


Article

User Education in Automated Driving: Owner's Manual and Interactive Tutorial Support Mental Model Formation and Human-Automation Interaction

Yannick Forster ^{1,2,*}, Sebastian Hergeth ¹, Frederik Naujoks ¹ , Josef Krems ² and Andreas Keinath ¹

¹ BMW Group, 80937 Munich, Germany; sebastian.hergeth@bmw.de (S.H.); frederik.naujoks@bmw.de (F.N.); andreas.keinath@bmw.de (A.K.)

² Department of Psychology, Chemnitz University of Technology, 09111 Chemnitz, Germany; josef.krems@psychologie.tu-chemnitz.de

* Correspondence: yannick.forster@bmw.de; Tel.: +49-89-382-21199

Received: 22 March 2019; Accepted: 15 April 2019; Published: 17 April 2019



Abstract: Automated driving systems (ADS) and a combination of these with advanced driver assistance systems (ADAS) will soon be available to a large consumer population. Apart from testing automated driving features and human-machine interfaces (HMI), the development and evaluation of training for interacting with driving automation has been largely neglected. The present work outlines the conceptual development of two possible approaches of user education which are the owner's manual and an interactive tutorial. These approaches are investigated by comparing them to a baseline consisting of generic information about the system function. Using a between-subjects design, $N = 24$ participants complete one training prior to interacting with the ADS HMI in a driving simulator. Results show that both the owner's manual and an interactive tutorial led to an increased understanding of driving automation systems as well as an increased interaction performance. This work contributes to method development for the evaluation of ADS by proposing two alternative approaches of user education and their implications for both application in realistic settings and HMI testing.

Keywords: automated driving; human-machine interface; user education; mental model formation; driving simulator study

1. Introduction

Automated driving development is flourishing at the present time. With rapidly increasing system functionalities, there is an urgent need to consider appropriate Human-machine interfaces (HMIs) to communicate system states via auditory, visual, and haptic elements and also to let the user interact with these systems through the same modalities. This proliferation of functionality and interaction possibilities raises the necessity for developing robust methods to evaluate such interfaces. Level 3 (L3) automated driving systems (ADS) take over longitudinal and lateral vehicle control and it is not the driver's responsibility anymore to monitor correct system functioning [1] as compared to level 2 (L2) driving automation. Through this transfer of driving task responsibility, the opportunity to engage in non-driving related tasks (NDRT) arises. Therefore, the driver's role shifts from monitoring correct system functioning in L2 to performing fallback performance in L3 in case the system function exceeds its operational design domain (ODD). The ODD is defined in the Society of Automotive Engineers (SAE) norm J3016R as "The specific conditions under which a given driving automation system or feature thereof is designed to function, including, but not limited to, driving modes" (p 12). Human factors research strives to accomplish the task of safer and more efficient human-automation interaction, where the first and most

important aspect that must be regarded is appropriate HMI design with a user-centered perspective. Following guidelines for proper interface design in automated driving (see e.g., [2]), developers can design HMIs that are simple, intuitive, and user-friendly. As well, one additional challenge for a successful market introduction is the development and deployment of user education with driving automation. Educating users is needed to support safe and efficient interaction when using driving automation [3]. It noted that through good interface design the demand for user education might be reduced to a certain degree, but most likely it cannot be eradicated. In the area of manual driving, there has been considerable interest in research for educating novice drivers in order to enhance road safety [4,5]. The body of knowledge in the automated driving domain, however, is rather scarce even though the National Highway Transportation Safety Administration's (NHTSA) automated vehicles policy [3] explicitly points towards the development of user education approaches. The importance of automation user education and the retrieval of manual operation skills in control transfer events, however, had been raised by Bainbridge [6] more than two decades ago. More recently, Endsley [7] described an impression of user education when first encountering a Tesla Model S equipped with autopilot. According to the author, there were brief verbal descriptions and after repeated questioning a rather unstructured familiarization drive was offered. Similarly, a report by Pradhan et al. [8] of a round table discussion on driver training showed that there is a pressing need to develop and evaluate user education for driving automation. The present work contributes to the challenge of conceptual development of user education procedures and ADS evaluation methodology development by outlining possible approaches for user education (i.e., owner's manual, interactive tutorial). The aim of the present study is to investigate the effects of different user education approaches on system understanding and interaction performance.

1.1. Background

Most research on automated driving has been directed at the assessment of transfer of control events (so called "take-over situations") from the automated vehicle to the human operator. It has revealed a variety of human factor issues such as fatigue [9], trust [10,11], mode awareness [12], and controllability [13]. The usability of L3 ADS HMIs, especially when coupled with driver assistance systems, is also an emerging topic [14,15]. Research on the usability of automated driving HMIs has shown that strong learning effects are at play when novice users interact with the driving automation systems for the first time [15]. At the same time, appropriate understanding of driving automation requires repeated interactions [16,17]. When it comes to using L3 ADS, users face complex system architectures that are communicated via graphical and auditory HMI elements. The difference between L3 ADS and lower levels of driving automation (i.e., L2 and lower) is that there are restrictions for system activation. Therefore, operators face different demands from driving automation depending on the level of automation. It depends on the level of automation [1], which and how many conditions have to be met when users initiate control transitions by themselves [18]. Since the supervision of correct functioning during L2 automation is the operator's responsibility, there is only a small number of necessary conditions from manual to L2 automated driving. With the shift of responsibility for the driving task from the driver to the ADS, the number of conditions to be met rises in L3 automation. These conditions for transitions between the levels of driving automation and the increasing complexity of input elements due to a rising number of available system functions pose a challenge to human operators when interacting with this technology.

One considerable aspect that arises between different transitions of control is the interference between driving and ADS operation. Users that are driving manually have to perform the entire dynamic driving task (DDT), for example, longitudinal and lateral vehicle guidance. At the same time, they have to successfully operate the automated vehicle's HMI to complete control transitions which require sensory processing, perception, response selection, and response execution [19]. These two tasks (i.e., DDT, operating HMI) are likely to interfere with each other. Thus, performance is impeded in both driving and human-automation interaction performance. Both activation performance (e.g., errors,

time on task) and driving performance (e.g., standard deviation of lateral position [20]) should be worse for untrained users as compared with more skilled users due to distraction effects (see e.g., [21,22]).

Until now, familiarization and user education in studies on automated driving evaluation have been rather unstructured and there is no commonly accepted standardization or best practice recommendation. On the one hand, there are some studies that only included a brief description of the automated driving function [23], while on the other hand, other studies included comprehensive familiarization drives with additional descriptions by the experimenters [24,25]. Comparing the automotive domain with research on automation in the aviation domain (see e.g., [26]) clearly shows that user education will be one of the inevitable challenges to face for successful introduction of this technology. Thus, there is a pressing need to develop user education approaches in order to enable novice users of driving automation to appropriately use this emerging technology. The following paragraphs outline considerations of two conceptual user education approaches and existing empirical evidence on an owner's manual (textual information) and an interactive tutorial. One important aspect that both approaches share is that they focus on building declarative knowledge rather than procedural knowledge [27]. Prior research on learning to interact with driving automation has shown that it takes considerably longer to build up declarative knowledge in the form of accurate mental models [17] as compared with procedural knowledge in the form of adequate interaction performance [15]. The two approaches explicitly target user education approaches supporting the formation of declarative knowledge.

1.2. Owner's Manual Based User Education

One traditional approach to educate users is the owner's manual [28], which, in principle, is available to all later users of driving automation systems. The owner's manual based approach can be characterized as passive and unguided learning based on a study by Wickens et al. [29]. Users of this technique only consume information without relating it to an applied context. In addition, the degree of guidance is low since learners decide for themselves how much they would read and even decide on the order of information they would process. Owner's manuals, if designed appropriately, do not convey erroneous operating paths but rather focus on correct ways of operating a certain technology. Therefore, it does not provide users with conflicting information and instead focuses on appropriate interaction. Owner's manuals are often designed such that they do not support users by providing procedural information on how to solve a task or problem when using a technology, but rather support users by providing knowledge and skills that they have to transfer to the task later on [30]. It follows that in an owner's manual, HMI elements are presented separately without relating different operating components to graphical elements or auditory elements. Therefore, learners need to transfer knowledge gained from abstract descriptions of HMI elements to the applied context in the vehicle interior. Additionally, due to the dominance of textual information, users might get lost in the complexity of the information [29]. A potential drawback of owner's manuals is that the information on automated driving composes only a small portion of an enormous amount of text that can be comprised of up to 500 pages. Therefore, it might be challenging for users of driving automation to ensure that the information is not lost within other components of an owner's manual such as technical data, maintenance, and repair information.

