

Trust in automated vehicles: a systematic review

Miranda Versteegh

16 June 2019

University of Twente

Faculty of Behavioral Science

Psychology

Supervisors:

Dr. Simone Borsci

Francesco Walker MSc

Contents

Abstract.....	3
Introduction.....	4
Trust.....	5
Calibration of trust.....	7
The aim of this study.....	11
Methods.....	12
Results.....	17
Discussion.....	27
Limitations.....	29
Conclusion.....	30
References.....	31
Appendix A.....	36
Appendix B.....	37
Appendix C.....	40

Abstract

This systematic review investigates literature regarding trust in automated vehicles and follows the PRISMA guidelines. First, the concept of trust is researched, as well as the different levels of automation that exist. Then, the different experiments that were done are being analyzed. The main focus of this analysis is the interest in the subject and the levels of automation, as well as investigating the experiment materials and measuring methods. The analysis on the levels of automation showed that level three and an unspecified level of automation were the most researched levels. Furthermore, the driving simulator was found to be the most used experiment material. Questionnaires were the most used measuring method. However, psychophysiological measurements did seem to be a promising addition. During the analysis, it became clear that some studies lacked some replicability. This was due to missing information on the levels of automation and the type of questionnaire that was used. The conclusion of this study is that it may point in the right direction for finding a reliable measure of trust and that replicability can be improved upon.

Introduction

In the last decade, the development of automated technologies has been increasing (Noah et al., 2017; Khastgir, Birrell, Dhadyalla, & Jennings, 2018). Automation is defined as a ‘technology that actively selects data, transforms information, makes decisions, or controls processes’ (Lee & See, 2004; Hoff & Bashir, 2015). Within this development of increasing automation in technology, automated vehicles have gotten a lot of attention (Khastgir et al., 2018). An automated vehicle can be defined as a ‘robotic vehicle that works without a human operator’ (Kaur & Rampersad, 2018). The technology that is being used to make this possible is called an advanced driver assistance system (ADAS), or otherwise called an automated driving system (ADS) (Walker, Boelhouwer, Alkim, Verwey, & Martens, 2018; Kelechava, 2018). This refers to the technology consisting of software and hardware that has the supporting role for the driver during a driving task.

The goal is to ultimately achieve fully automated driving (Payre, Cestac, & Delhomme, 2016). The reason for this is that automated systems can potentially lead to an increase in safety (Khastgir, Birrell, Dhadyalla, & Jennings, 2017; Payre, Cestac, & Delhomme, 2016; Khastgir et al., 2018; Molnar et al., 2018; Choi & Ji, 2015). Human errors on the road are the leading cause of traffic accidents (Khastgir et al., 2017; Hergeth, Lorenz, Vilimek, & Krems, 2016; Choi & Ji, 2015). The numbers vary between studies, but all indicate a percentage above 90% for the amount of accidents that are caused by a human. Other benefits of automated vehicles are to improve the comfort of the driver, decrease spent fuel, decreasing the driver’s workload, and improve mobility of elderly or disabled people (Payre et al., 2016; Molnar et al., 2018; Hergeth et al., 2016). Furthermore, automated systems can perform better and more efficient than humans in certain situations (Boubin, Rusnock, & Bindewald, 2017; Choi & Ji, 2015). This is due to the system’s capability to process a large amount of information very quickly.

These beneficial factors that automated vehicles bring, can differ with each level of automation. These levels go from no automation, to full automation, in which the vehicle completely takes over the driving tasks (Kaur & Rampersad, 2018). Different definitions of these levels have been developed. The Society of Automotive Engineers (SAE) developed a classification of 5 levels of automation (SAE, 2014). The first level is classified as no driving automation. This means that the driver performs all the driving tasks. Level 1 includes driver assistance. So the vehicle performs smaller tasks, while the driver is expected to still execute the majority of the driving tasks. At level 2, the driver is supervising the system, while the vehicle can

perform tasks simultaneously. Then, level 3 is the turning point at which the driver is more noticeably less engaged in performing the driving tasks. This is called conditional driving automation and refers to the vehicle performing an entire, specifically requested driving task. The system can request the driver to intervene if necessary. Level 4 differs from level 3 in that the driver is no longer asked to intervene in the driving task. Finally, the highest level of automation is level 5. At this level, the vehicle is no longer restricted to only performing a requested driving task, but can perform completely on its own.

So with a higher level of automation, human error decreases (Payre et al., 2016). However, despite these benefits of higher levels of automation, users do not trust automated vehicles yet (Khashtrgir et al., 2017; Dixon, Hart, Clarke, O'Donnell, & Hmielowski, 2018). The benefits of automated vehicles can only unfold if these technologies are adopted by drivers and society as a whole (Gold, Körber, Hohenberger, Lechner, & Bengler, 2015). The automated system needs to be designed properly as well. Designers need to be aware that these automated systems cannot replace a human operator completely (Boubin et al., 2017). Both have skills that the other has not. For example, humans are better at making judgements, and automated systems have the benefit regarding speed and performance. Therefore, the relationship and cooperation between humans and automation needs to be optimized.

