Piloting a Survey for Understanding Public Expectations Towards Autonomous Train Experiences as a Passenger

Sytse A. M. Vos (s3001547)

Bachelor Psychology, University of Twente

202000384: Bachelor Thesis

Simone Borsci, Sarah Kusumastuti

July 1st, 2025

Abstract

With 5% of all inland transport being done by trains in Europe in 2020 and spanning over 250 billion kilometers of rail across the entirety of Europe, it has to keep improving to make sure it is capable of transporting all these people. In this study, public expectations towards an autonomous train experience were investigated. This study is divided into 2 parts. The first part is a literature analysis, and the second part included a the pilot of a survey study based on the findings of earlier research, but also included 2 videos of a train experience to see which the participant attributed a human driver to. The literature analysis found that people find safety and reliability the most important and prefer a level of automation where the train is autonomous, but there are still humans aboard in case of emergency. The survey results indicated significant effects of Risk and Attitude on the willingness to use trains ($p = .003$), and the opinion surrounding autonomous trains were largely in line with earlier research concerning the most important needs and concerns, except for the effect of previous experience with autonomous vehicles and influence of Age which were not found in this study. Overall, the public seems open to the use of autonomous trains in the future.

Introduction

A lot of effort is put into improving the infrastructure within Europe, which is designed to allow people to travel within and between all the different countries with as much ease as possible. These infrastructures include multiple travel methods, such as roads used by cars and busses, airports which allow travelling by plane, ports used by boats, and railways providing a means of travel for both humans and goods. The frequency of use of these modalities differ significantly depending the country (Palm, 2022).

One of the main forms of transport is inland transport, which is transport used to travel within the same country and across land. This mode of transport includes methods like cars, busses, and trains. Of these forms of transport, cars were the most popular as of the year 2022, increasing from providing between 82.0% and 83.1% of all inland passenger transport between 2010 and 2019 to 87.2% in 2020 (Palm, 2022). Busses were at the second place, but experienced a decrease ranging from 9.5% and 10.4% between 2010 and 2019 to 7.4% in 2020. The third most used mode of transport was trains, increasing from 7.1% in 2010 to 8.0% in 2019, but then experienced a significant decrease to 5.4% in 2020, which is possible be explained by the emergence of covid. Even though some of these numbers seem small in comparison, they still represent billions of people that depend on these modes of transport. This shows that a large number of people still use and depend on these modalities.

The railways covering Europe include over 250 billion kilometers of rails in the year 2021. The group in charge of overseeing most of the rail network within Europe is Europe's Rail Join Undertaking also known as EU-Rail (EU-Rail, 2021). The EU-Rail has established the goal of "delivering a high capacity integrated European railway network by eliminating barriers to interoperability, providing solutions for full integration, and achieving faster uptake and deployment of innovation through funded projects.". The organization also focuses on promoting the digitalization and automation to reduce costs for the overall rail industry, strengthen the reliability and flexibility, and increase the capacity of the railway network.

The focus on automation of the railway network comes from the increasing automation and digitalization of the world which offers new technological and operational solutions to current capacity problems (EU-Rail, 2021). By automating trains it becomes possible to improve the current systems and create a higher capacity for passenger transport. There are multiple levels that exist within vehicle automation which are described as the Grades of Automation (GoA) (Reichmann et al., 2025). There exists a total of 5 levels within GoA with level 0 being the lowest and 4 the highest level of automation. Grade 0 relates to no automation at all, which means that there is a human train driver physically on-sight controlling the train without any computerized automation. A GoA of 1 there is still a human driver physically steering the train, however there is now also an Automated Train Protection (ATP) system present, assisting the driver by controlling speeds and slowing down if the train exceeds the speed limit. GoA 2 describes the situation where the train is controlled by both ATP and Automated Train Operations (ATO) and the driver takes a position leaning more towards assisting and supervising the automated systems comparted to driving the train. During the creation of this study, this is also the most common type of train. At GoA 3 is described as Driverless Train Operations (DTO) where the train is fully controlled by automated systems, but there is still a human present for the door systems and in case of system errors. At the final level of Unattended Train Operations (UTO), GoA 4, there is no longer a human present on the train who is controlling and maintaining the automated systems, meaning that the train is fully autonomous (Reichmann et al., 2025).

This research will focus on understanding public expectations towards autonomous train experiences as a passenger. Expectation is defined as the strong beliefs that something is or will be the case because it is likely (Oxford University Press, n.d.). Expectations are important to consider as they influence the likelihood of people wanting to take part in or experience a new concept, such as train automation. Currently, there has already been some literature on peoples' acceptance toward train automation, however, there is still a lack of conclusiveness and physical experiments on this topic. Therefore, this research will take the already existing literature and try to replicate and combine the findings with one comprehensive survey and compare the outcome to the results of previously performed studies. As in previous studies, this research will also investigate peoples' trust towards autonomous trains in a qualitative way considering the potential bias that people have when a train is not under direct human control.

The research will start with a summary of previous research done into the topic of autonomous trains, looking into the existing knowledge on the topic of people's expectations on automated trains and the factors that play a role in this with the goal of developing a survey from these studies. It will look into the findings and later on compare the results of these articles with the results from this study. It will also include a model created from previous findings which is used to predict the responses to the survey to validate the overall efficacy of the survey itself.

Previous work and results about passengers expectations towards autonomous trains

Independently from the model adopted by the researchers for their investigation, and how they categorized the factors of their investigation, the next section will propose a summary of previous research done, and the associated results, about the topic of passenger expectations towards autonomous trains. This will be used to identify key relationships and expectations to inform our study.

There are slight disagreements on the factors that play a role in people's expectations towards automated trains, where expectations are defined as the predisposing beliefs people have around automated trains before they first use it. Lemonnier et al. (2020) identified 22 different factors influencing the acceptance of automated vehicles. While the most prominent independent factors she identified are Age, Gender, and Personality (Lemonnier et al., 2020), these aspects cannot be influenced significantly by changing the technology of the trains unlike the dependent factors that she identified (Table 1). These factors are more useful when increasing people's expectations as they are able to be influenced and changed. These factors were labeled under the names "Preferences" and "Perception". Preferences include only 2 concepts like being able to perform onboard activities and the characteristics of the vehicle, while Perceptions includes a larger number of factors and includes the level of knowledge on automated trains, the perceived ease of use, the perceived usefulness, the social influence, the general attitude towards automation, the concerns around automation including privacy and environmental concerns, the general benefits as perceived by the public, the trust in the automated vehicle, the level of human control over the automated vehicle, and lastly the perceived risks associated with automated vehicles.

Table 1

List of factors found in earlier research

Aspects	**Description**
Preferences	
Onboard activities	Being able to perform activities onboard the train (Lemonnier et al., 2020).
Characteristics	Certain characteristics that belong to the automated train such as availability, comfort, sound environment and sight (Lemonnier et al., 2020).
Perceptions	

Level of knowledge	The level of understanding that the user has on automated vehicles (Lemonnier et al., 2020).
Perceived ease of use	How ease the user believes the automated train is to use (Lemonnier et al., 2020).
Perceived usefulness	How useful the user believes automated trains to be (Lemonnier et al., 2020, 2023).
Social influence	The effect of subjective social norms on the use of automated trains.
Attitude	The general opinion the user has on automated vehicles (Lemonnier et al., 2020).
Concerns	The types of concerns the user has with automated vehicles including both privacy concerns and environmental concerns (Hilgarter & Granig, 2020; Lemonnier et al., 2020, 2023).
Benefits	The perceived benefits that automated trains have over human driven trains according to the user (Lemonnier et al., 2023).
Trust	The level of trust the user has in automated trains(Cogan et al., 2022; Lemonnier et al., 2020).
Human control	The level of control that humans on the train have over the automated systems (Lemonnier et al., 2020).
Perceived risk	The risks that the user beliefs come with the use of automated trains (Lemonnier et al., 2020).

The list of factors proposed by Lemonnier et al. (2020) concerning the expectations of train automation are not entirely agreed upon by other experts, the division of “Preference” and “Perception” is one among other options. For instance, instead of labeling factors that influence automated train expectations under the names of Preferences and Perception, a different research carried out by Hilgarter and Granig (2020) chose to focus their attention and categorization of factors around the automation of trains under “Societal Challenges”, “Technological Challenges”, “Legal Challenges”, and “Economic Challenges”. In this research, Societal Challenges include topics similar to that of the previous work like the fear of job losses and acceptance/awareness. In this research, the Technological Challenges exclusively focus on reliability which entails matters concerning the problems within the technology. The Legal Challenges include a topic not yet discussed, which focusses on who is liable in the case of accidents occurring with faulty technology. Lastly the Economic Challenges are similar to that of the earlier research, but here also include concerns regarding the cost of implementing the infrastructure for automation. The differences in these research show that although they were carried out within the same year, some researchers may find or choose to focus on different structures or components of people's expectations while also finding overlapping topics.