Before encountering driving automation, the research on owner's manuals or textual information showed that they largely affect the behavior and attitude of users. When users of driving automation are provided with varying descriptions of degrees of completeness, correctness, or reliability, these largely influence the evolution of mental models, trust, and acceptance [31–34]. While most prior approaches of providing information have focused on trust and acceptance-related issues, the aim of the present study is to investigate building declarative knowledge as assessed through mental models and its association with the impact on interaction behavior (e.g., accuracy, speed) with driving automation.

1.3. Tutorial Based User Education

There are many different possible designs of interactive tutorials for user education. One design, the click dummy or minimum viable product version, is intentionally not considered here for two reasons. Using this version users freely try interacting with the system before actually proceeding to the respective application in the simulator or on the real road. The first reason for not considering this is that such an approach would rather build procedural knowledge through trial and error without the users being provided with appropriate descriptions of system functionality. Secondly, results from Forster et al. [15] have already shown how interactions with driving automation without prior declarative knowledge evolve over time. Another imaginable approach is the education during interaction approach. Such procedures would lead the user step by step through the operating process by giving explanations and providing feedback. While such a procedure bears the advantage, for example, in rental car scenarios where users want to start driving and use the function right away, there are also problems with this approach. The reason the user uses automated driving is to be relieved of the driving task and engage in different nondriving related tasks (NDRT). Comprehensive explanations by a system avatar [35] might take a long time and users want the function to execute the task right away. Therefore, we focus on an approach to educate users prior to first interaction. It might also be possible in car rental scenarios that by booking a tablet or computer, a brief tutorial could be completed. This procedure of educating users prior to first system encounters is also supported by inferences from aviation where only trained individuals operate automated systems and they are not trained during use. Therefore, the present work considers an interactive tutorial approach that is characterized by active and guided learning [29] in the form of a user quiz consisting of questions on how to operate the driving automation system in several use cases [18]. This tutorial is completed prior to the first interaction with driving automation. With the possibility of giving wrong answers, the tutorial supports learning from erroneous trials. If users take erroneous choices in a tutorial-based approach, these errors should not occur in subsequent interactions in the vehicle. Tutorials are characterized by a high degree of experimental control since users can only proceed to using the ADS in the vehicle after they have successfully completed the assigned questions. Thus, after an interactive tutorial, all users should have the same degree of system understanding before their first encounter with the driving automation system [5]. Another characteristic of an interactive tutorial is that it is somewhat focused on specific tasks that users can subsequently face as compared with abstract system descriptions in an owner's manuals [30]. Therefore, tutorials provide a meaningful context in which a driving automation system will be used and it is not necessary to transfer knowledge from abstract descriptions to specific use cases [6]. Interactive tutorial approaches offer the possibility to provide the user a close-to-reality impression of the HMI, providing a higher degree of immersion [36] as compared with owner's manual-based approaches. Designers can implement the HMI components (i.e., operating elements, buttons, graphical interface) in a way that it resembles later interior integration. As an active learning approach, the design makes the availability of erroneous answers indispensable. This can be considered a downside of this approach since learners might get confused from the sheer amount of wrong answers in relation to correct answers (e.g., 1:3 ratio). This becomes even more crucial considering that in order to foster active learning the distractors should be designed in a way that they are not too obviously wrong. As users do not have any knowledge of the system beforehand, it is likely that they frequently select wrong answers and become frustrated from the trial and error approach. Therefore, designers of such tutorials should give careful attention to the communication of the outcome of erroneous answers so that they do not lead to reactance from the learner [37].

Research on interactive user education approaches has led to positive results in educating novice drivers [38,39]. Considering that these systems are not yet available commercially and users are unfamiliar with driving automation in general, research on efforts towards educating users with ADS need to be undertaken. In the automated driving domain, research on trust and controllability regarding take-over situations provided the first evidence that the familiarization of users with the HMI by means of descriptions, HMI presentations or actual experience, can have beneficial effects [11,40].

However, conceptual approaches for teaching users how to interact with driving automation technology are still scarce. Table 1 compares the owner's manual and tutorial approaches with respect to education design characteristics.

Table 1. Summary of characteristics of manual and tutorial based user education approaches.

Characteristic	Source	Owner's Manual	Interactive Tutorial
Type of learning	[29]	Passive learning	Active learning
	[5,29]	Unguided/uncontrolled learning	Guided learning/experimental control
Focus of education	[30]	System focused	Use case focused
Type of information	N/A	Correct information only	Presence of distracting information
Degree of realism	[36]	Abstract representation of HM	Immersive representation of HMI

1.4. Aims and Objectives

The present work has two objectives. The first objective is to determine if an experimental manipulation, with respect to user education approaches consisting of an owner's manual or an interactive tutorial as compared with generic information about the system function, would result in more accurate mental models and superior operator performance in automated driving. Secondly, it focuses on a proof-of-concept for educating users in a risk-free environment on a desktop screen before they might be exposed to the task in a real vehicle on the road or in a driving simulator.

This study does not claim to explicitly investigate systematic differences between the two proposed user education approaches. Therefore, the analysis primarily focuses on evaluating each approach as compared with a baseline. Subsequently, it is the task of future basic research to discover more about the specific mechanisms of learning that underlie each of the two approaches.

On the basis of the considerations outlined above, we hypothesize that both the owner's manual and the tutorial would lead to more accurate mental models regarding the HMI for L2 and L3 automation as compared with mere generic information about the system's functionality. No hypothesis can be stated for a difference between the two treatments since the conceptual difference is based on a combination of education characteristics (see Table 1). Thus, potential differences cannot be traced back to a single cause. Nevertheless, an explorative approach has been undertaken by comparing mental model results for the two user education approaches.

Since we expect that the two treatments will support the evolution of accurate mental models of the HMI and the system, interaction performance in both treatment conditions should be enhanced compared to mere generic information. Similar to the experimental variations above, we follow an explorative approach when comparing the two user education procedures. In addition, due to the task-dependent characteristics of learning [41], we investigate whether the two user education approaches work as expected in all use cases considered in the study, or whether specific use case strengths and weaknesses are present.

2. Materials and Methods

2.1. Driving Simulation and Automated Driving System

The study was conducted in a fix-base driving simulator (see Figure 1). The vehicle mockup was identical to a BMW 5 series with automatic transmission and contained all necessary instrumentation. The front channels provided a field of view of 220° and were displayed through five front projectors. The rear view was displayed through three LED screens that were placed behind the vehicle.

The L2 and L3 driving automation carried out longitudinal and lateral vehicle guidance as soon as the driver activated the respective function. The L2 automation let the driver take his/her hands off the steering wheel for 15 seconds before displaying a hands-on request (HOR). The L3 ADS could execute independent lane change maneuvers such as decelerating due to slower vehicles ahead or pulling back to the right lane. There was no restriction in the adjustment of current velocity for L2 automation. At activation, the current velocity was the set speed of the system. For L3 automation, the speed could only be set to speeds slower than 130 km/h. If set to a higher speed, the system would suggest a transition to L2 driving automation. In all cases the set speed of the L3 automation was 130 km/h. The vehicle also included adaptive cruise control (ACC) which executed the longitudinal vehicle guidance as well as the speed limiter. In reference to SAE J3016R, the L3 automation is considered an ADS while the L2 automation is considered a driver support system.



Figure 1. Static driving simulator mockup used in the present study.

2.2. Study Design and Procedure

The study employed a single-factorial between-subjects design. The between-subjects factor, “education”, had three levels which were baseline information (BL), owner’s manual (ML), and interactive tutorial (TL). Participants were randomly assigned to the between-subjects factor. This work reported the findings of the mental model formation after the different treatments had been provided and of the interaction performance in a subsequent block of different transitions of control.

Upon arrival, the experimenter welcomed participants and informed consent was obtained. First, the experimenter provided a brief explanation of the purpose of the study. Then, the experimenter outlined which education approach participants would receive and gave a standardized instruction for each condition. The maximum time to complete the educational procedure was 10 minutes in each condition. Having finished the educational procedure, participants completed a mental model questionnaire (see section on the dependent variables) and were led to the vehicle mockup. Before proceeding to the experimental drive and operating the driving automation, participants completed a short manual familiarization drive to accustom themselves with handling the simulator. Since instructions for the use cases were given through recorded samples, participants were introduced to a sample telling them to change lanes in the familiarization trial. During the experimental drive, participants had to complete two blocks of six interactions each (see section on use cases). The use cases specific self-report data were collected after each interaction during the drive. Participants could not anticipate the instruction of use cases since the experimenter deliberately waited at least 30 seconds from the end of the use case specific inquiry before triggering the instruction to follow. Additionally, upcoming transitions could not have been indicated by external cues such as highway intersections or slow cars ahead. Having finished one block, the participants filled out the block inquiry at standstill on the right shoulder. Traffic density on the three-lane highway was low to medium. The drive lasted approximately 20 minutes. Table 2 schematically depicts the experimental procedure.

Table 2. Schematic outline of experimental procedure.

