to four adjective items to describe feelings while driving the AV: fearful, afraid, anxious, and uneasy. All the items were rated on 7-point Likert rating scale (1: describes very poorly; 7: describes very well).

4.6. Apparatus

Participants rode a programmed AV in a simulated environment with a high-fidelity advanced driving simulator (Figure 2). The simulator consisted of a Nissan Versa sedan providing all manual controls and a simulation system running with programmable software (version 2.63 of Realtime Technology’s RTI). Four projectors displayed the visual environment to participants on four flat walls. The forward road scenes were projected on three walls about 16 feet in front of the driver (120-degree field of view), and the rear view was shown on a rear wall located 12 feet from the steering (40-degree field of view). Each
forward screen was set at a resolution of 1,400 x 1,050 pixels and updated at 60 Hz, and the rear screen was set at a resolution of 1,024 x 768 pixels.

In this study, the automation features of the driving simulator were programmed to simulate an AV with SAE level 4, wherein the driver was not required to actively monitor the environment and the longitudinal and lateral vehicle control, navigation, and responses to traffic control devices and other traffic elements were all undertaken by the AV. All production vehicle controls (e.g., turn signals, headlights, shifter) function as normal. The AV was able to function in all driving situations as well as the average human driver and obeyed all traffic laws.

![Figure 2. Driving Simulator.](image)

After starting a simulated drive, each participant was instructed to engage automation manually by pushing a button located on the lower right side of the steering wheel labeled ON/OFF, and then she/he would no longer need to actively monitor the roadway or control the vehicle.

To present the event explanations to participants, the simulator employed a neutral tone of a male voice with a standard American accent. As shown in Table 2, the events across four AV explanation conditions were chosen from previous literature and corresponded to realistic unexpected situations in automated driving [8,53]. All the events were programmable considering the accessibility of the driving simulator.

4.7. Procedure

Upon arrival, participants were briefed on the experiment and signed a consent form. Participants then completed a demographics survey. Participants received a 3-minute training session prior to the actual experiment. In the training session, participants were instructed about the AV’s attributes. Specifically, the participants were told that the vehicle is able to drive safely entirely on its own; the car is able to function in all driving situations as well as the average human driver; it obeys all traffic laws; it receives navigation information from external sources similar to Google Maps, and can change routes to reach a destination more quickly if one is identified or available; and the autonomous vehicle maintains lanes by visually sensing the lane lines on the roadway.

Participants were shown how to transfer the AV from manual control to automated mode by placing the vehicle in the center of the right lane and pressing the automated mode activation button. Participants also practiced giving permission to the AV via their verbal input. After the training, participants experienced the 60-minute experimental session with the four explanation conditions, as described. In each AV explanation condition, participants engaged in a 6- to 8-minute drive with events occurring at prescribed times in the drive at intervals of 1–2 minutes.

After each explanation condition, which included three events each, participants completed a follow-up survey consisting of two questionnaires to measure trust and anxiety. All questionnaire items were adapted from validated prior research. There was a 2-minute break between AV explanation conditions.
5. Results

To determine whether the measurement constructs were valid and reliable, we assessed construct validity and reliability. Construct validity determines the extent to which a scale captures the concept it is supposed to measure [54]. Convergent and discriminant validity as two subtypes of validity that make up construct validity. Both were assessed through exploratory factor analysis. Scale items that load at .70 or above on their corresponding construct indicates convergent validity while scale items that loaded at .35 or below on other constructs indicate discriminant validity [55]. All scale items generally met or exceeded both requirements. Construct reliability is a measure of the internal consistency associated with a set of scale items [56]. Cronbach’s alpha is the most widely used measure of reliability [57,58]. All construct reliabilities were at or above the acceptable threshold of .70 [59]. In addition, Table 3 lists the means, standard deviations, mode, median, and correlations.

To test the hypotheses with data from the 40 participants, we used SPSS Statistical 24 mixed linear model package. The alpha was set at 0.05 for all statistical tests. All post hoc comparisons utilized a Bonferroni alpha correction.