Factors influencing the expectations do not limit themselves to just those named earlier. A more recent research carried out by Lemonnier et al. (2023) identified other key factors for train passengers when assessing the railway service through semi-structured interviews, such as: Attitude, Knowledge, Usefulness, and Environmental concerns including e.g., safety concerns and themes like job loss, dehumanization, the possibility of malfunctions, and concerns around communication. Some of these factors were in line with previous studies by Hilgarter and Granig (2020) and Lemonnier et al. (2020). A new factor identified by Lemonnier et al. (2023) was the one concerning the expected benefits for passengers that come with the automation of trains, including railway operator savings, the evolution of railway professions and the environmental benefit.

In past 10 years, it has been found that people with previous experience with autonomous vehicles have a significant influence on their expected use of autonomous trains in the future. These previous studies were done mostly in the form of structured interviews or surveys (Arzer et al., 2024; Cogan et al., 2022; Detjen et al., 2021; Fraszczyk et al., 2015; Fraszczyk & Mulley, 2017; Morast et al., 2023; Pakusch & Bossauer, 2017). For instance, in a study testing the acceptance of fully automated vehicles through a quantitative online study, Pakusch and Bossauer (2017) found that from their entire dataset, 77.6% of people were willing to use autonomous trains in the future. They also found that people who already had previous experience with autonomous vehicles were significantly more likely to be willing to use autonomous trains in the future as well with 88% on average (SD= 32.7%) being willing to use autonomous trains. Of those who had not yet experienced autonomous vehicles, only 72% on average (SD= 45.1%) were willing to use autonomous trains which is still a large percentage but also significantly smaller compared to the other group. Other factors like age, gender, and their current main form of transportation however had no significant effect.

The results found by Pakusch and Bossauer (2017) are also supported by other researchers. Cogan et al. (2022) found evidence that support for autonomous vehicles is high and even higher amongst those with previous experience with such vehicles. Besides willingness to use, this research also looked into peoples' preferences of onboard facilities. They found that people still prefer the presence of onboard personnel even in the train is driving by itself. They also found that GoA-3 is preferred more that GoA-4 which is in line with their other findings and those of Lemonnier et al. (2023) which also states that GoA-3 is the most accepted form of automation so far. The most important aspects to think of according to their respondents were those of reliability and safety. They also considered relay reduction and ticket costs to be important aspects together with cybersecurity being a common concern. Fraszczyk et al. (2015) however found that reduced ticket prizes, extended running periods, and increased train frequency were not convincing reasons for the majority of their participants to pick autonomous trains over human driven trains. This research also found that train driver unemployment is not a concern amongst the majority of the public with 62% stating that they do not think that autonomous vehicles are a threat to a driver's job security.

Another important aspect was highlighted by Arzer et al. (2024), indicating the effect of factors influencing public opinion on autonomous trains. They found similar results to the previous articles, stating that the factor that plays the largest role is that of safety and reliability. They also put an emphasis on their findings that people want to have a basic understanding of protective services onboard the train, preferably by onboard personnel or external recourses. This was further supported by Fraszczyk and Mulley (2017), emphasizing peoples' need for information and safety. They state that the public is not yet informed about safety measures onboard an autonomous train and is therefore hesitant and cautious of its use. Detjen et al. (2021) found conflicting interests within their sample with some people being enthusiastic about autonomous trains, but others being skeptical about its use. This paper states that it is best to create a targeted communication strategies to make sure that each group of people receives information is a way best suited to them. With this, it becomes easier to inform the public of autonomous trains and thereby increase their acceptance of using autonomous trains in the future.

In addition, research performed by Morast et al. (2023) provides further insight into the factors that influence the public acceptance of autonomous trains. They state that the largest part of the population is willing to accept and use autonomous trains, and that added to the previously mentioned factors this acceptance is influenced by having experience with similar technologies to that of autonomous trains and the frequency of train usage. According to this article however, age does play a role saying that younger people are more willing to accept autonomous trains compared to elderly people, unlike that of Pakusch and Bossauer (2017) in which age was not a factor. One important concern that is not looked into enough around autonomous trains according to Morast et al. (2023) is the fear of cyberattacks, but not only in autonomous trains, but also in daily life and automated transport in general.

Aims of the study

Based on previous studies we designed a survey that attempted to collect information about importance of the following aspects when people are assessing their current usage of trains and their intention to use future trains:

1. Demographics aspects. These aspects can be defied as people's characteristics such as age, gender etc. but also include aspects that are considered important to affect people like: regularity of usage (Morast et al., 2023) and previous experience with autonomous systems which seems to predict people trust toward such system (Cogan et al., 2022).
2. On board communication: This investigates the presence of onboard train personnel and/or the presence of onboard digital screens that can provide information to the passenger, and it is considered important to predict willingness to use (Cogan et al., 2022).
3. Presence of personnel: This describes the effect of the presence of onboard personnel who are able to provide information or step in during technical difficulties by for example taking over the control of the train (Cogan et al., 2022; Lemonnier et al., 2023).
4. Service quality: these are the aspects that people deem important when assessing the overall quality of train service (Fraszczyk & Mulley, 2017).
5. Attitude towards driverless trains: This shows the effect of the general attitude people have towards autonomous trains (Lemonnier et al., 2020)

The current pilot aims to inform the design of the survey, explore the main expected relationship among the components of the survey e.g., how the different factors affect the overall intention of usage. In the present pilot, we do not expect to draw a conclusion about passenger expected experience, but instead aim to check if the expected main relationship among the aspects can be tested by the survey.

In particular, after checking by means of descriptive and correlational analysis if the results of the present survey are in line with previous studies, we will also attempt to explore if we can use the survey data to test the following main expectations based on previous studies:

(RQ1) is the intention to use future trains declared by the passenger in the survey predicted by declared trust towards autonomous trains, participant age, their attitude and their willingness to take risks?

(RQ2) does previous experience with autonomous vehicles affect people's declared trust towards autonomous trains and their attitude towards trains?

Moreover, considering that passenger cannot currently experience (or be aware of experiencing) autonomous trains in the real world if not in rare cases, and in the attempt to model if people answers, when exposed to realistic scenarios we created two virtual reality videos of a train journey from the passenger perspective. Each video showed the same journey with a different level of intensity of break and acceleration of the train (we refer to this as bumpiness).

The videos were used to explore people tendency to attribute of bumpiness to the fact that the train is automated or driven by a human. In line with that we also tested if (RQ3) seeing video before or after the survey affects people declared level of trust and intention of usage.

Methods

The survey study was mixed with an online experiment in which people were divided between conditions to experience two videos before doing the survey (condition 1) or after the survey (condition 2). In each of the conditions the videos were showing a train journey which was either more or less bumpy, the order in which the 2 videos were presented was randomized.

Participants

This study involved a total of 53 participants. Of these participants 32 were male, 21 were female, and 0 described themselves as “Other”. The age of the participants ranged from 20 to 73 years old (M = 39.53, SD = 18.408). Most of the participants were from the Netherlands with 46 people being Dutch (71.6%) and other participants being from Australia, France, India, Italy, Kazakhstan, Mexico, Spain, and the United Kingdom. Participants were recruited through a convenience sample consisting mostly of respondents from acquaintances and personal networks together with students at the University of Twente. The inclusion criteria for the participants contained a fluent understanding of the English language. All participants gave a written informed consent included in the beginning of the survey where they were also informed about their rights to withdraw from the study at any point and refuse the use of their data for the study. The study was approved by the ethics committee of the Faculty of Behavioral, Management and Social Sciences (BMS) at the University of Twente, ethics number 250464.

Materials

For this survey study Qualtrics was used for the creation of the survey (Appendix A). The survey started with questions to test if the results in the current study are similar to that of the previous studies(Arzer et al., 2024; Cogan et al., 2022; Detjen et al., 2021; Fraszczyk et al., 2015; Fraszczyk & Mulley, 2017; Morast et al., 2023; Pakusch & Bossauer, 2017). In line with the aspects we identified as important in literature, the survey had a total of 7 demographic questions, nationality, age, proficiency in English, gender, being acquainted with someone who works in the railway sector, the living area, and risk taking behavior. This was followed by 2 questions about the level of preference concerning automation. Next were 5 questions on transportation habits, after which came 1 question on future concerns from the participant. The survey ended with 1 question for gathering feedback on the survey from the participant. The answers to all the items were given on a scale from 1-5 e.g., 1 - "Strongly Disagree", and 5- "Strongly agree". The willingness to take risk the answers were given on a scale with 10 points from 0 – not willing at all to 10 – very willing. Included in the survey study were also 2 videos made from an experiment setup created by colleagues. This included a VR recording of a train experiences as a passenger together with a split screen recording of the person experiencing the VR.