User Education		Preparation		Experimental Drive	
Education application	Mental Model questionnaire	Familiarization drive (manual)	Instructions	Control transitions	Inquiry

2.3. Human-Machine Interface

The visual HMI showed the vehicle and its surroundings in L3 automation and this was displayed on the instrument cluster. While the L2 automation was engaged, icons on the left side of the instrument cluster depicted lateral and longitudinal vehicle guidance. There was no display of the vehicle and its surroundings during active L2 automation. Active L2 and L3 automation were colored differently. Similar solutions for visual automated vehicle HMIs have been proposed by Forster et al. [42] and outlined in Manca et al. [43]. Thus, the present conceptual approach constituted a representative solution for an automated vehicle HMI. Generally, the HMI is comprised of steering assistance and adaptive cruise control interface features [44]. Referring to design considerations by Naujoks et al. [45], the present HMI also contained information about the main state such as lateral and longitudinal guidance, current velocity, and set speed information. It also followed the design principle of redundancy by showing vehicle and surroundings plus icons indicating an active L3 ADS. A countdown provided information about the duration of the automated driving availability during active functioning. The system states were also different if the system was not activated and not available versus not activated and available. This was indicated by the L3 ADS availability icon in the instrument cluster. There were buttons on the left side of the steering wheel for both the ADS and the driver support functions for the participants to initiate the respective control transition and also to deactivate the function to return to manual driving. The buttons were not illuminated in any automated mode. Participants could switch between different driver support functions (i.e., L2 automation, adaptive cruise control, speed limiter) with one button. The input device included three more buttons (e.g., set speed). For adjusting cruising speed in both L2 and L3, there was a rocker switch among the buttons. The buttons were easily reachable for participants using their left thumb. No restrictions were present for the activation of the L2 automation. In contrast, participants had to meet certain conditions for completing a transition to L3 automation [1]. These conditions were: (1) availability of the ADS, (2) velocity below 130 km/h, and (3) lane keeping within a certain lateral margin. Besides pressing the L3 automation button with both hands on the steering wheel, participants could deactivate the ADS and drive manually by accelerating, braking, or applying a small force to the steering wheel. Participants could only deactivate the L2 automation by pressing the “driver support” button or braking. Accelerating temporarily led to a L1 automation state (i.e., lateral support only). Likewise, applying a force on the steering wheel only led to a temporary L1 automation state of longitudinal support. These temporary states were active until the drivers let go of the respective input.

2.4. Use Cases

The present experiment included driver initiated transitions between manual, L2, and L3 automated driving [18]. A total of six use cases per experimental block were possible from all combinations of upward and downward transitions. Transitions to manual were not regarded in this context since there is evidence that these UCs do not require any learning [15]. Table 3 outlines use cases with transition type, initiation automation level, target automation level, and number of the respective use case. There were six possible sequences of use cases for each block, to which participants were randomly assigned. Therefore, sequential effects could be ruled out. The upward transitions of automated driving were motivated by the driver wanting to be relieved of the driving task itself (UC1, UC2) or the supervising task (UC 3). The downward transition from L3 to L2 automation could be motivated by the user’s wish to deactivate certain functionalities or drive at a higher speed [18].

Table 3. Overview of use cases for one experimental block.

Transition Type	Scenario	Automation Level at UC Initiation	Automation Target Level	Use Case Number
Upward transition	Activation L3	L0	L3	1
	Activation L3	L2	L3	3
	Activation L2	L0	L2	2
Downward transition	Deactivation L3	L3	L2	4

The experimenter triggered standardized samples for the onset of each use case that were recorded prior to the experiment. There was a 10-second delay from the onset of the standardized sample instruction for UC1 and UC3 (transitions to L3) and the availability of the ADS. Therefore, it was not possible for participants to activate the L3 ADS within these 10 seconds. The instruction for transitions to L3 thus included an additional indication to activate the ADS (i.e., “(. . .) as soon as it is available”).

2.5. User Education Approaches

2.5.1. Baseline Information

Baseline information conveyed merely generic information about the equipped driving automation as reflected in SAE J3016R [1]. Participants received a description that the simulated vehicle incorporated a L2 and L3 driving automation. Depending on the situation, both functions would break or accelerate the vehicle and keep distance to the vehicle ahead. In addition, both functions would execute lateral vehicle guidance. The L3 ADS would execute lane changes if necessary. With the L3 ADS, certain defined nondriving related tasks could be engaged in and the driver did not need to continuously monitor the traffic conditions. However, the driver needed to resume vehicle guidance upon system notice. In contrast, with the L2 automation, the driver had to continuously monitor the traffic conditions. The baseline information preceded both the owner’s manual and the interactive tutorial.

2.5.2. Owner’s Manual

The owner’s manual was designed in accordance to existing BMW manuals which are characterized by short sentences, and listings if possible. They also incorporated graphical elements of buttons when these were referenced. The document was clearly structured by inserting headings and subheadings. If an information was especially important, it was shown in a separate textbox. The owner’s manual was a four-page DIN-A4 text document that users had to read prior to system use. It contained all relevant icons in the instrument cluster and buttons on the operating element. The owner’s manual was structured covering the L2 and L3 relevant HMI elements in successive steps. It provided information about activations, deactivations, and the ODD of the respective functions. Figure 2 shows an excerpt of the owner’s manual (originally presented in German language). Note that icons had been disguised due to confidentiality and are represented as grey squares. Specific function names have been replaced by “L3 ADS” and “L2 automation” for the same reason.

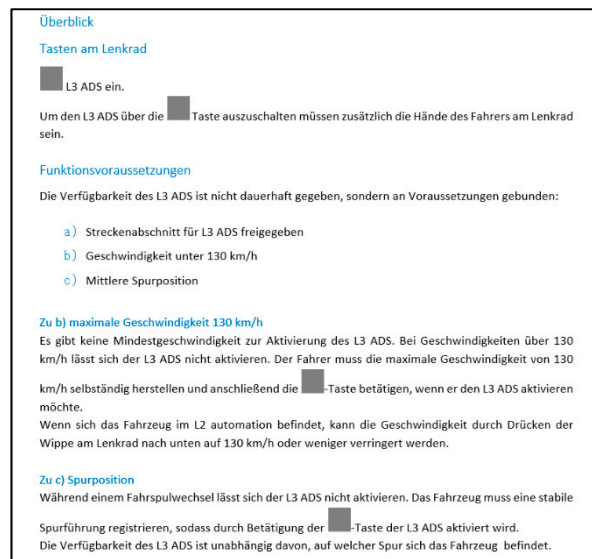


Figure 2. Excerpt from owner’s manual. Note that icons had been disguised due to confidentiality and are represented as black squares. Specific function names have been replaced by “L3 ADS” and “L2 automation”.

2.5.3. Interactive Tutorial

The interactive tutorial was designed using Microsoft PowerPoint. The tutorial showed the operating element (i.e., buttons on the left of the steering wheel), the instrument cluster with the present system state, and a driver’s view of the vehicle interior. Next, there was a description of the task including the current system state, target system state, and target system name. The four answers included the respective buttons and/or icons that could appear in the instrument cluster. Participants could select an answer by clicking on the respective box. If a wrong answer was given, a box appeared in the foreground explaining what would happen and that the answer was not correct and redirected the participant to the task. If the answer was correct, the same box would appear telling the test taker what HMI elements would appear and that the answer was correct. The tutorial was finished when all tasks were completed successfully. In the tutorial, each task had only one correct answer. Incorrect answers were derived from an inquiry of a prior experiment [15]. Table 4 provides an overview of the seven questions in regard to UCs, transition type, and restriction as reported in SAE J3016R [1]. Figure 3 shows a screenshot of tutorial question 1. Note that buttons on the steering wheel and icons in the answer options were disguised due to confidentiality. The function name of the L2 automation has been replaced by “L2 automation” for the same reason.

Table 4. Overview of tutorial questions with corresponding UC, transition, and type of restriction for the transition.

Tutorial Question	Relevant UC	Relevant Transition	Restriction as of SAE J3016R [1]
1	2	L0 → L2	None
2	1	L0 → L3	Availability restriction
3	1	L0 → L3	Speed restriction
4	1	L0 → L3	Lateral guidance restriction
5	3	L2 → L3	Availability restriction
6	3	L2 → L3	Speed restriction
7	4	L3 → L2	None



Figure 3. Layout of interactive tutorial with operating element (upper left), instrument cluster (upper middle), driver’s view (upper right), user task, and four alternative answers.