The mixed design controlled for the individual differences on prior driving automation experience. Nonetheless, we tested for individual differences in prior experience using various forms of ADAS or ADS. More specifically, individuals reported their prior experience with cruise control systems, adaptive cruise control systems, lane-

<table>
<thead>
<tr>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficiency Route Change</td>
<td>The AV rerouted in view of road construction ahead</td>
</tr>
<tr>
<td>Swerving Vehicle Ahead</td>
<td>The vehicle ahead was swerving, so the AV slowed down until the swerving vehicle exited the highway</td>
</tr>
<tr>
<td>Stopped Police Vehicle on Shoulder</td>
<td>A police vehicle stopped on shoulder, so the AV changed lane to avoid collision</td>
</tr>
<tr>
<td>Oversized Vehicle Ahead</td>
<td>There was an oversized vehicle ahead blocking roadway, so the AV slowed down until the oversized vehicle took turns at the intersection</td>
</tr>
<tr>
<td>Heavy Traffic Rerouting</td>
<td>Heavy traffic jam was reported ahead, so the AV rerouted</td>
</tr>
<tr>
<td>Police Vehicle Approaching</td>
<td>A police vehicle approached the AV from behind and activated siren. Then the AV pulled over and stopped</td>
</tr>
<tr>
<td>Stopped Police Vehicle on Shoulder</td>
<td>A police vehicle stopped on shoulder, so the AV changed lane to avoid collision</td>
</tr>
<tr>
<td>Abrupt Stopped Truck Ahead</td>
<td>There was roadway obstruction ahead. The AV changed lanes</td>
</tr>
<tr>
<td>Road Hazard Rerouting</td>
<td>The AV rerouted because it identified road hazard ahead</td>
</tr>
<tr>
<td>Police Vehicle Approaching</td>
<td>A police vehicle approached the AV from behind and activated siren. Then the AV asked driver’s permission whether to pull over</td>
</tr>
<tr>
<td>Unclear Lane Markings Rerouting</td>
<td>When the AV approached the intersection, the lane marking ahead was not clear. Then the AV asked driver’s permission whether to reroute</td>
</tr>
<tr>
<td>Vehicle with Flashing Hazard Lights Ahead</td>
<td>A vehicle in the left front lane was flashing the hazard light.</td>
</tr>
</tbody>
</table>

To test the hypotheses with data from the 40 participants, we used SPSS Statistical 24 mixed linear model package. The alpha was set at 0.05 for all statistical tests. All post hoc comparisons utilized a Bonferroni alpha correction.

The mixed design controlled for the individual differences on prior driving automation experience. Nonetheless, we tested for individual differences in prior experience using various forms of ADAS or ADS. More specifically, individuals reported their prior experience with cruise control systems, adaptive cruise control systems, lane-
departure warning systems, lane-keeping assistance systems, collision warning systems, and emergency braking systems. We found no significant differences among the three age groups in terms of their experience with ADAS or ADS (F=1.097, p=0.345).

Table 3. Descriptive Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev</th>
<th>Mode</th>
<th>Median</th>
<th>Trust</th>
<th>Anxiety</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trust</td>
<td>5.66</td>
<td>0.94</td>
<td>6.00</td>
<td>5.86</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>2.50</td>
<td>1.25</td>
<td>1.00</td>
<td>2.13</td>
<td>-0.36*</td>
<td>1</td>
</tr>
</tbody>
</table>

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

5.1. The Effect of Age and AV Explanation on Trust

The results of the Two-Way Mixed ANOVA showed that there was a significant interaction between age and AV explanation (F(6, 49) = 2.336, p = .035) as shown in Table 4. The following subsections presents the results of the post hoc comparisons.