The 2 videos looked the same, but there was quite a substantial difference in the parameters of the trains which made events such as accelerating and braking more abrupt in one compared to the other. The video clearly present the movement of the chair to convey the abruptness of the events. One video had all the events without any abrupt movements in the chair (i.e., smooth journey), the other video showed 2 abrupt events with very perceivable chair movements i.e., very bumpy journey. Further, a computer on other digital device such as a telephone or tablet/iPad was needed in order to access the survey.

Procedures

After contacting the individual with regards to taking part in the study, they were informed about the general purpose of the research and the overall expected time it would take for them to complete. They were also informed about the information that will be gathered for research purposes together with their rights to withdraw from the study at any point in time. After this, a URL was sent to them which would bring them to the Qualtrics questionnaire, starting with the informed consent and then continuing over to the actual survey. After filling in the basic questions about automated trains, they were confronted with the questions regarding the overall state of the survey itself. For the VR recordings, 2 conditions were created, one where the recordings appeared at the start of the survey, and one where they appeared near the end of the survey. This was done to check for any differences caused by the priming of seeing the recordings of a possible autonomous train experience. After finishing the survey, the participants were given the chance to make some final comments and remarks to the researcher concerning anything they were still curious about or any concerns they might still have. When the participant did not have any further comments or questions, they were debriefed and asked how likely they were to recommend or invite friends/family/acquaintances to also take part in the study.

Data analysis

Rstudio was used for data analysis and gathering important information about the performance of the survey. Before any analyses could be performed, the data from Qualtrics was transported into Rstudio in csv format. The first steps included installing the correct packages needed to perform the analysis within Rstudio, these included “ggplot”, “janitor”, “tidyverse”, and “ltm”. After cleaning the data by screening for any missing or irrelevant data and omitting it from the dataset, descriptive statistics were computed on all the items and aspects of the survey. Participants’ responses to all items on the survey were transmuted from Likert scales into percentages. These were then put into a table to gain a quick overview of the answers that were given in order to see if there were already some items that showed a need for further analysis. Correlation analysis was used to check if results and relationships emerged in previous studies were replicated and in line with expectations in our pilot study.

To answer the question whether the intention to use future trains as declared by the passenger in the survey is predicted by declared trust towards autonomous trains, participant age, their attitude and their willingness to take risks? (RQ1), the answers to the item regarding intention of use were analyzed together with the results on the items asking about their attitude towards autonomous trains, their declared trust, age, and willingness to take risks. Firstly, descriptive statistics were computed to gain some insight into the way the participants answered the questions on risk taking behaviour and their intention to use autonomous trains. To test our expectations, a linear model was made with the intention of use as the dependent variable and the items on declared trust as the independent variables. For intention of use, only the answers to the first statement “Once I will have access to driverless train system, I predict that I would use them regularly” were used. This was because this statement directly indicates the intention of the participant to use autonomous trains in the future, and the second statement “I believe that implementations of driverless trains contribute to increase of operators unemployment” is not directly linked to this aspect. A predictive effect was declared if there existed a correlation with a significance of $\alpha < 0.05$. To test the other factors included in the research question, a multiple linear model was created to test for any effects between the intention of use and the participants’ willingness to take risks, their age, and the attitude towards autonomous trains.

To answer the second research question, does previous experience with autonomous vehicles affect peoples’ declared trust towards autonomous trains and their intention to use autonomous trains (RQ2), analysis was done exploring the effects of previous experience with autonomous systems towards participants declared trust and attitudes towards future trains. This was done in the form of a multiple linear regression model where Previous experience was coded as the independent variable and both Trust and Intention to use as dependent variables.

For the third research question, does the condition of seeing the videos at the beginning or end of the survey have an effect on intention to use autonomous trains in the future, a MANOVA was used to test if there exists a significant difference between the means of each condition regarding the responses on the intention to use and the trust in autonomous trains.

Results

Overview of participants answers to the items

To check whether or not the results from the current study were in line with earlier results, the significance of factors and the responses to the items were put into tables (Table 2, Table 3), which shows an overview of the average scores of all the main aspects from the items on the survey according to the responses from the participants. This includes items on Attitude, Current Concerns, Trust, and Future Concerns.

Table 2

Results from previous research

Factor	Current study Mean (%) + (SD)	Previous studies	Source
		Attitude	
Willingness to use	56 (30.19)	High willingness to use (79.2%)	Cogan et al. (2022); Morast et al. (2023); Pakusch and Bossauer (2017)
Support	66 (24.13)	65% support autonomous trains	Cogan et al. (2022)
Risk taking behaviour	55 (18.56)	N/A	N/A
Experience	35% had experience	36% had experience	Cogan et al. (2022); Morast et al. (2023);

			Pakusch and Bossauer (2017),
Age	39 (18.40)	39 (16.47)	Morast et al. (2023); Pakusch and Bossauer (2017)
Grade of automation	GoA-3 58 (24.03)	GOA-3 53%	Cogan et al. (2022); Lemonnier et al. (2023)
Trust			
Trust	51 (16.80)	N/A	N/A
Operator unemployment	57 (29.48)	62%	Fraszczyk et al. (2015)
Fear of cyber attacks	34 (31.48)	28% rated 5 on scale from 1-5	Cogan et al. (2022); Morast et al. (2023)
Importance of Current Concerns			
Safety	87 (15.98)	66% rated 5 on scale from 1-5	Arzer et al. (2024); Fraszczyk and Mulley (2017)
Reliability	81 (18.28)	49% rated 5 on scale from 1-5	Cogan et al. (2022)
Sustainability	57 (24.09)	27% rated 5 on scale from 1-5	Cogan et al. (2022); Fraszczyk and Mulley (2017)
Accessibility	63 (25.77)	41% rated 5 on scale from 1-5	Cogan et al. (2022); Fraszczyk and Mulley (2017)
Comfort	61 (20.52)	25% rated 5 on scale from 1-5	Cogan et al. (2022); Fraszczyk and Mulley (2017)

Cost of tickets	70 (23.03)	40% rated 5 on scale from 1-5	Cogan et al. (2022); Fraszczyk and Mulley (2017)
Importance of Future Concerns			
Future Reduced ticket price	68 (26.92)	18% rated 5 on scale from 1-5	Cogan et al. (2022); Fraszczyk et al. (2015)
Future Running periods	65 (23.17)	23% rated 5 on scale from 1-5	Cogan et al. (2022); Fraszczyk et al. (2015)
Future train frequency	70 (22.50)	34% rated 5 on scale from 1-5	Cogan et al. (2022); Fraszczyk et al. (2015)
Future Reduced risk for operators	58 (29.40)	N/A	N/A
Future sustainability	53 (25.96)	N/A	N/A
Presence of attendant who can drive	64 (30.45)	20% rated 5 on scale from 1-5	Cogan et al. (2022)
Presence of attendant in driver cab	48 (33.91)	40% rated 5 on scale from 1-5	Cogan et al. (2022)
Presence of digital information screens	66 (26.90)	Utility value of .242 (0.09)	Cogan et al. (2022)
Presence of personnel for information	54 (26.13)	Utility value of -0.332 (0.06)	Arzer et al. (2024); Cogan et al. (2022)

To gain insight into what people find most important with regards to the current state of train transport, the responses to the items of “Transportation Habits” were put into a bar chart to gain a clear view of these results (Figure 1). This figure shows how important the factors of “Safety”, “Reliability”, “Sustainability”, “Accessibility”, “Comfort”, and “Cost of tickets" are according the participants. It can be seen that overall the highest mean belongs to "Safety" (M = 87, SD = 15.98) and the lowest on Sustainability (M = 57, SD = 24.09). None of the participants gave a 0 for either "Safety", "Reliability", or "Comfort".

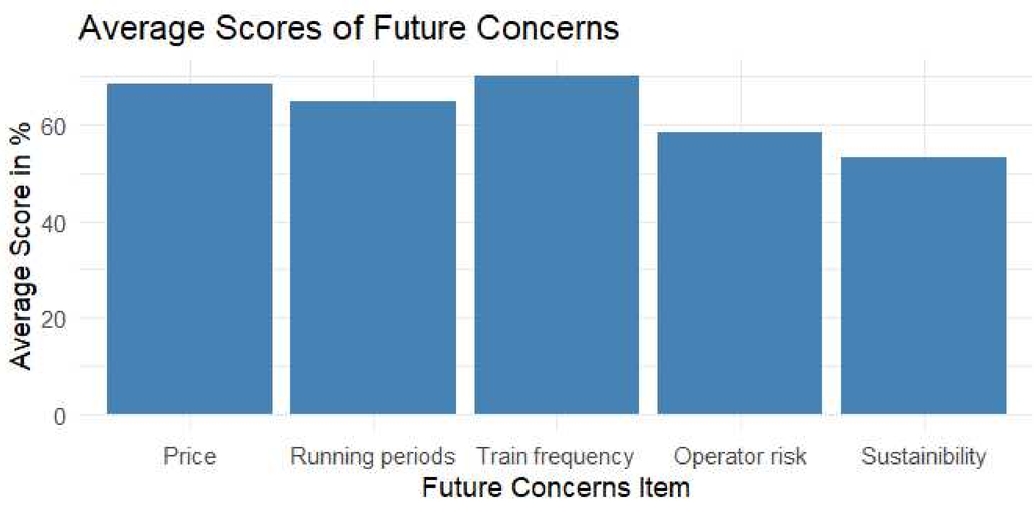
Figure 1

Mean Results of Items on Rail Quality

To understand the importance of certain changes because of future autonomous trains, the results of these items were also put into a bar chart (Figure 2). This chart shows the means of how the participants rated the importance of the factors "Reduced ticket price", "Extended running periods", "Increased train frequency", "Reduced risk for human operators", and "Sustainability". This chart shows that the factor that received the highest mean was "Train Frequency" (M = 70, SD = 22.50), and the lowest on "Sustainability" (M = 53, SD = 25.96). Only "Increased train frequency" received no scores of 1.