2.6. Dependent Variables

The mental model questionnaire developed by Beggiato et al. [16] was adapted for L2 and L3 driving after automation. In a prior experiment, Forster et al. [17] used this adaptation to evaluate the development of mental models in interaction with driving automation. Participants completed the questionnaire after they had received the treatment and then after the six interactions. However, the present work only reports results from the mental model questionnaire before the drive. The mental model questionnaire included 11 items on a seven-point Likert scale from 1 (“strongly disagree”) to 7 (“strongly agree”) for both the L2 and L3 automation. Two items covered the participant’s understanding of HMI in general, four items covered the understanding of the HMI with a focus on transition restrictions [1], and five items served as distractors. In the present study, the four items covering HMI and transition restrictions were analyzed (see Table 5). Items were derived to specifically detect differences between the L2 and L3. Thus, for each item of interest, the opposite end of the scale was correct for the two levels of automation. If “strongly disagree” was the correct answer for L2, “strongly agree” was correct for L3 and vice versa. In the present experiment, the correct answer for all four items of relevance was 1 (“strongly disagree”) for L2 and 7 (“strongly agree”) for L3 automation.

Table 5. Mental Model questionnaire with item number and item wording.

Number	Wording
1	There is a speed limitation that must not be exceeded to activate the system
2	Lane keeping is relevant for the system activation
3	The system displays availability to the driver
4	There are road sections where the system is not available

Interaction performance was assessed by means of a five-point rating scale, which is shown in Table 6. The interaction performance was rated during the experiment right after the participant had completed the respective use case. To counteract potential bias in experimenter ratings, specific behavioral observations were assigned to each category per use case to ensure objectivity of the rater. Each erroneous input that did not result in the required control transition was counted as an error. There was also the possibility to push a correct button (e.g., L3 ADS) without the transition being executed by the system since conditions were not met. Such interactions were also counted as errors. Moreover, additional strong lane deviations due to extensive correction of erroneous inputs and glance allocation away from the road towards the in-vehicle HMI led to assignments of category four (i.e., massive errors). For simplification purposes, Table 6 only provides the generic descriptions.

2.7. Sample Characteristics

The sample consisted of $N = 24$ participants (6 female, 18 male) with $n = 8$ participants assigned randomly to each experimental condition. No dropouts were recorded. Mean age of the sample was 33.96 years ($SD = 12.99$, $max = 62$, $min = 20$). All participants were BMW Group employees, held a German driver's license, and had normal or corrected to normal vision.

Table 6. Experimenter rating with label and description.

Category	Value	Description
No Problem	1	Quick processing
Hesitation	2	Independent solution without errors But: hesitation, very conscious operating and full concentration
Minor errors	3	Independent solution without or with minor errors which were corrected confidently But: longer pauses for reflection Evaluation of potential operating steps
Massive errors	4	One or multiple errors Clearly impaired operation flow Excessive correction of errors No help of experimenter necessary
Help of experimenter	5	Multiple errors Massive errors require to restart task, Help of experimenter necessary

3. Results

No missing data were recorded for self-report or observational measures. The present study included a small number of participants per between-subjects factor, and therefore the use of estimation methods that require assumptions on sample distributions are problematic [46]. Therefore, we applied bootstrapping for analyzing the data since this procedure has been found to be robust in cases of heterogeneity and non-normality [47]. We calculated 95% confidence intervals (CI) for comparisons between the experimental conditions as a state-of-the-art inferential statistics technique [48,49]. We report means, upper, and lower bounds of 95% CIs in Table 7 (mental model L2 automation), Table 8 (mental model L3 automation) and Table 9 (interaction performance). Inferential p -values are reported after the rule described in Cumming and Finch [50]. Bootstrap analysis was performed using Matlab R2015b. The code of the bootstrapping procedure to generate $n = 10,000$ data set replicates is supported in Appendix A.

3.1. Mental Model Questionnaire

Results for the mental model are provided separately for the L2 (see Table 7 and Figure 4) and L3 automation (see Table 8 and Figure 5). Mental models for the L2 automation showed that knowledge about the system in item 1 (speed limitation) and item 3 (availability display) was enhanced in both treatment conditions in a comparable manner. The results revealed bimodal distributions in many cases (e.g., BL item 1 and item 4, Figure 4). However, the means of these distributions fell right in between the two distribution modes. Consequently, a mean value on the center of the scale as reported in the tables must regard the respective distribution in the according figure. While participants in the baseline condition strongly assumed a speed limitation and availability display for the L2 automation, participants in the treatment conditions were surer that there was no speed limitation ($p < 0.05$) and did not as strongly assume an availability display ($p < 0.01$). A large overlap between the conditions was present for item 2 (lane keeping) as indicated by the CIs (see Table 7). On a descriptive level, participants in the baseline condition showed a trend towards the most correct answers. Thus, no statistically differences in regard to lane keeping relevance for L2 automation were present. Item 4 (road section

availability) showed that participants in the baseline condition strongly assumed a restricted ODD, while the distribution for the tutorial condition was spread almost over the entire seven-point Likert scale, and participants in the owner’s manual condition showed the best answers (i.e., no restriction of ODD). The difference between baseline and tutorial was not statistically significant. The owner’s manual condition differed significantly from both the baseline ($p < 0.01$) and the tutorial ($p < 0.05$).

Results for the mental model L3 automation showed that the participants in the baseline condition had to guess whether there was a speed limitation or not (item 1). In comparison, mental models for the owner’s manual ($p < 0.05$) and tutorial ($p < 0.05$) differed significantly from the baseline condition with most participants indicating a strong approval for the statement. Participant understanding of lane keeping relevance, (item 2) in the tutorial condition, was more accurate as compared to the baseline ($p < 0.01$). On a descriptive level, the owner’s manual condition was superior to the baseline, but inferior to the tutorial condition, however, due to a large margin of error, no statistical significance can be inferred. In regard to the availability display (item 3), participants in all three experimental conditions indicated strong approval for the existence of an availability display for the L3 automation. Thus, CIs overlapped to a large degree and no significant differences are present. Answers for item 4 (road section availability) revealed that participants in the baseline were indifferent whether a restricted ODD is present on the highway or not, which is reflected in a mean of 3.87 (see Table 7). Compared to the baseline, participants in the owner’s manual condition indicated existence of a restricted ODD ($p < 0.01$). Participants in the tutorial condition were more accurate on a descriptive level as compared with the baseline, without reaching statistical significance.

Table 7. Means and 95% CI upper and lower bounds of bootstrapped mental model distributions for the four relevant items on L2 automation by the three experimental conditions.

	Speed Limitation		Lateral Guidance		Availability Display		Road Section	
	Distr. Mean	95% CI Bounds	Distr. Mean	95% CI Bounds	Distr. Mean	95% CI Bounds	Distr. Mean	95% CI Bounds
Baseline	5.00	[4.00–5.8]	3.75	[2.88–4.2]	6.75	[6.50–7.0]	5.63	[4.63–6.0]
Manual	2.61	[1.13–4.13]	3.88	[2.38–5.50]	4.40	[2.88–5.88]	2.37	[1.38–3.63]
Tutorial	2.62	[1.63–3.75]	5.01	[3.88–6.00]	4.76	[3.25–6.13]	4.37	[3.50–5.25]

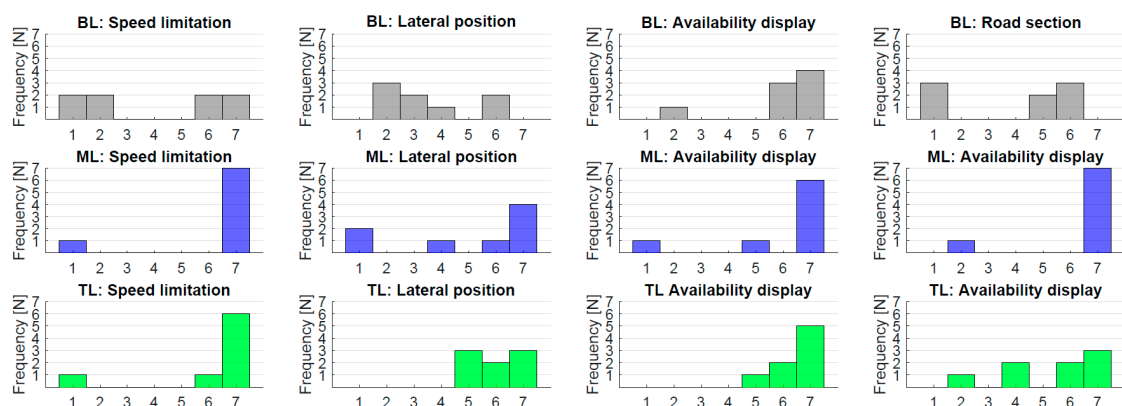


Figure 4. Frequency [n] of L3 ADS mental model ratings [1–7] for the four relevant items (columns) by experimental condition (rows 1–3: baseline grey, owner’s manual blue, tutorial green).

Table 8. Means and 95% CI upper and lower bounds of bootstrapped mental model distributions for the four relevant items on L3 automation by the three experimental conditions.