Table 4. ANOVA Summary Table of Trust

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>Sum of squares</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>18.235</td>
<td>1</td>
<td>18.235</td>
<td>29.342</td>
<td>.000</td>
</tr>
<tr>
<td>Explanation Condition</td>
<td>4.617</td>
<td>3</td>
<td>1.539</td>
<td>2.477</td>
<td>.064</td>
</tr>
<tr>
<td>Age Groups</td>
<td>4.301</td>
<td>2</td>
<td>2.151</td>
<td>3.461</td>
<td>.034</td>
</tr>
<tr>
<td>Explanation condition x Age groups</td>
<td>8.710</td>
<td>6</td>
<td>1.452</td>
<td>2.336</td>
<td>.035</td>
</tr>
<tr>
<td>Physical Demand</td>
<td>14.968</td>
<td>1</td>
<td>14.968</td>
<td>24.084</td>
<td>.000</td>
</tr>
<tr>
<td>Trust Propensity</td>
<td>7.709</td>
<td>1</td>
<td>7.709</td>
<td>12.404</td>
<td>.001</td>
</tr>
<tr>
<td>Error</td>
<td>89.493</td>
<td>144</td>
<td>0.621</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>138.944</td>
<td>157</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: “df” indicates degree of freedom; “F” indicates F statistic; “p” indicates p value

5.1.1. Trust among age groups

For the No AV explanation condition, results showed that the trust for middle-age drivers (µMiddle-age = 5.49) was significantly lower than trust for the younger drivers (µYounger = 5.88). However, no difference was found between middle-age and older drivers (µOlder = 5.54), as well as between younger and older drivers (p>.05). Table 5 provides the means
and standard deviations for each condition. The means and their corresponding significant p values are depicted in Figure 3a.

For the AV explanation before action, there was no significant difference in trust (p>.05) among three age groups in the AV-explanation-before-action condition (μYounger = 6.17; μMid-age = 5.96; μOlder = 5.91). Table 5 provides the means and standard deviations for each condition. Figure 3b visually depicts the means and their corresponding significant p values.

The AV explanation after action results showed that the trust for middle-age drivers (μMid-age = 4.94) was significantly lower (p<.05) than for both the younger (μYounger = 5.99) and older (μOlder = 5.54) drivers in the AV-explanation-after-action condition. No difference in trust (p>.05) was found between younger and older drivers. Table 5 provides the means and standard deviations for each condition. Figure 3c visually depicts the means and their corresponding significant p values.

Request for permission results showed that the trust for older drivers (μOlder = 6.00) was significantly higher (p<.05) than for middle-age drivers (μMid-age = 5.54) in the request-for-permission condition. However, no difference in trust (p>.05) was found between younger and older drivers (μOlder = 5.54), or between younger and middle-age drivers (p>.05) in this condition. Table 5 provides the means and standard deviations for each condition. Figure 3d visually depicts the means and their corresponding significant p values.

5.1.2. Trust within age groups

For younger drivers, trust was significantly lower (p<.05) in the request-for-permission condition (μPermReq = 5.55) than in the conditions no explanation (μNExpl = 5.88), explanation before action (μBExpl = 6.17) and explanation after action (μAExpl = 5.99). Also, the AV-explanation-before-action condition led to the higher (p<.05) trust compared to
the no-AV-explanation and request-for-permission conditions. However, there were no differences in trust between the no-AV-explanation condition and the request-for-permission condition (p>.05). Table 5 provides the means and standard deviations for each condition. Figure 4a visually depicts the means and their corresponding significant p values.

For middle-age drivers, trust was significantly higher (p<.05) in the AV-explanation-before-action condition (µBExpl = 5.96) than in the no-AV-explanation (µNExpl = 5.49), AV-explanation-after-action (µAExpl = 4.94), and request-for-permission conditions (µPermReq = 5.54). conditions. However, the AV-explanation-after-action condition led to the lowest trust (p<.001). No significant difference was found in trust between no-AV-explanation and request-for-permission conditions (p>.05). Table 5 provides the means and standard deviations for each condition. Figure 4b visually depicts the means and their corresponding significant p values.