Figure 2

Means for Item on Changes of Future Autonomous Trains



In order to compare previous results to those of the current study, we also investigated the importance of certain aspects that are part of autonomous trains in deciding to use autonomous trains (Table 2). In this table, the results were shown of how important it is for the participant that certain aspects are part of the autonomous trains in percentages. These aspects were “Presence of a train attendant who is able to drive the train”, “Presence of a train attendant in the driver cab who is able to drive the train”, “Having information available through screens”, and “Having information available through train personnel”. It can be seen that the aspect with which was rated 100% the most was “Presence of a train attendant who is able to drive the train” (28.3%), and the aspect which was rated 100% the least was that of “Having information available through train personnel” (9.4%). The aspect with the highest mean score was “Having information available through screens” (M = 66, SD = 26.90), and the aspect with the lowest mean was “Presence of a train attendant in the driver cab who is able to drive the train” (M = 48, SD = 33.91).

The data gathered for comparison were the items on automation level preference (Table 3). This table shows what preference the participants had when it comes to the level of automation of future trains. The first item asked if the participant preferred to ride in an autonomous train, a driver train, or if there is no preference. The second item asked the preference with regard to the level of automation (GoA 1-4). This table shows that 66% of people do not have a specific preference for the type of train (M = 80, SD = 29.97). The largest part of the participants also indicated that they preferred GoA type 3, with 58.5% of people choosing this (M = 52, SD = 24.03).

Table 3

Level of Automation Preference

Type of Train	**Autonomous**	**Human driver**	**No preference**		**Mean in %**	**SD in %**
Times chosen	3 (5.7%)	15 (28.3%)	35 (66.0%)		80	29.97
Level of automation	**GoA 1**	**GoA 2**	**GoA 3**	**GoA 4**	**Mean in %**	**SD in %**
Times chosen	5 (9.4%)	15 (28.3%)	31 (58.5%)	2 (3.8%)	52	24.03

The responses to the Virtual Reality videos were put into a table to see if there was an observable difference between the given answers (Table 4). The responses showed that in condition 1 the majority of participants guessed that the human was driving in the bumpy train experience (60%), while in condition 2 the majority stated that they believed the human to be driving in the more comfortable experience (67.9%). Overall, the slight majority of people guessed that the human was driving in the comfortable experience (54.7%). The reasoning for the answers to this question for most participants came down to the ride being bumpy/shaky, this answer was popular for assigning the human driver both to the bad and the good experience. It was also a common answer that the participant did not know or just guessed.

Table 4

Participants' Answer to the Question about Which Train was Driven by a Human.

Number of people who believed the train experience was driven by a human	Survey First (Condition 1)	Video First (Condition 2)	Combined
Human driver is causing a bumpy train experience.	15 (60%)	9 (32.1%)	24 (45.3%)
Human driver is causing a comfortable train experience.	10 (40%)	16 (67.9%)	29 (54.7%)
total	25 (100%)	28 (100%)	53 (100%)

A chi-square test suggested no differences due to the conditions between the attribution to a bad or a good journey more to humans or autonomous driven train $p > .05$.

Is intention to use predicted by declared trust towards autonomous trains, participant age, their attitude and their willingness to take risks (RQ1)?

Descriptive statistics was performed by contingency tables to observe the answers people gave to the questions on willingness to take risks, and their associated intention to use automated train in the future (Appendix B). Overall, most participants rate themselves a 6 (22.6%) on a scale of 1-10 for risk attitude with no one rating themselves a 10, and most people tending towards the middle of the scale scoring between 4 and 7 (M = 5.434, SD = 1.814). For condition 1, the most popular response was also 6 (28%) with the lowest chosen score being 1 and the highest being 8 (M = 5.36, SD = 1.705). In condition 2, the option that was chosen most was 5 with 1 also being the lowest and 9 being the highest (M = 5.536, SD = 2.009).

For future intention to use autonomous trains, most people rated themselves a 50 on the scale from 1-5 with 19 participants (35.8%) choosing this option (M = 3.245, SD = 1.207). In condition 1, the most people picked 4 (32.0%) out of 25 participants (M = 3.320, SD = 1.345). For condition 2, most people opted for option 3 with 13 people (46.4%) choosing this out of 28 (M = 3.179, SD = 1.090).

To answer the research question, a multiple linear regression model was made between the variables "Intention to use autonomous trains" created from the first item on willingness to use automated trains in the future, and the variable "Risk" created from the item on risk taking behaviour, "Age", "Attitude", and "Trust". This model showed a significant predictive effect for Intention to use autonomous trains in the future $R^2 = .158$, $F(1, 48) = 3.44$, $p = .015$ and the results can be seen in Table 5.

Table 5

Results from a Multiple Linear Regression Model Including "Intention", "Attitude", "Trust", "Risk", and "Age" Predicting for Intention to Use

Coefficients	Estimate	SD	t-value	p-value
Intention	-6.09	19.451	-0.313	0.756
Attitude	0.35	0.193	1.810	0.077
Trust	0.215	0.272	0.791	0.433
Risk	0.402	0.216	1.864	0.068
Age	0.163	0.216	0.756	0.453

The model showed significance, but we were not able to determine which of the factors caused this significance from this model alone. Because of this, a second model was created with just the variables “Attitude” and “Risk” compared to the variable “Intention to use autonomous trains” as these factors showed p-values very close to .05. The results of this model showed even higher significance (Table 6) with an overall p-value of .003.

Table 6

Results from a Multiple Linear Regression Model Including “Attitude” and “Risk” Predicting for Intention to Use

Coefficients	Estimate	SD	t-value	p-value
Intention	7.703	14.233	0.541	0.591
Attitude	0.398	0.163	2.449	0.018
Risk	0.409	0.211	1.936	0.059

Effects of previous experience with autonomy on trust and attitude towards future trains (RQ2)

To explore if there is a significant effect between having previous experience with autonomous trains and the level of trust in autonomous trains, a multiple linear regression model was created consisting of the average trust and the attitude towards autonomous trains compared to previous experience. This model also showed no significant effect, $p > .05$.

Presenting video before or after the survey affected people answers regarding the intention to use (RQ3)

To investigate if the difference in the Conditions for the survey had any significant effect on the intention to use autonomous trains, a MANOVA that combined the average trust and intention to use compared to the 2 conditions was also made. This MANOVA revealed that there exists no significant differences due to the conditions in terms of the scores on trust and intention to use $p > .05$.

Discussion

This study provided insight into the perception of public expectations toward autonomous train experiences, investigating the effects of multiple factors, such as Age, Attitude, Concerns, Experience, and Trust.

The first part of this research investigated if the current findings are in line with previous research. In earlier research, it is stated that the majority of people support the future use of autonomous trains (Cogan et al., 2022; Morast et al., 2023; Pakusch & Bossauer, 2017), this is in line with the current findings as 79.2% stated that they are at the least not opposed to the use of autonomous trains. According to Arzer et al. (2024) and Fraszczyk and Mulley (2017) the most important factors surrounding the current concerns are Safety and Reliability, as many people also mention a fear of cyberattacks (Morast et al., 2023). Cogan et al. (2022) also states the importance of providing information through either human onboard or external personnel (Arzer et al., 2024). The current research further supports these findings, except for the fear of cyberattacks. The factor Safety and Reliability scored the highest of the measured aspects with a mean of 87.26% for Safety and a mean of 81.10 for Reliability, showing that these factors are of high importance when considering autonomous trains, but only 34.43% mentioned that they worry about cyberattacks. The presence of onboard personnel with a mean score of 53.77 was not as important to the participants as the use of digital screens for providing information with a mean of 66.04. In relation to the presence of onboard personnel, previous studies by Cogan et al. (2022) and Lemonnier et al. (2023) say that the majority of people prefer GoA-3 which describes a state of automation in which the train is fully driven by autonomous systems but there is still a human present who can take over in case of emergency. This study found results similar to those of Cogan et al. (2022) and Lemonnier et al. (2023) with 58.5% preferring GoA-3 of any other GoA, and 63.68% of participants stating that they prefer the presence of a human who is able to drive the train if needed. Regarding the automation of driving the train, Fraszczyk et al. (2015) found that 62% were concerned about the losses in jobs for train drivers, but this outcome was not reproduced in the current study as 56.60% said that it is not a concern to them. With regards to the benefits of the automation of trains, Cogan et al. (2022) states that a reduction in ticket costs and relay reductions are of high importance, but Fraszczyk et al. (2015) opposes this by saying that ticket prices, extended running periods, and increased train frequencies are not considered important enough. According to the current study, an increase in train frequency was considered as a strong benefit with a mean score of 70.28%, reduced ticket prices scored a mean of 68.40, and extended running periods resulted in a mean of 65.09 showing that people do find it important, which is in accordance with the results found by Cogan et al. (2022).