	Speed Limitation		Lateral Guidance		Availability Display		Road Section	
	Distr. Mean	95% CI Bounds	Distr. Mean	95% CI Bounds	Distr. Mean	95% CI Bounds	Distr. Mean	95% CI Bounds
Baseline	4.00	[2.50–5.38]	3.50	[2.63–4.50]	6.01	[5.00–6.75]	3.87	[2.63–5.13]
Manual	6.27	[4.75–7.00]	5.01	[3.50–6.38]	5.99	[4.75–7.00]	6.38	[5.13–7.00]
Tutorial	6.12	[4.75–7.00]	6.00	[5.50–6.50]	6.50	[6.21–6.87]	5.37	[4.38–6.38]

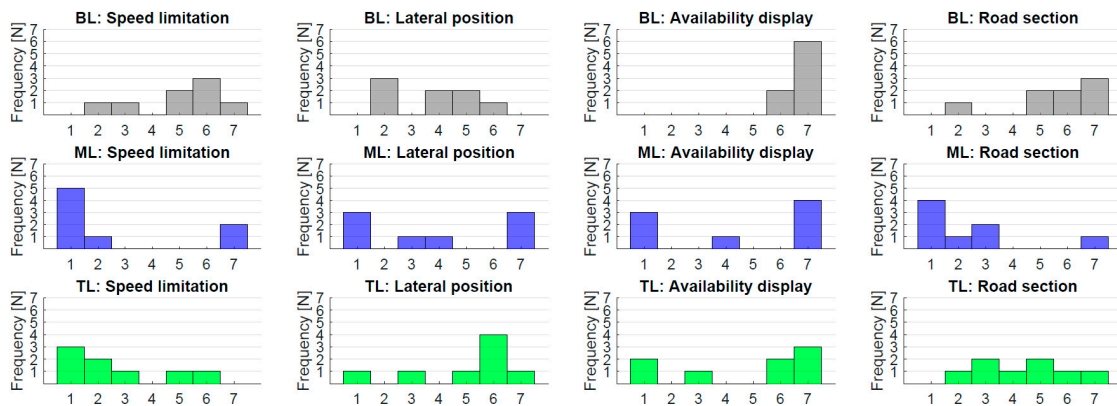


Figure 5. Frequency [n] of L2 automation mental model ratings [1–7] for the four relevant items (columns) by experimental condition (rows 1–3: baseline grey, owner’s manual blue, tutorial green).

3.2. Interaction Performance Ratings

Results for observational measures (see Table 9 and Figure 6) in UC1 revealed that both the owner’s manual and the tutorial led to superior performance in activation of the L3 ADS as compared with the baseline. In addition, the participants that had read the owner’s manual showed better performance than the participants that had received the tutorial. Activation of L2 (UC2) showed that both the owner’s manual and the tutorial significantly increased performance with no difference between the two conditions. Transitions from L2 to L3 (UC3) showed that best performance was achieved in the tutorial condition where participants performed significantly better than in the baseline and in the owner’s manual condition. Baseline and manual conditions did not differ significantly. Eventually, no differences among all three conditions were observed for the transition from L3 to L2 automated driving (UC4).

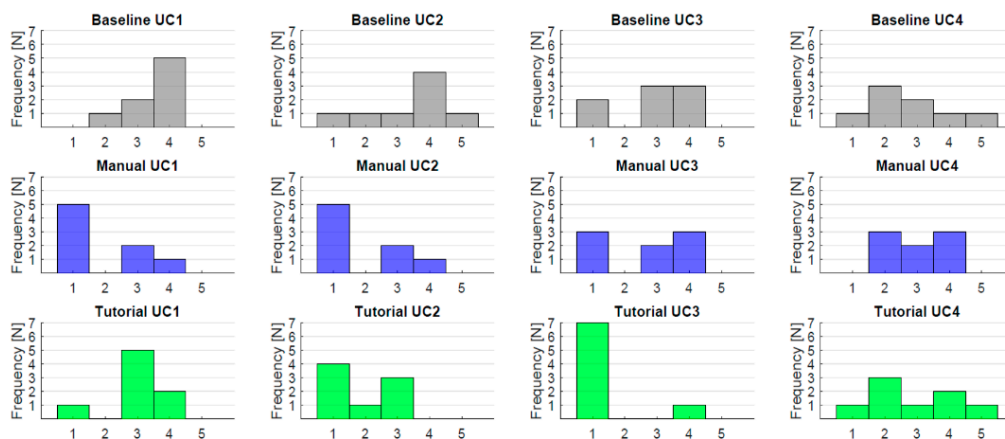


Figure 6. Frequency [n] of experimenter ratings [1–5] in each use case (columns) by experimental condition (rows 1–3: baseline grey, manual blue, tutorial green).

Table 9. Means and 95% CI upper and lower bounds of bootstrapped experimenter rating distributions by the three experimental conditions for the four UCs.

	UC1		UC2		UC3		UC4	
	Distr. Mean	95% CI Bounds	Distr. Mean	95% CI Bounds	Distr. Mean	95% CI Bounds	Distr. Mean	95% CI Bounds
Baseline	3.50	[3.00–3.88]	3.38	[2.63–4.00]	2.87	[2.13–3.50]	2.75	[2.13–3.50]
Manual	1.88	[1.25–2.63]	1.87	[1.25–2.63]	2.63	[1.87–3.38]	3.00	[2.50–3.50]
Tutorial	3.00	[2.50–3.50]	1.87	[1.38–2.38]	1.38	[1.00–2.13]	2.87	[2.13–3.63]

4. Discussion

The present driving simulator study investigated the effects of two different user education approaches which were (a) the owner’s manual, and (b) an interactive tutorial on mental model formation and human-automation interaction, in comparison to (c) a generic functionality description [1]. We hypothesized that participants in the owner’s manual and interactive tutorial conditions would have a better understanding and show superior operator behavior as compared with the group that had only received a generic description. In a between-subjects design, $N = 24$ participants were randomly assigned to one of three treatment groups (i.e., baseline, owner’s manual, tutorial). Having finished the respective education, the mental model was assessed and subsequently, drivers completed several transitions between manual, L2, and L3 automated driving. Our results showed that both owner’s manual and tutorial led to more accurate mental models and improved interaction performance as compared with the baseline. In addition, a use case specific trend was observed since only trained participants showed better performance than untrained participants in first-order use cases (i.e., transitions from manual to L2/L3) as compared with second-order use cases (i.e., transitions between two levels of automation). The results suggest that user education approaches, as conceptualized in the present study, significantly add to a better understanding of automated vehicle HMIs and this knowledge is also transferable to user’s interactions with the driving automation system. The following paragraphs discuss findings and limitations of the present results and derive practical implications.

4.1. Mental Model

Interpretation of the results of mental model formation must consider that the items focused on system restrictions that existed for L3 and not for L2. Both user education approaches communicated this difference by introducing the restrictions for the L3 function while not mentioning these for the L2 function. As a consequence, the only way to identify the non-existence of the restrictions for L2 automation was by inferring this from the knowledge that it was not mentioned in the respective treatment.

Results for L2 mental models (Figure 4 and Table 7) showed a trend towards more correct answers in the treatment conditions, especially for the speed (item 1) and availability restriction (item 3 and item 4). Both approaches were successful in conveying that the ODD restrictions do not exist for L2 automation. No difference was apparent for the lateral guidance item (item 2). This result might be due to the fact that understanding and terminology of lane keeping is difficult for users, and it also corroborates prior evidence on mental model formation for the aspect of lane keeping [17]. A slightly different picture emerged for the formation of mental models of L3 automation where ODD restrictions for activating driving automation are present. While both user education approaches could deliver first evidence for conveying speed (item 1) and availability of road sections (item 4), there was no difference apparent for an availability display (Figure 5 and Table 8). However, this result is due to user expectations in the baseline. This group showed a strong trend towards assuming that an availability display existed. Consequently, there is no room for additional improvement here. Therefore, the diagnostic benefit of this item for L3 mental model assessment is rather small. What has not been assessed, however, is the degree of users’ certainty when answering mental model items. It is speculated that although the present results did not show a difference between the conditions,

participants in the two treatment groups should be more certain about their indication, while baseline participants might rather have indicated their expectation of the HMI [17]. Concerning lateral position ODD restriction, the present study found a descriptive trend that users in the owner's manual condition gave more correct answers as compared with the baseline. A strong shift towards the correct answer was observed for the tutorial. This finding is explained through the use case focus of the tutorial, since one of seven questions specifically addressed this restriction. In comparison, the owner's manual only mentioned the importance of lane keeping within a larger set of paragraphs. This result supports considerations based on Wickens et al. [29], who found that use case specific user education is more beneficial than system specific where users can get lost in the complexity of the information.