For older drivers, trust was significantly lower (p<.05) in the no-AV-explanation (µNExpl = 5.54) and AV-explanation-after-action (µAExpl = 5.54) conditions than in the AV-explanation-before-action (µBExpl = 5.91) and request-for-permission (µPermReq = 6.00) conditions. However, no difference in trust was found between AV-explanation-before-action and request-for-permission conditions (p>.05), and no difference was found between no-AV-explanation and AV-explanation-after-action (p>.05) conditions. Table 5 provides the means and standard deviations for each condition. Figure 4c visually depicts the means and their corresponding significant p values.

![Figure 4](image-url)

**Figure 4.** The average trust within age groups under four different conditions. (a) the average trust within younger drivers; (b) Middle-aged drivers; (c) Older drivers.
Table 5. Mean and Standard Deviation of Trust

<table>
<thead>
<tr>
<th>Age Groups</th>
<th>NExpl</th>
<th>BExpl</th>
<th>AExpl</th>
<th>PermReq</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Younger</td>
<td>5.88</td>
<td>0.25</td>
<td>6.17</td>
<td>0.20</td>
</tr>
<tr>
<td>Middle-age</td>
<td>5.49</td>
<td>0.20</td>
<td>5.96</td>
<td>0.15</td>
</tr>
<tr>
<td>Older</td>
<td>5.54</td>
<td>0.31</td>
<td>5.91</td>
<td>0.24</td>
</tr>
</tbody>
</table>

Note: “NExpl” indicates no-AV-explanation condition; "BExpl" indicates AV-explanation-before-action condition; "AExpl" indicates AV-explanation-after-action condition; "PermReq" indicates request-for-permission condition; “M” indicates mean; “SD” indicates standard deviation.

5.2. The Effect of Age and AV Explanation on Anxiety

The results of the Two-Way Mixed ANOVA showed that there was no significant interaction between age and AV explanation (F(6, 48) = 0.652, p = .689). The following subsections use the exploratory data analysis as the approach to explore the data and to summarize the main characteristic of the anxiety among and within age groups.

Table 6. ANOVA Summary Table of Anxiety

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>Sum of squares</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>15.657</td>
<td>1</td>
<td>15.657</td>
<td>10.875</td>
<td>.001</td>
</tr>
<tr>
<td>Explanation Condition</td>
<td>1.794</td>
<td>3</td>
<td>0.598</td>
<td>0.415</td>
<td>.742</td>
</tr>
<tr>
<td>Age Groups</td>
<td>14.995</td>
<td>2</td>
<td>7.498</td>
<td>5.208</td>
<td>.007</td>
</tr>
<tr>
<td>Explanation condition x Age groups</td>
<td>5.630</td>
<td>6</td>
<td>.938</td>
<td>.652</td>
<td>.689</td>
</tr>
<tr>
<td>Physical Demand</td>
<td>19.423</td>
<td>1</td>
<td>19.423</td>
<td>13.491</td>
<td>.000</td>
</tr>
<tr>
<td>Trust Propensity</td>
<td>2.275</td>
<td>1</td>
<td>2.275</td>
<td>1.580</td>
<td>.211</td>
</tr>
<tr>
<td>Error</td>
<td>207.319</td>
<td>144</td>
<td>1.440</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>246.333</td>
<td>157</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: “df” indicates degree of freedom; “F” indicates F statistic; “p” indicates p value
5.2.1. Anxiety among age groups

For the No AV explanation condition, the results showed that the anxiety for younger drivers ($\mu_{\text{Younger}} = 3.18$) was significantly higher ($p<.05$) than for middle-age ($\mu_{\text{Mid-age}} = 2.40$) and the older ($\mu_{\text{Older}} = 2.13$) drivers. No significant difference was found ($p>.05$) between middle-age and older drivers. Table 7 provides the means and standard deviations for each condition. The means and their corresponding significant $p$ values are depicted in Figure 5a.

For the AV explanation before action condition there was no significant difference in anxiety ($p>.05$) among the three age groups in the before-explanation condition ($\mu_{\text{Younger}} = 2.81; \mu_{\text{Mid-age}} = 2.36; \mu_{\text{Older}} = 2.34$). Table 7 provides the means and standard deviations for each condition. Figure 5b visually depicts the means and their corresponding significant $p$ values.