To answer the first research question, A multiple linear regression model with the factors “attitude”, “age”, “risk”, and “trust” showed a significant effect for predicting the intention to use autonomous trains. A second models was created for “attitude” and ”risk”. This model showed an even more significant result, indicating that a model including just “attitude” and “risk” can more accurately predict for the willingness to use autonomous trains in the future without the factors “age” and “trust”. These results are in line with earlier findings by Pakusch and Bossauer (2017) who also found no significance for age, but opposed to Morast et al. (2023) who did find that age had a significant effect.

In order to answer the second research question, does previous experience with autonomous vehicles affect people’s declared trust towards autonomous trains and their attitude towards trains. A multiple linear regression model was created combining the effect of “experience” on both “trust” and attitude”. This model indicated no significance, meaning that experience does not affect people’s trust and attitude surrounding autonomous vehicles. These outcomes were different from previous research done by Cogan et al. (2022); Morast et al. (2023); Pakusch and Bossauer (2017), as they did find that people who already experienced an autonomous vehicle did show higher willingness to use autonomous vehicles in the future compared to people without previous experience.

To answer the final research question, does seeing the video before or after the survey affect peoples’ declared level of trust and intention of use, a MANOVA combining both the factors of “intention to use” and “trust” compared to the conditions indicates that there is no meaningful relation between seeing the video at the beginning or at the end of the survey and the level of trust and intention of use that the participant reported on the items in the survey. This shows that the videos do not prime or predispose the participant to certain opinions surrounding autonomous trains.

Improvements for future research

This study encountered multiple limitations during its research which can be learned from to improve future research. The biggest limitation is that the current study is a pilot study, meaning that most of the research is still exploratory. Because of this, it is difficult to say with certainty that the results found in this study are fully 100% accurate and comparable with the public opinion. As there are still some factors on which there is not yet a consensus on their significance and importance, these need to be further analyzed in the future to determine their role in the public perception of autonomous trains. These factors include among others, Age, Previous experience with autonomous trains, and the dangers of Cyberattacks. Another important limitation is that the setting in which the participants completed the survey was not standardized. This makes it uncertain if the surrounding environment was equal for all participants at the time of doing the survey. Because of this, it is possible that the answers that the participants gave on the survey are not entirely unbiased or without any outside influence. For example, it is possible that the participant got distracted or bored during the survey and decided to just give "random" answers instead of answering the questions truthfully. Therefore, it is recommended that when performing these researches in the future, the environment in which the participants take part in the study are equal across all participants to exclude unknown outside influences, as this would result in higher validity for the results. A final limitation is the relatively small sample size. A sample size of 53 should statistically be enough for any significant outcomes, however, it is still small compared to the overall public. Therefore it is possible that the results of this study are not representative of the public view. A larger sample size is almost always recommended to make the sample size more representative of the actual population which is why a study should aim for as many participants as possible.

Conclusion

This study found results that were in some cases similar to those of previous studies, such as the role of concerns surrounding autonomous trains and the preferred grade of automation (GoA-3) where humans are still present but the trains is driven autonomously. However, not all results were in line with previous findings, as this study did not find significant effects for the factors of Previous experience or the roll of age, unlike studies done by Cogan et al. (2022); Morast et al. (2023); Pakusch and Bossauer (2017) who did encounter significance. The research performed in this study did find significant effect from willingness to take risks on the intention to use autonomous trains in the future. The 2 different conditions that the participants could be put in did not indicate to have any influence on their trust and intention of use. The limitations of this study mostly consisted of it being a pilot study, having an unregulated environment for most participants, and the relatively small sample size. Because of all this, it is recommended that for future research, the role of the factors that possibly influence the perception of autonomous trains are further investigated, whilst making sure that the sample size is much larger, and the environment is made equal for all participants.

References

Arzer, A., Beehler, L., & Vredenborg, M. (2024, 2024//). Next Stop: Passenger Perspectives on Autonomous Trains. HCI in Mobility, Transport, and Automotive Systems, Cham.

Cogan, B., Tandetzki, J., & Milius, B. (2022). Passenger Acceptability of Teleoperation in Railways. *Future Transportation*, *2*(4), 956-969. https://www.mdpi.com/2673-7590/2/4/53

Detjen, H., Nurhas, I., & Geisler, S. (2021). Attitudes towards autonomous public transportation. 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications,

EU-Rail. (2021). *Europe's Rail Mission and Objectives*. European Union. Retrieved 8 February 2025 from https://rail-research.europa.eu/about-europes-rail/europes-rail-mission-and-objectives/

Fraszczyk, A., Brown, P., & Duan, S. (2015). Public perception of driverless trains. *Urban Rail Transit*, *1*, 78-86.

Fraszczyk, A., & Mulley, C. (2017). Public perception of and attitude to driverless train: a case study of Sydney, Australia. *Urban Rail Transit*, *3*(2), 100-111.

Hilgarter, K., & Granig, P. (2020). Public perception of autonomous vehicles: A qualitative study based on interviews after riding an autonomous shuttle. *Transportation Research Part F: Traffic Psychology and Behaviour*, *72*, 226-243. https://doi.org/10.1016/j.trf.2020.05.012

Lemonnier, A., Adelé, S., & Dionisio, C. (2020). The determinants of acceptability and behavioural intention of automated vehicles - A review. *Le travail humain*, *83*, 297. https://doi.org/10.3917/th.834.0297

Lemonnier, A., Adelé, S., & Dionisio, C. (2023). Acceptability of autonomous trains with different grades of automation by potential users: A qualitative approach. *Travel Behaviour and Society*, *33*, 100641. https://doi.org/10.1016/j.tbs.2023.100641

Morast, A., Voß, G. M. I., Dautzenberg, P. S. C., Urban, P., & Nießen, N. (2023). A survey on the acceptance of unattended trains. *Journal of Rail Transport Planning & Management*, *25*, 100370. https://doi.org/10.1016/j.jrtpm.2023.100370

Pakusch, C., & Bossauer, P. (2017). User Acceptance of Fully Autonomous Public Transport. *ICE-B*, *2*, 52-60.

Palm, V. (2022). *Key figures on European transport*. Publications Office of the European Union. https://ec.europa.eu/eurostat/documents/15216629/15589759/KS-07-22-523-EN-N.pdf

Reichmann, M., Himmelbauer, G. S., Wagner, A., Zajicek, J., Stadlmann, B., Wancura, H., Furian, N., & Vössner, S. (2025). Introducing the concept of grades of automation for shunting operations. *Journal of Rail Transport Planning & Management*, *33*, 100500. https://doi.org/10.1016/j.jrtpm.2024.100500

Appendix A

The pilot survey

This appendix includes the link to a preview of the pilot survey used in the second part of this study.