Mental model findings for L2 and L3 automation highlight the importance of user education for forming an accurate understanding of the automated vehicle HMIs. There was no clear superiority of either of the education approaches. It seems like when it comes to abstract questions (e.g., item 4, "there are road sections where the system is not available"), owner's manual approaches are more beneficial than tutorial approaches. On the other hand, if items are not abstract but rather align closely with a scenario that had been taught in the tutorial (e.g., item 2 and tutorial question 4), the results turned out to favor the latter.

The results of the mental model questionnaire might also hold important implications for mode awareness issues. Feldhuetter et al. [12] among others have shown that users frequently cannot discriminate between L2 and L3 automation based on HMI and functionality alone. Therefore, we assume that users also have a better understanding of current system states when being supported with knowledge on how to complete control transitions. From their own actions they should be able to infer the correct system state and not confuse L2 and L3 as frequently as without information on control transitions. However, to verify this assumption future research is necessary.

4.2. Interaction Performance

The present findings support the a priori stated hypothesis that user education improves interaction performance with driving automation. However, this finding does not apply to all use cases, but is limited to a specific pattern (see Figure 6 and Table 9). The owner's manual provided the most support when executing the transition from manual to L3 automation. In contrast, the interactive tutorial delivered the best interaction results when participants were driving with L2 automation before performing the transition to L3. Both education approaches showed the same superiority over the baseline information for the transition from manual to L2 automated mode. However, this superiority was not present for either condition when participants were driving in L3 automated mode before performing the transition to L2. Therefore, perhaps the target system state is not only important for interaction performance, but also the preceding system state. In the present study, the owner's manual approach did not provide any better results as compared with the baseline for transitions between automation levels. UC3 and UC4 require preceding transitions to the respective level (i.e., L2 for UC3 and L3 for UC4). Therefore, learning mechanisms from the preceding interactions come into play. These might interfere with the information that users have gained from either the manual or the tutorial. Similarly, Forster et al. [15] describe learning effects between different transitions of automation levels when evaluating HMIs for driving automation. Referring to results from the mental model, interaction measures support the reasoning that user education is especially beneficial in first-order use cases. However, the procedural knowledge about the automated driving HMI that is gained through interaction significantly influences subsequent trials. Once there was a successful interaction with the driving automation (which was always the case in the present study), this experience more strongly influenced subsequent behavior than declarative knowledge from the prior treatment. On the other hand, this reasoning is contradicted by the results of extremely good performance in UC3 in the tutorial group in comparison with the owner's manual and the baseline groups. This data show barely any variation of the experimenter rating. It might be that this observation is an artifact of the ordinal nature of the experimenter rating itself. The step from the category at the top end of the scale

(i.e., category “no problem”) to the middle of the scale (i.e., category “minor problems”) is considerably small. As a consequence, there is a need for future research that applies interval scale observational measures such as error rates and time on task [19] in order to more closely determine the differences between first and second-order interactions.

Even though experience with the system seemed to largely influence interaction behavior, the present results emphasize the important role of user education for successful interaction with driving automation. It seems that an interplay of prior education and learning from these first-order interactions is present. Therefore, vehicle manufacturers and employers of driving automation might consider supporting users with information material prior to first use, as these are influential for subsequent interactions. At the same time, automated driving development needs to consider that interface design itself results in learning that adds to the content from tutorials and owner’s manuals. Thus, even when users activate L3, they learn how to activate parts of the L2 automation, or at least can rule out other possible operating paths.

On the basis of prior considerations of the role of task interference between operating in-vehicle HMIs and performing the DDT, this study provides evidence that user education has beneficial effects for a safer and more efficient transition from manual to automated driving mode. User education in the area of driving automation is of utmost importance because first contacts with combined automation and ADAS are challenging for users [15,17]. The mitigation of ineffective and inefficient human operator performance with driving automation plays an important role not only for proximal goals of safety and effectiveness, but also for the more distal adoption of this technology with ease of use being a precursor of intention to use [51]. Valuable insights for technology adoption can be gained by considering research on automation in aviation (see e.g., [26,52]) where the training approaches for pilots have been successfully established. Trösterer et al. [26] outlined how pilots practiced operating an airplane under fully functional and defective conditions in both simulation and real-world environments. Such educational approaches, however, are most likely not applicable for driving automation since the user population is much larger as compared with the small number of pilots and the enormous time and financial costs of such trainings. While automated driving users must be considered as novices when it comes to using automation [53], aircraft pilots are highly skilled and trained individuals when operating an automated system. We must conclude that the driving automation domain faces different conditions from that of the aviation sector. Nevertheless, there is the same need for user training and regulation thereof.

A combined view of results in the mental model questionnaire and interaction performance shows that at first, user education positively impacts the knowledge on how to interact with driving automation. This knowledge is then transferred to actual application when encountering a combined ADAS and automation function. However, one has to take into consideration that the declarative knowledge is rather transferred to first-order use cases of transitions from manual as compared with second-order use cases of transitions between levels of automation.

Directly comparing the owner’s manual and the tutorial showed that no clear advantage of one approach over the other was present. There was no larger issue detected for the transfer of information from abstract system description in the owner’s manual to the application context in the automated vehicle. The use case focus of the tutorial did not show a clear advantage for the outcome of the experimenter rating and this finding does not support prior considerations based on Wickens et al. [29]. We explain this is the fact that the tutorial, as designed for this study, also conveyed information in a comparably complex way for the participants through the presence of erroneous answers. In addition, there was no clear advantage of active and guided learning as compared with passive and unguided learning. This result might also be due to the combined complexity of the questions and distractors. The users might also have gotten lost in the tutorial itself rather than only in the owner’s manual [29].

4.3. Limitations and Future Research

The two user education approaches in this study were used as stand-alone procedures. Hence, both of them suffer from specific drawbacks (see Table 1). For example, the owner's manual does not support strict experimental control in terms of objectivity and is passive in its nature of learning. Major drawbacks of the tutorial are the degree of complexity which is mainly due to the high amount of incorrect answers and the necessary trial-and-error approach that users have to take without being provided with any information on the HMI beforehand. Thus, a combination of both approaches in the form of a system-focused information presentation by means of the owner's manual and subsequent guided-learning objectives test might be more suitable. Such approaches are frequently used in web-based trainings by large companies to ensure that their employees have an understanding of their roles and responsibilities within the working context [54].

The number of participants in the present study was rather small when considering the between-subjects design leading to a sample size of $n = 8$ per condition. Cumming and Finch [50] recommend a number of $n = 10$ when applying CI approaches for determining statistical inference. The aim of the present research, however, was to find out about differences and trends of user education concepts and these can well be inferred from the present data set and the applied statistical methodology. Subsequent investigations should consider collecting larger samples to ensure robustness of inferential statistics.

Furthermore, there was only one behavioral measure included in the present analysis. The question of how behavioral indicators on a more operational level [55] such as speed and accuracy [19] benefit from user education remains to be answered. One also has to consider that the experimenter rating, as used in the present study, represents a rather conservative performance indicator compared to time-on-task and error rates, and in a prior study did not show strong performance increases in repeated interaction with driving automation [15]. In this study, despite the standardized approach in applying the experimenter rating, there still remained room for interpretation for the experimenter. As far as possible we counteracted this by assigning specific errors (e.g., pressing a wrong button, unintentionally deactivating the function) to each category for all use cases. Nevertheless, we cannot categorically rule out a certain bias. Hence, further research efforts are necessary, including a blind rating process by multiple raters and subsequent comparison of these by means of inter-rater reliability measures (see e.g., [56]).

Additionally, the open question of how long performance gains from user education are apparent over repeated trials of interaction remains to be answered. Considering that first-order interactions (UC1, UC2) affect performance in second-order interactions (UC3, UC4), it is likely that benefits from user education are not only evident at the very first but also at the same time as most critical encounters with driving automation. The same holds true for mental models. With a superiority of user education present, it remains unclear, whether and how fast non-educated users catch up with informed users. In that vein, it is also the role of future research to learn the implications of training on mental models and operator performance over a longer course of time. The present study found effects of education on single trials within a short simulation drive. However, users of driving automation might also be trained for long-term effects, and retrained in case new functionalities are added to the automated vehicle via over-night updates. Moreover, automated vehicle HMIs will most likely differ between OEMs and thus the question arises as to how to train users so that when changing to another automated vehicle their understanding and appropriate performance are ensured.

This study showed that an approach based on the owner's manual contributes to formation of accurate mental models and increases interaction performance. However, for an application under real-world conditions, there is the issue that less than half of the users will read the owner's manuals [28]. From this group, there might still be a large estimated number who preferred to skim through the owner's manual and not fully understood the presented information. Therefore, it is questionable whether the positive effects of stand-alone owner's manual-based user education procedures emerge for the introduction of combined automation and ADAS to the consumer market. In addition, as

outlined in the prior considerations, information about automated vehicle HMIs is not presented separately but rather within a booklet that can be composed of up to 500 pages. Despite the positive effects of the owner's manual found here, it remains to be seen whether future customers find this information and apply it accordingly after having read additional information on vehicle maintenance, warnings, and much more.