One of the most important and influential factors that determines this relationship is trust (Khashtrgir et al., 2017; Walker, et al., 2018; Noah et al., 2017; Schaefer, Chen, Szalma, & Hancock, 2016; Choi & Ji, 2015; Lazanyi & Maraczi, 2017; Dixon et al., 2018; Molnar et al., 2018). Trust was found to be a determining factor in the adoption of automated vehicles (Kaur & Rampersad, 2018). This will be discussed more in depth to get a better understanding of what the concept of trust entails.

Trust

It is important to distinguish between interpersonal trust, so trust between humans, and trust in automation (Lee & See, 2004; Hoff & Bashir, 2015). The two kinds of trust may seem similar on the surface, but should not be confused with each other. The similarities are that both concern a situation in which a cooperative relationship exists, an exchange takes place, and there is a certain level of uncertainty (Hoff & Bashir, 2015). Despite these similarities, the differences that exist between the two kinds of trust make them quite distinct. Interpersonal trust is based on perceived

ability, integrity, and benevolence. While trust in automation is based on performance, process, and purpose. Moreover, the development of both kinds of trust is also different (Lee & See, 2004; Hoff & Bashir, 2015). The initial basis of interpersonal trust, or the level of trust measured at the beginning of the interaction, is based on the predictability of the trustee (Hoff & Bashir, 2015). When the relationship progresses, the dependability and integrity of the trustee becomes more important. Finally, in a fully developed relationship, faith or benevolence become important. The development of trust in automation is quite the opposite. Initial trust is based on faith. When the system shows errors, dependability and predictability will become the basis of trust (Hoff & Bashir, 2015). From now on, the term trust will be used to indicate trust in automation and not interpersonal trust. The basis of trust will be explored more in-depth later in this paper.

Additionally, trust in automation needs to be properly defined. The definition of trust in automation was developed over the last years (Lee & See, 2004). Some definitions are based on expectancy; others use intention or willingness to act (Lazanyi & Maraczi, 2017; Payre et al., 2016; Noah et al., 2017; Choi & Ji, 2015; Lee & See, 2004; Hoff & Bashir, 2015; Kaur & Rampersad, 2018). The most widely used definition is based on vulnerability or risk, uncertainty, goal-oriented tasks, and the dynamic nature of trust. This definition is as follows: ‘the attitude that an agent will help achieve an individual’s goal in a situation characterized by uncertainty and vulnerability’ (Lee & See, 2004, p. 51). Yet other definitions are phrased differently, but all of them are at least based on the risk factor and the uncertainty (Lazanyi & Maraczi, 2017; Hoff & Bashir, 2015; Kaur & Rampersad, 2018). In this paper, the definition made by Lee and See (2004) will be used from this point forward.

To dive in even deeper into the concept of trust in automation, there are different types of trust. The first is the difference in three layers of variability (Hergeth et al., 2016; Noah et al., 2017; Hoff & Bashir, 2015). The first layer is called dispositional trust (Hergeth et al., 2016; Noah et al., 2017). This kind of trust represents the operator’s tendency to trust the system. This trust exists before interaction with the system and is influenced by demographic factors, as well as personality traits. The second layer is situational trust. This is dependent on the external environment, as well as the specific reactions in each situation of the operator. These characteristics can also dependent on the context. This trust is built through interaction with the system. The third and final layer is learned trust. This layer is based on the knowledge formed by past experiences and interactions. It draws from perceived system performance and reliability.

Furthermore, this last layer consists of two categories: initial learned trust and dynamic learned trust (Hoff & Bashir, 2015). Initial trust is the trust that exists before system interaction and can be based on the reputation that a system has built and on the past experiences in a similar situation. Dynamic trust is trust during system interaction. This is based on the performance of a system, and can therefore change during an interaction with said system. These layers and their differences are sometimes described differently, but their essence remains the same (Lazanyi & Maraczi, 2017).

Furthermore, a difference was indicated in dimensions of trust based on beliefs (Choi & Ji, 2015). These are system transparency, technical competence, and situation management. System transparency is trust based on the belief that a system is predictable and understandable. Technical competence is based on the perception of the system's performance. Finally, situation management is based on the belief that the operator can recover control whenever required.

Finally, a distinction can be made between trust *in* automation, and trust *with* automation (Khastgir et al., 2017; Khastgir et al., 2018). Trust *in* automation, or the system, is the trust that the system is functioning like it is supposed to. The driver would be guided by the perceived capabilities of the system, whether those are accurate or not. Trust *with* the system means that the driver has accurate knowledge about the true capabilities and limitations of the system. This knowledge is then used to get the most benefits out of the system.