Link to survey: Pilot survey in Qualtrics

Appendix B

Results on Willingness to Take Risks and Intention to Use

Table 7

Results from Items on Risk Behaviour and Future Intention

Willingness to take risks	Condition 1	Condition 2	Combined
10	1 (4.0%)	2 (7.1%)	3 (5,7%)
20	1 (4.0%)	0 (0.0%)	1 (1.9%)
30	1 (4.0%)	1 (3.6%)	2 (3.8%)
40	4 (16.0%)	4 (14.3%)	8 (15.1%)
50	4 (16.0%)	7 (25.0%)	11 (20.8%)
60	7 (28.0%)	5 (17.9%)	12 (22.6%)
70	6 (24.0%)	5 (17.9%)	11 (20.8%)
80	1 (4.0%)	3 (10.7%)	4 (7.5%)
90	0 (0.0%)	1 (3.6%)	1 (1.9%)
100	0 (0.0%)	0 (0.0%)	0 (0%)
Total	25 (100%)	28 (100%)	53 (100%)
Mean	53	55	54
SD	17.05	20.09	18.56
Intention to use automated trains in the future	**Condition 1**	**Condition 2**	**Combined**
0	5 (16.0%)	3 (10.7%)	7 (13.2%)
25	2 (8.0%)	2 (7.1%)	4 (5.5%)

50	6 (24.0%)	13 (46.4%)	19 (35.8%)
75	8 (32.0%)	7 (25.0%)	15 (28.3%)
100	5 (20.0%)	3 (10.7%)	8 (15.1%)
Total	25 (100%)	28 (100%)	53 (100%)
Mean	58%	54%	56%
SD	33.63	27.26	30.18