In the present analysis, we calculated mean instead of median values as the measure of central tendency even though the experimenter rating is ordinal in nature. This procedure is supported by Norman [57] who describes the applicability of parametric statistical methods on ordinal data. Prior research on controllability have also used Cooper–Harper scale items and applied parametric approaches on them [44,58,59].

The present work followed an applied research purpose to investigate the applicability of two user education concepts. Thus, results of increased mental models and interaction performance cannot be traced back to a single variation between the two conditions but rather to a combination of several factors. There is a need for future basic research to determine how each approach individually might be improved or impeded by systematically varying conceptual characteristics.

5. Conclusions

The present study presents first evidence for the effectiveness of educating users prior to operating an HMI for automated driving by means of an owner's manual or an interactive tutorial. Both these approaches might be considered in realistic settings when selling automated vehicles to customers. Since they have the potential to facilitate understanding and first use of the functions, it is likely that such approaches would lead to an increase in acceptance and adoption of driving automation technology. When it comes to supporting safe and efficient human-automation interaction, however, we need to initially pursue these two goals through user-centered design. Appropriate design must come first and should reduce the need for user training as far as possible. Hence, rather than following a “blame and train” [60] approach of adapting human behavior to technology, we need to ensure that all possibilities of technology design are exhausted before proceeding to the design of user education.

In addition, there are implications for research methodology when evaluating HMIs for driving automation. On the one hand, it is possible to quickly train users and thus assess performance of more skilled users without the necessity of training the users by using a high-fidelity driving simulator, which is often cost and time consuming. On the other hand, over-educating participants in user studies might conceal observations about the intuitiveness and learnability of HMIs. Eventually, researchers and developers of automated vehicle HMIs need to decide, based on the aim of the respective study, whether and how to include user education in their experiments.

Author Contributions: Conceptualization, Y.F., S.H., F.N., and J.K.; methodology, Y.F., S.H., and F.N.; formal analysis, Y.F.; data curation, Y.F.; writing—original draft preparation, Y.F.; writing—review and editing, Y.F., S.H., F.N., and J.K.; visualization, Y.F.; supervision, S.H., J.K., and A.K.

Funding: This research received no external funding.

Acknowledgments: The authors would like to thank Viktoria Geisel for her support in study preparation and data collection during her internship.

Conflicts of Interest: We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled.

Appendix A

Matlab Code for Bootstrap Statistics (exemplary for one item of mental model L3 automation)

```
%Set number of bootstrap replicates
nrep=10000;
%Create space for bootstrap replicates
mean_BL=zeros(1, nrep);
mean_ML=zeros(1, nrep);
mean_TL=zeros(1, nrep);
%Start for loop
for i=1:nrep
    %draw a random sample with replacement from each condition
    sample_Baseline=randsample(MM_L3_BL, length(MM_HAF_BL), true);
    sample_Manual=randsample(MM_L3_ML, length(MM_HAF_ML), true);
    sample_Tutorial=randsample(MM_L3_TL, length(MM_HAF_TL), true);
    %calculate mean from each random sample and store in array
    mean_BL(i)=mean(sample_Baseline);
    mean_ML(i)=mean(sample_Manual);
    mean_TL(i)=mean(sample_Tutorial);
end
```

References

1. SAE. *Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems*; SAE International: Warrendale, PE, USA, 2018.
2. Naujoks, F.; Wiedemann, K.; Schömig, N.; Hergeth, S.; Keinath, A. Towards guidelines and verification methods for automated vehicle HMIs. *Transp. Res. Part F Traff. Psychol. Behav.* **2019**, *60*, 121–136. [[CrossRef](#)]
3. NHTSA. *Federal Automated Vehicles Policy. Accelerating the Next Revolution in Roadway Safety*; NHTSA: Washington, DC, USA, 2016.
4. Crundall, D.; Andrews, B.; Van Loon, E.; Chapman, P. Commentary training improves responsiveness to hazards in a driving simulator. *Accid. Anal. Prev.* **2010**, *42*, 2117–2124. [[CrossRef](#)] [[PubMed](#)]
5. Pradhan, A.K.; Pollatsek, A.; Knodler, M.; Fisher, D.L. Can younger drivers be trained to scan for information that will reduce their risk in roadway traffic scenarios that are hard to identify as hazardous? *Ergonomics* **2009**, *52*, 657–673. [[CrossRef](#)]
6. Bainbridge, L. Ironies of automation. *Automatica* **1983**, *19*, 775–779. [[CrossRef](#)]
7. Endsley, M.R. Autonomous driving systems: A preliminary naturalistic study of the Tesla Model S. *J. Cogn. Eng. Decis. Mak.* **2017**, *11*, 225–238. [[CrossRef](#)]
8. Pradhan, A.K.; Sullivan, J.; Schwarz, C.; Feng, F.; Bao, S. Training and Education: Human Factors Considerations for Automated Driving Systems. In *Road Vehicle Automation 5*; Meyer, G., Beiker, S., Eds.; Springer: Cham, Switzerland, 2019; pp. 77–84.
9. Jarosch, O.; Kuhnt, M.; Paradies, S.; Bengler, K. It's Out of Our Hands Now! Effects of Non-Driving Related Tasks during Highly Automated Driving on Drivers' Fatigue. In Proceedings of the 9th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, Manchester Village, VT, USA, 26–29 June 2017.
10. Forster, Y.; Naujoks, F.; Neukum, A. Increasing anthropomorphism and trust in automated driving functions by adding speech output. In Proceedings of the IEEE Intelligent Vehicles Symposium, Redondo Beach, CA, USA, 11–14 June 2017.
11. Hergeth, S.; Lorenz, L.; Krems, J.F. Prior familiarization with takeover requests affects drivers' takeover performance and automation trust. *Hum. Factors* **2017**, *59*, 457–470. [[CrossRef](#)] [[PubMed](#)]

12. Feldhütter, A.; Segler, C.; Bengler, K. Does Shifting Between Conditionally and Partially Automated Driving Lead to a Loss of Mode Awareness? In Proceedings of the International Conference on Applied Human Factors and Ergonomics, Los Angeles, CA, USA, 17–21 July 2017.
13. Naujoks, F.; Mai, C.; Neukum, A. The effect of urgency take-over requests during highly automated driving under distraction conditions. *Adv. Hum. Asp. Transp.* **2014**, *7*, 431.
14. Forster, Y.; Hergeth, S.; Naujoks, F.; Krems, J.F. How Usability can Save the Day. Methodological Considerations for Making Automated Driving a Success Story. In Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Toronto, ON, Canada, 23–25 September 2018; Donmez, B., Walker, B.N., Fröhlich, K., Eds.; ACM: New York, NY, USA, 2018.
15. Forster, Y.; Hergeth, S.; Naujoks, F.; Beggiato, M.; Krems, J.F.; Keinath, A. Learning to Use Automation: Behavioral Changes in Interaction with Automated Driving Systems. *Transp. Res. Part F Traff. Psychol. Behav.* **2019**, *62*, 599–614. [[CrossRef](#)]
16. Beggiato, M.; Pereira, M.; Petzoldt, T.; Krems, J. Learning and development of trust, acceptance and the mental model of ACC. A longitudinal on-road study. *Transp. Res. Part F Traff. Psychol. Behav.* **2015**, *35*, 75–84. [[CrossRef](#)]
17. Forster, Y.; Hergeth, S.; Naujoks, F.; Beggiato, M.; Krems, J.F.; Keinath, A. Learning and Development of Mental Models in Interaction with Driving Automation: A Simulator Study. In Proceedings of the 10th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, Santa Fe, NM, USA, 24–27 June 2019.
18. Naujoks, F.; Hergeth, S.; Wiedemann, K.; Schömig, N.; Keinath, A. Use Cases for Assessing, Testing, and Validating the Human Machine Interface of Automated Driving Systems. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2018**, *62*, 1873–1877. [[CrossRef](#)]
19. Wickens, C.D. Multiple resources and performance prediction. *Theor. Issues Ergon. Sci.* **2002**, *3*, 159–177. [[CrossRef](#)]
20. Knappe, G.; Keinath, A.; Meinecke, C. Empfehlungen für die Bestimmung der Spurhaltequalität im Kontext der Fahrsimulation. *MMi-Interaktiv* **2006**, *11*, 3–13.
21. AAM. *Statement of Principles, Criteria and Verification Procedures on Driver Interactions with Advanced in-Vehicle Information and Communication Systems*; Alliance of Automobile Manufacturers: Washington, DC, USA, 2006.
22. NHTSA. *Visual-Manual NHTSA Driver Distraction Guidelines for in-Vehicle Electronic Devices*; National Highway Traffic Safety Administration (NHTSA), Department of Transportation (DOT): Washington, DC, USA, 2012.
23. Forster, Y.; Naujoks, F.; Neukum, A.; Huestegge, L. Driver compliance to take-over requests with different auditory outputs in conditional automation. *Accid. Anal. Prev.* **2017**, *109*, 18–28. [[CrossRef](#)] [[PubMed](#)]
24. Beller, J.; Heesen, M.; Vollrath, M. Improving the driver–automation interaction: An approach using automation uncertainty. *Hum. Factors* **2013**, *55*, 1130–1141. [[CrossRef](#)]
25. Louw, T.; Merat, N. Are you in the loop? Using gaze dispersion to understand driver visual attention during vehicle automation. *Transp. Res. Part C Emerg. Technol.* **2017**, *76*, 35–50. [[CrossRef](#)]
26. Trösterer, S.; Meschtscherjakov, A.; Mirnig, A.G.; Lupp, A.; Gärtner, M.; McGee, F.; McCall, R.; Tscheligi, M.; Engel, T. What We Can Learn from Pilots for Handovers and (De)Skilling in Semi-Autonomous Driving. In Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Oldenburg, Germany, 24–27 September 2017; Boll, S., Pfleging, B., Donmez, B., Eds.; ACM: New York, NY, USA, 2017; pp. 173–182.
27. Fitts, P.M.; Posner, M.I. *Human Performance*; Brooks/Cole: Oxford, UK, 1967.
28. Mehlenbacher, B.; Wogalter, M.S.; Laughery, K.R. On the Reading of Product Owner’s Manuals: Perceptions and Product Complexity. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2002**, *46*, 730–734. [[CrossRef](#)]
29. Wickens, C.D.; Hollands, J.G.; Banbury, S.; Parasuraman, R. *Engineering Psychology & Human Performance*; Psychology Press: Hove, UK, 2015.
30. Van Loggem, B. User Documentation: The Cinderella of Information Systems. In *Advances in Information Systems and Technologies*; Rocha, Á., Correia, A.M., Wilson, T., Stroetmann, K.A., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 167–177.
31. Beggiato, M.; Krems, J.F. The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. *Transp. Res. Part F Traff. Psychol. Behav.* **2013**, *18*, 47–57. [[CrossRef](#)]
32. Blömacher, K.; Nöcker, G.; Huff, M. The role of system description for conditionally automated vehicles. *Transp. Res. Part F Traff. Psychol. Behav.* **2018**, *54*, 159–170. [[CrossRef](#)]