For the AV explanation after action condition, no significant difference in anxiety was found ($p>.05$) among the three age groups in the after-explanation condition ($\mu_{\text{Younger}} = 2.50; \mu_{\text{Mid-aged}} = 2.71; \mu_{\text{Older}} = 2.25$). Table 7 provides the means and standard deviations for each condition. Figure 5c visually depicts the means and their corresponding significant $p$ values.

For the request for permission there was no significant difference in anxiety ($p>.05$) among the three age groups in the permission-required condition ($\mu_{\text{Younger}} = 2.81; \mu_{\text{Mid-aged}} = 2.29; \mu_{\text{Older}} = 1.97$). Table 7 provides the means and standard deviations for each condition. Figure 5d visually depicts the means and their corresponding significant $p$ values.

![Figure 5](image-url)

**Figure 5.** The average anxiety between age groups under four different conditions. (a) the average anxiety between age groups under no explanation condition; (b) AV explanation before action condition; (c) AV explanation after action condition; (d) request for permission condition.
5.2.2. Anxiety within age groups

For younger drivers, anxiety was significantly higher ($p<.05$) in the no-AV-explanation condition ($\mu_{NExp}=3.18$) than in the AV-explanation-before-action ($\mu_{BExp}=2.81$), AV-explanation-after-action ($\mu_{AExp}=2.50$), and the request-for-permission ($\mu_{PermReq}=2.81$) conditions. However, there were no differences in anxiety among the AV-explanation-before-action, AV-explanation-after-action, and the request-for-permission conditions ($p>.05$). Table 7 provides the means and standard deviations for each condition. Figure 6a visually depicts the means and their corresponding significant $p$ values.

For middle-age drivers, the highest anxiety was generated ($p<.05$) in the AV-explanation-after-action condition ($\mu_{AExp}=2.71$) compared to the no-AV-explanation ($\mu_{NExp}=2.40$), AV-explanation-before-action ($\mu_{BExp}=2.36$), and the request-for-permission ($\mu_{PermReq}=2.29$) conditions. However, there were no differences in anxiety among the no-explanation, AV-explanation-before-action, and request-for-permission conditions ($p>.05$). Table 7 provides the means and standard deviations for each condition. Figure 6b visually depicts the means and their corresponding significant $p$ values.

For older drivers, no difference ($p>.05$) in anxiety was found among the no-AV-explanation condition ($\mu_{NExp}=2.13$), AV-explanation-before-action condition ($\mu_{BExp}=2.34$), AV-explanation-after-action condition ($\mu_{AExp}=2.25$), and the request-for-permission condition ($\mu_{PermReq}=1.97$). Table 7 provides the means and standard deviations for each condition. Figure 6c visually depicts the means and their corresponding significant $p$ values.

![Figure 6](image1.png)  
**Figure 6.** The average anxiety within age groups under four different conditions. (a) the average anxiety within younger drivers; (b) Middle-aged drivers; (c) Older drivers.
Table 7. ANOVA Summary Table of Anxiety

<table>
<thead>
<tr>
<th>Age Groups</th>
<th>NExpl</th>
<th>BExpl</th>
<th>AExpl</th>
<th>PermReq</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Younger</td>
<td>3.18</td>
<td>1.06</td>
<td>2.81</td>
<td>1.06</td>
</tr>
<tr>
<td></td>
<td>2.50</td>
<td>1.06</td>
<td>2.81</td>
<td>1.06</td>
</tr>
<tr>
<td>Middle-age</td>
<td>2.40</td>
<td>1.04</td>
<td>2.36</td>
<td>1.04</td>
</tr>
<tr>
<td></td>
<td>2.71</td>
<td>1.03</td>
<td>2.29</td>
<td>1.04</td>
</tr>
<tr>
<td>Older</td>
<td>2.13</td>
<td>1.09</td>
<td>2.34</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>2.25</td>
<td>1.09</td>
<td>1.97</td>
<td>1.09</td>
</tr>
</tbody>
</table>

Note: “NExpl” indicates no-AV-explanation condition; "BExpl" indicates AV-explanation-before-action condition; "AExpl" indicates AV-explanation-after-action condition; "PermReq" indicates request-for-permission condition; ‘M’ indicates mean; ‘SD’ indicates standard deviation.