Appendix C

Rstudio Code for Data Analysis

```
#loading libraries
library(janitor)
library(tidyverse)
library(dplyr)
library(ggplot2)
library(ltm)
library(tidyr)
#importing dataset
Full_Dataset <- read.csv("Survey Dataset 1.csv")
Final_Dataset <- read.csv("Thesis_Full_Dataset.csv")
#removing unnecessary columns and rows
Full_Dataset <- Full_Dataset[, -c(3, 4, 9:17)]
Final_Dataset <- Final_Dataset[, -c(3, 4, 9:17)]
#removing unfinished attempts
Final_Dataset <- Final_Dataset[-c(1, 2, 40, 53, 58:61), ]
#Descriptive statistics
Final_Dataset %>%
  tabyl(D2_age)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset %>%
  tabyl(D4_Gender)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset %>%
  tabyl(D1_nation)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset$D2_age <- as.numeric(Final_Dataset$D2_age)
```

```
Final_Dataset$D2_age %>% summary()
Final_Dataset$D2_age %>% sd()
#SETTING UP VARIABLES FOR ANALYSES
#Adding column for expected situation as driver or autonomous
#1 represent the participant guessed the good situation is driven by a human
#0 represent the participant guessed the good situation is autonomous
Final_Dataset$Bet.A.Bad.B.Good_1 <- recode(Final_Dataset$Bet.A.Bad.B.Good_1,
                    "2" = 1,
                    "1" = 0
)
Final_Dataset$Bet.A.Bad.B.Good_1.1 <- recode(Final_Dataset$Bet.A.Bad.B.Good_1.1,
                    "2" = 0,
                    "1" = 1
)
#Creating a new variable with the Condition merged so that there are no missing values
Final_Dataset$Merged <- ifelse(!is.na(Final_Dataset$Bet.A.Bad.B.Good_1),
Final_Dataset$Bet.A.Bad.B.Good_1, Final_Dataset$Bet.A.Bad.B.Good_1.1)
Final_Dataset %>%
  tabyl(Merged)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset$Video_Explan <- ifelse(!is.na(Final_Dataset$explan1) & Final_Dataset$explan1 != "",
                   Final_Dataset$explan1,
                   Final_Dataset$Explan2
                   )
#transforming intention + risk taking to use scales to %scales
Final_Dataset$D6_Risk.Attitude <- as.numeric(Final_Dataset$D6_Risk.Attitude)
Final_Dataset$FC1_Intention.to.use_1 <- as.numeric(Final_Dataset$FC1_Intention.to.use_1)
Final_Dataset$FC1_Intention.to.use_2 <- as.numeric(Final_Dataset$FC1_Intention.to.use_2)
Final_Dataset$Intention1 <- (Final_Dataset$FC1_Intention.to.use_1 -1) / 4*100
Final_Dataset$Intention2 <- (Final_Dataset$FC1_Intention.to.use_2 -1) / 4*100
Final_Dataset$Risk <- (Final_Dataset$D6_Risk.Attitude) / 10*100
#seperating the dataset into the 2 conditions
#Condition 1 is the video at the beginning, condition 2 is the video in the end
Condition1 <- Final_Dataset[Final_Dataset$Condition == 1, ]
```

```
Condition2 <- Final_Dataset[Final_Dataset$Condition == 2, ]
#checking the responses to the video question per condition
Condition1 %>%
  tabyl(Merged)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Condition2 %>%
  tabyl(Merged)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
#creating tables for the willingess to use and the willingness to take risks
Final_Dataset$D6_Risk.Attitude <- as.numeric(Final_Dataset$D6_Risk.Attitude)
str(Final_Dataset$FC1_Intention.to.use_1)
str(Final_Dataset$D6_Risk.Attitude)
Final_Dataset %>%
  tabyl(FC1_Intention.to.use_1)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset$FC1_Intention.to.use_1 %>% summary()
Final_Dataset$FC1_Intention.to.use_1 %>% sd()
Final_Dataset %>%
  tabyl(D6_Risk.Attitude)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset$D6_Risk.Attitude %>% summary()
Final_Dataset$D6_Risk.Attitude %>% sd()
Condition1 %>%
  tabyl(FC1_Intention.to.use_1)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Condition1$FC1_Intention.to.use_1 %>% summary()
```

```
Condition1$FC1_Intention.to.use_1 %>% sd()
Condition1 %>%
  tabyl(D6_Risk.Attitude)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Condition1$D6_Risk.Attitude %>% summary()
Condition1$D6_Risk.Attitude %>% sd()
Condition2 %>%
  tabyl(FC1_Intention.to.use_1)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Condition2$FC1_Intention.to.use_1 %>% summary()
Condition2$FC1_Intention.to.use_1 %>% sd()
Condition2 %>%
  tabyl(D6_Risk.Attitude)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Condition2$D6_Risk.Attitude %>% summary()
Condition2$D6_Risk.Attitude %>% sd()
#creating multiple linear regression models
model1 <- lm(Intention1 ~ FC2_attitude_1 + Av_Trust + Risk + D2_age, data = Final_Dataset)
summary(model1)
#modeling intention to use with factors
model2 <- lm(Intention1 ~ FC2_attitude_1 + Risk, data = Final_Dataset)
summary(model2)
#Creating MANOVA
manova_result <- manova(cbind(Av_Trust, Intention1) ~ Condition, data = Final_Dataset)
summary(manova_result)
summary.aov(manova_result)
#running a t-test
#t-test for intention to use between condition 1 and 2
t.test(Condition1$FC1_Intention.to.use_1, Condition2$FC1_Intention.to.use_1)
t.test(Final_Dataset$Intention1 ~ Final_Dataset$Condition)
```

```
t.test(Final_Dataset$Intention2 ~ Final_Dataset$Condition)
t.test(Final_Dataset$Risk ~ Final_Dataset$Condition)
#Logistic model
#comparing video answer with the condition
Final_Dataset$Condition <- as.numeric(Final_Dataset$Condition)
Final_Dataset$D6_Risk.Attitude <- as.factor(Final_Dataset$D6_Risk.Attitude)
Final_Dataset$situation <- (Final_Dataset$Condition -1)
Logistic_Model <- glm(Final_Dataset$Merged ~ Final_Dataset$Situation, family = binomial)
summary(Logistic_Model)
#Comparing the video answer to the condition USE A CHI SQUARE TEST
chisq.test(table(Final_Dataset$Situation, Final_Dataset$Merged))
# Create the contingency table
train_data <- matrix(c(15, 10, 9, 16),
           nrow = 2,
           byrow = FALSE,
           dimnames = list(c("Bumpy", "Comfortable"),
                   c("Condition1", "Condition2")))
# View the table train_data
train_data %>% View()
# Perform chi-square test
chisq.test(train_data)
# If you want expected frequencies
chisq.test(train_data)$expected
# For additional details including standardized residuals
chisq.test(train_data)$residuals
#creating a column, 1 for people with prior experience and 0 for people without
#people who responded 4 indicated no prior experience with autonomous vehicles
Final_Dataset$D7_expereince.auton <- as.numeric(Final_Dataset$D7_expereince.auton)
Final_Dataset$Experience <- ifelse(Final_Dataset$D7_expereince.auton == 4,
                  "0",
                  "1")
#running a t-test + linear model between intention to use differing in experience
t.test(Final_Dataset$Intention1 ~ Final_Dataset$Experience)
model4 <- lm(Intention1 ~ Experience, data = Final_Dataset)
summary(model4)
#running a linear model between declared trust differing in experience
```

```
#creating an average score for trust (1,2,3,6,7,8 are inverted as they indicate DIStrust)
Final_Dataset$TH5_Trust_1 <- as.numeric(Final_Dataset$TH5_Trust_1)
Final_Dataset$TH5_Trust_2 <- as.numeric(Final_Dataset$TH5_Trust_2)
Final_Dataset$TH5_Trust_3 <- as.numeric(Final_Dataset$TH5_Trust_3)
Final_Dataset$TH5_Trust_4 <- as.numeric(Final_Dataset$TH5_Trust_4)
Final_Dataset$TH5_Trust_5 <- as.numeric(Final_Dataset$TH5_Trust_5)
Final_Dataset$TH5_Trust_6 <- as.numeric(Final_Dataset$TH5_Trust_6)
Final_Dataset$TH5_Trust_7 <- as.numeric(Final_Dataset$TH5_Trust_7)
Final_Dataset$TH5_Trust_8 <- as.numeric(Final_Dataset$TH5_Trust_8)
#create a code for reversing scores on Trust_1:3,6:8
reverse_code <- function(x) {
  return(6 - x)
}
# Apply reverse coding where needed
Final_Dataset$TH5_Trust_1 <- reverse_code(Final_Dataset$TH5_Trust_1)
Final_Dataset$TH5_Trust_2 <- reverse_code(Final_Dataset$TH5_Trust_2)
Final_Dataset$TH5_Trust_3 <- reverse_code(Final_Dataset$TH5_Trust_3)
Final_Dataset$TH5_Trust_6 <- reverse_code(Final_Dataset$TH5_Trust_6)
Final_Dataset$TH5_Trust_7 <- reverse_code(Final_Dataset$TH5_Trust_7)
Final_Dataset$TH5_Trust_8 <- reverse_code(Final_Dataset$TH5_Trust_8)
#changing the scores of trust into percentages
Final_Dataset$TH5_Trust_1 <- (Final_Dataset$TH5_Trust_1 -1) / 4*100
Final_Dataset$TH5_Trust_2 <- (Final_Dataset$TH5_Trust_2 -1) / 4*100
Final_Dataset$TH5_Trust_3 <- (Final_Dataset$TH5_Trust_3 -1) / 4*100
Final_Dataset$TH5_Trust_4 <- (Final_Dataset$TH5_Trust_4 -1) / 4*100
Final_Dataset$TH5_Trust_5 <- (Final_Dataset$TH5_Trust_5 -1) / 4*100
Final_Dataset$TH5_Trust_6 <- (Final_Dataset$TH5_Trust_6 -1) / 4*100
Final_Dataset$TH5_Trust_7 <- (Final_Dataset$TH5_Trust_7 -1) / 4*100
Final_Dataset$TH5_Trust_8 <- (Final_Dataset$TH5_Trust_8 -1) / 4*100
#Creating a column with the average of all 8 trust items
Final_Dataset$Av_Trust <- rowMeans(Final_Dataset[, paste0("TH5_Trust_", 1:8)], na.rm = TRUE)
#Creating linear regression models for trust with intention to use and previous experience
model_experience <- lm(Experience ~ Av_Trust + FC2_attitude_1, data = Final_Dataset)
model_experience %>% summary()
#intention to use per condition
Condition1 %>%
```

```
  tabyl(FC2_attitude_1)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Condition2 %>%
  tabyl(FC2_attitude_1)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
#creating demographics for current and future Transportation Habits
#Setting values to numeric instead of character
#Rail quality
Final_Dataset$TH2_Rail_quality_1 <- as.numeric(Final_Dataset$TH2_Rail_quality_1)
Final_Dataset$TH2_Rail_quality_2 <- as.numeric(Final_Dataset$TH2_Rail_quality_2)
Final_Dataset$TH2_Rail_quality_3 <- as.numeric(Final_Dataset$TH2_Rail_quality_3)
Final_Dataset$TH2_Rail_quality_4 <- as.numeric(Final_Dataset$TH2_Rail_quality_4)
Final_Dataset$TH2_Rail_quality_5 <- as.numeric(Final_Dataset$TH2_Rail_quality_5)
Final_Dataset$TH2_Rail_quality_6 <- as.numeric(Final_Dataset$TH2_Rail_quality_6)
#Future Concerns
Final_Dataset$TH3_DriverlessF_R_1 <- as.numeric(Final_Dataset$TH3_DriverlessF_R_1)
Final_Dataset$TH3_DriverlessF_R_2 <- as.numeric(Final_Dataset$TH3_DriverlessF_R_2)
Final_Dataset$TH3_DriverlessF_R_3 <- as.numeric(Final_Dataset$TH3_DriverlessF_R_3)
Final_Dataset$TH3_DriverlessF_R_4 <- as.numeric(Final_Dataset$TH3_DriverlessF_R_4)
Final_Dataset$TH3_DriverlessF_R_5 <- as.numeric(Final_Dataset$TH3_DriverlessF_R_5)
#Future needs