33. Forster, Y.; Kraus, J.; Feinauer, S.; Baumann, M. Calibration of Trust Expectancies in Conditionally Automated Driving by Brand, Reliability Information and Introductory Videos: An online study. In Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Toronto, ON, Canada, 23–25 September 2018; Donmez, B., Walker, B.N., Fröhlich, K., Eds.; ACM: New York, NY, USA, 2018.
34. Körber, M.; Baseler, E.; Bengler, K. Introduction matters: Manipulating trust in automation and reliance in automated driving. *Appl. Ergon.* **2018**, *66*, 18–31. [[CrossRef](#)]
35. Hock, P.; Kraus, J.; Walch, M.; Lang, N.; Baumann, M. Elaborating Feedback Strategies for Maintaining Automation in Highly Automated Driving. In Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Ann Arbor, MI, USA, 24–26 October 2016; pp. 105–112.
36. Slater, M. A note on presence terminology. *Presence Connect* **2003**, *3*, 1–5.
37. Brehm, J.W. *A Theory of Psychological Reactance*; Academic Press: New York, NY, USA, 1966.
38. Pradhan, A.K.; Fisher, D.L.; Pollatsek, A. The effect of PC-based training on novice driver' risk awareness in a driving simulator. In Proceedings of the 3rd International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, Rockport, ME, USA, 27–30 June 2005.
39. Yamani, Y.; Samuel, S.; Knodler, M.A.; Fisher, D.L. Evaluation of the effectiveness of a multi-skill program for training younger drivers on higher cognitive skills. *Appl. Ergon.* **2006**, *52*, 135–141. [[CrossRef](#)] [[PubMed](#)]
40. Payre, W.; Cestac, J.; Dang, N.T.; Vienne, F.; Delhomme, P. Impact of training and in-vehicle task performance on manual control recovery in an automated car. *Transp. Res. Part F Traff. Psychol. Behav.* **2017**, *46*, 216–227. [[CrossRef](#)]
41. Gagne, R. Learning Outcomes and Their Effects. Useful Categories of Human Performance. *Am. Psychol.* **1984**, *39*, 377–385. [[CrossRef](#)]
42. Forster, Y.; Naujoks, F.; Neukum, A. Your Turn or My Turn? Design of a Human-Machine Interface for Conditional Automation. In Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Ann Arbor, MI, USA, 24–26 October 2016; pp. 253–260.
43. Manca, L.; de Winter, J.C.F.; Happee, R. Visual Displays for Autoamted Driving: A Survey. In *Adjunct Proceedings of the 7th International Conference on Automotive User Interfaces and Iternactive Vehicular Applications*; Association for Computing Machinery: New York, NY, USA, 2015.
44. Naujoks, F.; Purucker, C.; Neukum, A.; Wolter, S.; Steiger, R. Controllability of Partially Automated Driving functions—Does it matter whether drivers are allowed to take their hands off the steering wheel? *Transp. Res. Part F Traff. Psychol. Behav.* **2015**, *35*, 185–198. [[CrossRef](#)]
45. Naujoks, F.; Forster, Y.; Wiedemann, K.; Neukum, A. A Human-Machine Interface for Cooperative Highly Automated Driving. In *Advances in Human Aspects of Transportation*; Springer: Cham, Switzerland, 2016; pp. 585–595.
46. Nevitt, J.; Hancock, G.R. Performance of Bootstrapping Approaches to Model Test Statistics and Parameter Standard Error Estimation in Structural Equation Modeling. *Struct. Equ. Model.* **2001**, *8*, 353–377. [[CrossRef](#)]
47. Keselman, H.J.; Wilcox, R.R.; Othman, A.R.; Fradette, K. Trimming, Transforming Statistics, and Bootstrapping: Circumventing the Biasing Effects of Heterescedasticity and Nonnormality. *J. Mod. Appl. Stat. Methods* **2002**, *1*, 288–309. [[CrossRef](#)]
48. Schäfer, T. Die New Statistics in der Psychologie. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* **2018**, *50*, 3–18.
49. Cumming, G. *Understanding the New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis*; Routledge: Abingdon, UK, 2013.
50. Cumming, G.; Finch, S. Inference by eye: Confidence intervals and how to read pictures of data. *Am. Psychol.* **2005**, *60*, 170. [[CrossRef](#)]
51. Davis, F.D. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Q.* **1989**, *13*, 319. [[CrossRef](#)]
52. Sarter, N.B.; Mumaw, R.J.; Wickens, C.D. Pilots' monitoring strategies and performance on automated flight decks: An empirical study combining behavioral and eye-tracking data. *Hum. Factors* **2007**, *49*, 347–357. [[CrossRef](#)]
53. Hart, C.A. Self Driving Safety Steps into the Unknown. Available online: <http://www.thedrive.com/tech/26896/self-driving-safety-steps-into-the-unknown> (accessed on 14 March 2019).

54. McIlwraith, A. *Information Security and Employee Behaviour: How to Reduce Risk through Employee Education, Training and Awareness*; Routledge: Abingdon, UK, 2016.
55. Michon, J.A. A critical view of driver behavior models: What do we know, what should we do? In *Human Behavior and Traffic Safety*; Springer: Boston, MA, USA, 1985; pp. 485–524.
56. Forster, Y.; Hergeth, S.; Naujoks, F.; Krems, J.F. Unskilled and Unaware: Subpar Users of Automated Driving Systems Make Spurious Decisions. In Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Toronto, ON, Canada, 23–25 September 2018; ACM: New York, NY, USA, 2018.
57. Norman, G. Likert scales, levels of measurement and the “laws” of statistics. *Adv. Health Sci. Educ. Theory Pract.* **2010**, *15*, 625–632. [[CrossRef](#)] [[PubMed](#)]
58. Kauffmann, N.; Winkler, F.; Vollrath, M. What Makes an Automated Vehicle a Good Driver? In Proceedings of the CHI Conference, Montreal, QC, Canada, 21–26 April 2018.
59. Neukum, A.; Lübbecke, T.; Krüger, H.-P.; Mayser, C.; Steinle, J. ACC-Stop&Go: Fahrerverhalten an Funktionalen SYSTEMGRENZEN, 5. Workshop Fahrerassistenzsysteme—FAS. 2008. Available online: http://www2.psychologie.uni-wuerzburg.de/izvw/texte/2008_Neukum_etal_ACC_Stop&Go.pdf (accessed on 16 April 2019).
60. Norman, D. *The Design of Everyday Things: Revised and Expanded Edition*; Basic Books: New York, NY, USA, 2013.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).