6. Discussion

The goal of this research was to understand how driver’s age determines how effective AV explanations are at promoting driver’s trust and reducing driver’s anxiety. Results of this study highlight the importance of driver’s age in understanding the effects of AV explanations. For younger drivers the AV explanation before and after action conditions led to the highest trust, while the request-for-permission condition led to the lowest trust. For middle-age drivers the AV explanation before action condition had the highest trust, while the AV explanation after action condition had the lowest trust. For older drivers the request-for-permission and the AV explanation before action conditions both produced the highest trust. Contrary, the AV explanation-after-action condition resulted in the lowest trust.

There were also significant differences in drivers’ trust and anxiety between age groups across the AV explanation conditions. The AV-explanation-before-action condition produced the highest trust for all drivers regardless of age group. For the younger drivers, the AV-explanation-after-action condition was equally good, whereas for the older drivers the request-for-permission condition was equally good. In all, when available the AV explanation before action is the preferred approach. The request-for-permission approach seems to be best suited for older drivers. For older drivers, the request-for-permission approach produced the highest trust. The positive impact of the request-for-permission can be explained by prior research suggesting that older drivers struggle with handing over control of the driving [31]. The request-for-permission approach simply gives more of the control to the driver. Nonetheless, the AV-explanation-before-action condition produced similar benefits for older drivers as the permission condition. The request-for-permission approach, however, led to the lowest trust and highest anxiety for both the younger and the middle-age drivers. That being said, this does not explain why younger and middle-age drivers showed lower trust and higher anxiety compared to older drivers. Future research is needed to fully explore these findings.

The AV-explanation-after-action approach seems to be best suited for younger drivers. The AV-explanation-after-action approach led to the lowest trust for the middle-age and older drivers, along with the highest anxiety for the middle-age drivers. On the contrary, the AV-explanation-after-action approach led to the highest trust for younger drivers. Providing the explanation after the AV takes action could increase the uncertainty...
compared to the AV-explanation-before-action condition and force drivers to retrieve information and to understand why the AV took that action. To be clear, future research is needed to investigate this question.

The no-explanation approach produced mixed results for younger drivers. The no-explanation condition resulted in the most anxiety for younger drivers when compared to the other conditions (i.e., within age-group analysis). In addition, the no-explanation approach led to the highest anxiety for younger drivers when compared to middle-age and older drivers (i.e., among age-group analysis). This means that providing explanations regardless of timing could significantly decrease younger drivers’ anxiety. That being said, younger drivers’ trust in the AV was also least impacted by not having an explanation. We would expect trust and anxiety to be negatively related. Future research should be conducted to better understand the contradictory results for younger drivers.

6.1. Research Implications

Our results have several implications for research on age and driving automation. Our findings highlight the important role that a driver’s age has in understanding the impact of AV explanation on trust in driving automation. For example, if we assume that the request-for-permission condition represents heightened control, our findings support prior literature on age and driving control. We found that the request-for-permission condition led to the highest trust only for older drivers. This aligns with prior literature that suggests that older drivers have the greatest difficulty with the loss of driving control [31]. For older drivers, the request-for-permission condition actually helps to alleviate this issue. This and other findings highlight the need to account for drivers’ age when theorizing and designing AVs and their corresponding explanations.

This study has implications for theories related to drivers’ age and driving automation. Several of our findings did not align with the prior literature on drivers’ age and driving automation. There are at least three ways to view these findings that appear to run counter to what we might expect given the prior literature. One, the differences from our findings might be a result of the level of automation (i.e., level 3 vs. level 4 vs. level 5). Two, the impacts of a driver’s age are not uniform or linear but instead vary in a way that is hard to predict. For example, the assertion that younger drivers trust technology the most, followed by middle-age drivers, then older drivers is a uniform linear approach to predicting the impacts of drivers’ age on driving automation. Our results do not support this assertion; instead, we found that although a driver’s age is important, its effect is often difficult to predict. Third, both the level of automation and the driver’s age might have joint effects on driving automation outcomes such as trust.