Final_Dataset$TH4_factors_1 <- as.numeric(Final_Dataset$TH4_factors_1)
Final_Dataset$TH4_factors_2 <- as.numeric(Final_Dataset$TH4_factors_2)
Final_Dataset$TH4_factors_3 <- as.numeric(Final_Dataset$TH4_factors_3)
Final_Dataset$TH4_factors_4 <- as.numeric(Final_Dataset$TH4_factors_4)
#Support
Final_Dataset$FC2_attitude_1 <- as.numeric(Final_Dataset$FC2_attitude_1)
#Changing scores to percentages
#Intention to use
Final_Dataset$FC1_Intention.to.use_1 <- (Final_Dataset$FC1_Intention.to.use_1 -1) / 4*100
Condition1$FC1_Intention.to.use_1 <- (Condition1$FC1_Intention.to.use_1 -1) / 4*100
Condition2$FC1_Intention.to.use_1 <- (Condition2$FC1_Intention.to.use_1 -1) / 4*100
```

```
#Rail quality
Final_Dataset$TH2_Rail_quality_1 <- (Final_Dataset$TH2_Rail_quality_1 -1) / 4*100
Final_Dataset$TH2_Rail_quality_2 <- (Final_Dataset$TH2_Rail_quality_2 -1) / 4*100
Final_Dataset$TH2_Rail_quality_3 <- (Final_Dataset$TH2_Rail_quality_3 -1) / 4*100
Final_Dataset$TH2_Rail_quality_4 <- (Final_Dataset$TH2_Rail_quality_4 -1) / 4*100
Final_Dataset$TH2_Rail_quality_5 <- (Final_Dataset$TH2_Rail_quality_5 -1) / 4*100
Final_Dataset$TH2_Rail_quality_6 <- (Final_Dataset$TH2_Rail_quality_6 -1) / 4*100
#Future Concerns
Final_Dataset$TH3_DriverlessF_R_1 <- (Final_Dataset$TH3_DriverlessF_R_1 -1) / 4*100
Final_Dataset$TH3_DriverlessF_R_2 <- (Final_Dataset$TH3_DriverlessF_R_2 -1) / 4*100
Final_Dataset$TH3_DriverlessF_R_3 <- (Final_Dataset$TH3_DriverlessF_R_3 -1) / 4*100
Final_Dataset$TH3_DriverlessF_R_4 <- (Final_Dataset$TH3_DriverlessF_R_4 -1) / 4*100
Final_Dataset$TH3_DriverlessF_R_5 <- (Final_Dataset$TH3_DriverlessF_R_5 -1) / 4*100
#Future needs
Final_Dataset$TH4_factors_1 <- (Final_Dataset$TH4_factors_1 -1) / 4*100
Final_Dataset$TH4_factors_2 <- (Final_Dataset$TH4_factors_2 -1) / 4*100
Final_Dataset$TH4_factors_3 <- (Final_Dataset$TH4_factors_3 -1) / 4*100
Final_Dataset$TH4_factors_4 <- (Final_Dataset$TH4_factors_4 -1) / 4*100
#Support
Final_Dataset$FC2_attitude_1 <- (Final_Dataset$FC2_attitude_1 -1) / 4*100
#safety
Final_Dataset %>%
  tabyl(TH2_Rail_quality_1)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset$TH2_Rail_quality_1 %>% summary()
Final_Dataset$TH2_Rail_quality_1 %>% sd()
#Reliability
Final_Dataset %>%
  tabyl(TH2_Rail_quality_2)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset$TH2_Rail_quality_2 %>% summary()
Final_Dataset$TH2_Rail_quality_2 %>% sd()
```

```
#Sustainability
Final_Dataset %>%
  tabyl(TH2_Rail_quality_3)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset$TH2_Rail_quality_3 %>% summary()
Final_Dataset$TH2_Rail_quality_3 %>% sd()
#Accessibility
Final_Dataset %>%
  tabyl(TH2_Rail_quality_4)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset$TH2_Rail_quality_4 %>% summary()
Final_Dataset$TH2_Rail_quality_4 %>% sd()
#Comfort
Final_Dataset %>%
  tabyl(TH2_Rail_quality_5)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset$TH2_Rail_quality_5 %>% summary()
Final_Dataset$TH2_Rail_quality_5 %>% sd()
#Ticket prices
Final_Dataset %>%
  tabyl(TH2_Rail_quality_6)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset$TH2_Rail_quality_6 %>% summary()
Final_Dataset$TH2_Rail_quality_6 %>% sd()
#Creating Bar charts
Long_Concerns <- Final_Dataset %>%
  pivot_longer(
    cols = starts_with("TH3_DriverlessF_R_"),
```

```
    names_to = "Future_Concerns_Item",
    values_to = "Score"
  )
ggplot(Long_Concerns, aes(x = Future_Concerns_Item, y = Score)) +
  stat_summary(fun = mean, geom = "bar", fill = "steelblue") +
  scale_x_discrete(labels = c(
    "TH3_DriverlessF_R_1" = "Price",
    "TH3_DriverlessF_R_2" = "Running periods",
    "TH3_DriverlessF_R_3" = "Train frequency",
    "TH3_DriverlessF_R_4" = "Operator risk",
    "TH3_DriverlessF_R_5" = "Sustainibility"
  )) +
  labs(title = "Average Scores of Future Concerns",
       x = "Future Concerns Item",
       y = "Average Score in %") +
  theme_minimal()
Long_Rail <- Final_Dataset %>%
  pivot_longer(
    cols = starts_with("TH2_Rail_quality_"),
    names_to = "Rail_Quality_Item",
    values_to = "Score"
  )
ggplot(Long_Rail, aes(x = Rail_Quality_Item, y = Score)) +
  stat_summary(fun = mean, geom = "bar", fill = "steelblue") +
  scale_x_discrete(labels = c(
    "TH2_Rail_quality_1" = "Safety",
    "TH2_Rail_quality_2" = "Reliability",
    "TH2_Rail_quality_3" = "Sustainability",
    "TH2_Rail_quality_4" = "Accessibility",
    "TH2_Rail_quality_5" = "Comfort",
    "TH2_Rail_quality_6" = "Cost"
  )) +
  labs(title = "Average Scores of Rail Quality Items",
       x = "Rail Quality Item",
       y = "Average Score in %") +
  theme_minimal()
```

```
#creating tables for Future Transportation changes
#Setting values to numeric instead of character
Final_Dataset$TH3_DriverlessF_R_1 <- as.numeric(Final_Dataset$TH3_DriverlessF_R_1)
Final_Dataset$TH3_DriverlessF_R_2 <- as.numeric(Final_Dataset$TH3_DriverlessF_R_2)
Final_Dataset$TH3_DriverlessF_R_3 <- as.numeric(Final_Dataset$TH3_DriverlessF_R_3)
Final_Dataset$TH3_DriverlessF_R_4 <- as.numeric(Final_Dataset$TH3_DriverlessF_R_4)
Final_Dataset$TH3_DriverlessF_R_5 <- as.numeric(Final_Dataset$TH3_DriverlessF_R_5)
#Reduced ticket price
Final_Dataset %>%
  tabyl(TH3_DriverlessF_R_1)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset$TH3_DriverlessF_R_1 %>% summary()
Final_Dataset$TH3_DriverlessF_R_1 %>% sd()
#Extended running periods
Final_Dataset %>%
  tabyl(TH3_DriverlessF_R_2)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset$TH3_DriverlessF_R_2 %>% summary()
Final_Dataset$TH3_DriverlessF_R_2 %>% sd()
#Increased train frequency
Final_Dataset %>%
  tabyl(TH3_DriverlessF_R_3)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset$TH3_DriverlessF_R_3 %>% summary()
Final_Dataset$TH3_DriverlessF_R_3 %>% sd()
#Reduced risk for human operators
Final_Dataset %>%
  tabyl(TH3_DriverlessF_R_4)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
```

```
  view()
Final_Dataset$TH3_DriverlessF_R_4 %>% summary()
Final_Dataset$TH3_DriverlessF_R_4 %>% sd()
#Sustainability
Final_Dataset %>%
  tabyl(TH3_DriverlessF_R_5)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset$TH3_DriverlessF_R_5 %>% summary()
Final_Dataset$TH3_DriverlessF_R_5 %>% sd()
#creating tables for importance of Future Transportation Aspects
#Setting values to numeric instead of character
Final_Dataset$TH4_factors_1 <- as.numeric(Final_Dataset$TH4_factors_1)
Final_Dataset$TH4_factors_2 <- as.numeric(Final_Dataset$TH4_factors_2)
Final_Dataset$TH4_factors_3 <- as.numeric(Final_Dataset$TH4_factors_3)
Final_Dataset$TH4_factors_4 <- as.numeric(Final_Dataset$TH4_factors_4)
#Attendant availlable who can drive
Final_Dataset %>%
  tabyl(TH4_factors_1)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset$TH4_factors_1 %>% summary()
Final_Dataset$TH4_factors_1 %>% sd()
#Attendant availlable in the cab
Final_Dataset %>%
  tabyl(TH4_factors_2)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset$TH4_factors_2 %>% summary()
Final_Dataset$TH4_factors_2 %>% sd()
#Information through screens
Final_Dataset %>%
  tabyl(TH4_factors_3)%>%
```

```
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset$TH4_factors_3 %>% summary()
Final_Dataset$TH4_factors_3 %>% sd()
#Information through attendant
Final_Dataset %>%
  tabyl(TH4_factors_4)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset$TH4_factors_4 %>% summary()
Final_Dataset$TH4_factors_4 %>% sd()
#Level of automation preference
#Type of train
Final_Dataset$ALP1_Preference <- as.numeric(Final_Dataset$ALP1_Preference)
Final_Dataset$ALP1_Preference <- (Final_Dataset$ALP1_Preference -1) / 2*100
Final_Dataset %>%
  tabyl(ALP1_Preference)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset$ALP1_Preference %>% summary()
Final_Dataset$ALP1_Preference %>% sd()
#Level of automation
Final_Dataset$ALP2_Type.of.automat <- as.numeric(Final_Dataset$ALP2_Type.of.automat)
Final_Dataset$ALP2_Type.of.automat <- (Final_Dataset$ALP2_Type.of.automat -1) / 3*100
Final_Dataset %>%
  tabyl(ALP2_Type.of.automat)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
Final_Dataset$ALP2_Type.of.automat %>% summary()
Final_Dataset$ALP2_Type.of.automat %>% sd()
#Averages of all important aspects
mean(Final_Dataset$Intention1)
```

```
sd(Final_Dataset$Intention1)
mean(Final_Dataset$Intention2)
sd(Final_Dataset$Intention2)
mean(Final_Dataset$Risk)
sd(Final_Dataset$Risk)
mean(Condition1$Risk)
sd(Condition1$Risk)
mean(Condition2$Risk)
sd(Condition2$Risk)
mean(Final_Dataset$Av_Trust)
sd(Final_Dataset$Av_Trust)
mean(Final_Dataset$TH2_Rail_quality_1)
sd(Final_Dataset$TH2_Rail_quality_1)
mean(Final_Dataset$TH2_Rail_quality_2)
sd(Final_Dataset$TH2_Rail_quality_2)
mean(Final_Dataset$TH2_Rail_quality_3)
sd(Final_Dataset$TH2_Rail_quality_3)
mean(Final_Dataset$TH2_Rail_quality_4)
sd(Final_Dataset$TH2_Rail_quality_4)
mean(Final_Dataset$TH2_Rail_quality_5)
sd(Final_Dataset$TH2_Rail_quality_5)
mean(Final_Dataset$TH2_Rail_quality_6)
sd(Final_Dataset$TH2_Rail_quality_6)
mean(Final_Dataset$TH3_DriverlessF_R_1)
sd(Final_Dataset$TH3_DriverlessF_R_1)
mean(Final_Dataset$TH3_DriverlessF_R_2)
sd(Final_Dataset$TH3_DriverlessF_R_2)
mean(Final_Dataset$TH3_DriverlessF_R_3)
sd(Final_Dataset$TH3_DriverlessF_R_3)
mean(Final_Dataset$TH3_DriverlessF_R_4)
sd(Final_Dataset$TH3_DriverlessF_R_4)
mean(Final_Dataset$TH3_DriverlessF_R_5)
sd(Final_Dataset$TH3_DriverlessF_R_5)
mean(Final_Dataset$TH4_factors_1)
sd(Final_Dataset$TH4_factors_1)
mean(Final_Dataset$TH4_factors_2)
```

```
sd(Final_Dataset$TH4_factors_2)
mean(Final_Dataset$TH4_factors_3)
sd(Final_Dataset$TH4_factors_3)
mean(Final_Dataset$TH4_factors_4)
sd(Final_Dataset$TH4_factors_4)
mean(Final_Dataset$FC2_attitude_1)
sd(Final_Dataset$FC2_attitude_1)
mean(Final_Dataset$FC1_Intention.to.use_1)
sd(Final_Dataset$FC1_Intention.to.use_1)
mean(Condition1$FC1_Intention.to.use_1)
sd(Condition1$FC1_Intention.to.use_1)
mean(Condition2$FC1_Intention.to.use_1)
sd(Condition2$FC1_Intention.to.use_1)
mean(Final_Dataset$TH5_Trust_7)
sd(Final_Dataset$TH5_Trust_7)
sd(Final_Dataset$ALP2_Type.of.automat)
Final_Dataset %>%
  tabyl(Experience)%>%
  adorn_totals("row")%>%
  adorn_pct_formatting()%>%
  view()
```