Our results also contribute to the literature on socially inclusive AI. Many scholars have highlighted the problems of biased AI and the need to build an AI system that is more inclusive [16,17]. AI explainability has been shown to be important to the promotion of trust between humans and AI, yet to date little research has been conducted to understand how individual differences might help determine AI effectiveness. This study might be the first study to not only explore the impact of individual differences on AV explanations specifically, but also to explore the impact of individual differences on AI explanations generally. Results of our study clearly highlight that individual differences are important to understanding the effectiveness of AI explanations. We hesitate to generalize the results of this study or any one study to other settings or populations. However, we believe the results of this study do justify the need for future research to better understand how individual differences impact the effectiveness of AI explanations. In doing so, we take a step forward in designing AI that is more socially inclusive.

6.2. Design Implications
The findings in this study have several implications for AV design. First, AV explanations should be designed, in part, based on the driver’s age. Over and beyond this, our results provide guidelines for designing AVs. That being said, our findings apply across all age groups. For example, providing an AV explanation could be a universal approach to accommodating all age groups.

Second, there are important and meaningful differences across age groups. Based on our results, for older drivers the AV should be designed to ask for permission to take action before making any changes. On the contrary, for younger and middle-age drivers the AV should be designed to avoid this option because of the lower trust and higher anxiety introduced.

Additionally, this study focused on the AV’s domain, but our results can be applied to other domains that involve AI explanations. The implication in such areas is that explanation is a key factor that provides transparency and explainability. But to be inclusive, designers have to acknowledge that the user population consists of different age groups. This study indicates that age impacts the relationship of AI explanations with trust and anxiety. Designers should consider such differences and focus on decreasing anxiety and increasing trust in different age groups.

7. Conclusion

AV explanations are vital to promoting drivers’ trust and reducing their anxiety. Findings in this study highlight the importance of drivers’ age in understanding these effects. The implications for this study highlight the need for future research on AV explanations and the driver’s age. Implications for this study also provide opportunities for future research to build and expand on the ideas in this paper toward socially inclusive AI.

However, there are also limitations in this study that need to be addressed in the future research. First, although the experimental setting provides high internal validity, it has limitations with regard to external validity. For example, all participants in this study were recruited from a university-related subject pool. These individuals might be different with regard to their AV-related knowledge and experience when compared to others in the general population. In addition, participants might have engaged in hypothesis guessing and altered their responses based on what they thought the researcher desired. To be sure, we found no evidence of this in our study. Ultimately, future studies should be conducted in field settings to increase external validity. Second, the average level of trust is relatively high (i.e., 5.66 out of 7), and the anxiety level were low (i.e., 2.50 out of 7) in this study. However, all averages were typical to levels found in the prior study that examines the level 4 or 5 AVs (e.g., [60-62]).

Third, this study did not examine many other attributes associated with AV and AV explanations. These include AV driving behaviors, explanations related to the definition, generation, selection, and evaluation of alternative courses of action for the driver as well as the presentation of the explanations and the modality used to deliver the explanations [63]. For example, this study only examined auditory modality. Future studies should be conducted to examine these and other possible attributes associated with AV and its explanations. Future research might even focus on what an AV should and should not explain. In all, there is clearly more research needed in this new area.

Author Contributions: Conceptualization, L.P. R. and X.J.Y.; methodology, L.P. R., X.J.Y., and Q.Z.; validation, L.P. R and X.J.Y; formal analysis, L.P.R., X.J.Y., and Q.Z.; writing—original draft preparation, L.P.R., X.J.Y., and Q.Z.; writing—review and editing, L.P.R., X.J.Y., and Q.Z.; visualization, L.P.R., X.J.Y., and Q.Z.; supervision, L.P.R. and X.J.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by University of Michigan’s Mcity, grant number: 2017-2018

Conflicts of Interest: The authors declare no conflict of interest.
References