Calibration of trust

Now that a background on the concept of trust is investigated, it is important to look into what is called the 'calibration of trust'. While maximizing trust might seem a convincing course of action, calibrating trust is actually more important (Khastgir et al., 2017; Walker, Boelhouwer, et al., 2018; Dikmen & Burns, 2017). This is defined as 'the process of adjusting trust to correspond to an objective measure of trustworthiness' (Khastgir et al., 2017, p. 542), or the 'match between abilities of the automation and the person's trust in automation' (Payre et al., 2016, p. 230). Although there are more definitions, all of them have the same core of matching the user's trust with the actual capabilities of the system (e.g. Walker, Boelhouwer, et al., 2018; Hergeth et al., 2016; Khastgir et al., 2018; Lee & See, 2004; Hoff & Bashir, 2015). The calibration of trust is important due to the risk of over-trust and distrust. In the case of over-trust, the operator has too much trust in a system. This can cause over-reliance in the system and using the system beyond its capabilities (Boubin et al., 2017; Noah et al., 2017). In the case of distrust, the operator has too

little trust in the system, causing under-reliance. This will lead to the operator using the system less, if at all. If trust is calibrated, these problems can be avoided. Knowledge about both the limitations and capabilities of a system can help reach the appropriate level of trust (Noah et al., 2017).

To determine more ways to calibrate trust, it is of great importance to determine the factors that influence trust. Firstly, some researchers indicate the factors performance, process, and purpose as influencers of trust (Dikmen & Burns, 2017; Noah et al., 2017; Choi & Ji, 2015; Lee & See, 2004). Performance refers to the operator's observation of the result of the actions of the system (Dikmen & Burns, 2017). Process is the observation of the functioning of the system, followed by an understanding of how it makes decisions (Dikmen & Burns, 2017; Noah et al., 2017). Purpose relates to the understanding of the intention of the system (Dikmen & Burns, 2017; Noah et al., 2017). Each of these factors and how they are perceived by the operator should align with the objective, real-world situation to achieve the appropriate amount of trust. To realize this, information on all three dimensions should be provided (Lee & See, 2004).

Besides these three dimensions, many more factors are mentioned by different researchers. Examples of these are automation error, experience, transparency, certification, situation awareness, workload, consequence, willingness and self-confidence (Khastgir et al., 2017; Dikmen & Burns, 2017; Khastgir et al., 2018; Walker, Martens, & Verwey, 2018; Hoff & Bashir, 2015). Mentioning all the factors that affect trust in automation is beyond the scope of this paper. Still, there is an overall agreement in the existing literature that accurate knowledge about the system is a very important factor. The right knowledge could lead to calibrated trust, which in turn can lead to the appropriate use of the system (Khastgir et al., 2017). There are three kinds of knowledge: static knowledge, real time knowledge, and internal mental model (Khastgir et al., 2017; Khastgir et al., 2018). Static knowledge is the understanding of how the system works. This knowledge exists before an interaction with an automated vehicle and can be built up over time as the driver gains experience. Real time knowledge refers to the state of the system and the environment. This kind of knowledge is dynamic and requires the driver to stay in the loop (Khastgir et al., 2018). To stay in the loop refers to the driver being informed of the state and performance of the system in real-time. In other words, the system is transparent, which could also be explained as an increase in awareness and knowledge. That transparency can be reached by providing information about the system to the user, either beforehand or in real-time.

Understanding the influence that external sources have is what the internal mental model draws on (Khastgir et al., 2017). These sources can be the media or marketing campaigns that affect the driver's trust and perception (Khastgir et al., 2018).

Awareness of the above-mentioned factors, and possibly even more, is necessary to determine how trust can be measured. In the past years, different kinds of measurements were used. The majority of the studies measured trust in a driving simulator (Payre et al., 2016; Hergeth et al., 2016; Khastgir et al., 2018; Molnar et al., 2018; Gold et al., 2015; Walker, Martens, & Verwey, 2018; Hergeth, Lorenz, & Krems, 2017). Most used an interactive interface, whilst others used non-interactive video material (Walker, Martens, & Verwey, 2018). Another method was to use Manual Control Recovery (MCR) (Payre et al., 2016). The idea behind this was that the more trust there is, the less a driver will monitor the system. Therefore, if the system indicates that the driver needs to regain MCR, the reaction time will be dependent on the amount of trust. If the driver has a high level of trust, and therefore is not monitoring the system very often, the reaction time is expected to increase due to the driver not being prepared.

To gather data on the participants' trust, most studies used self-report measures, such as questionnaires and open questions (e.g. Lazányi, 2018; Payre et al., 2016; Dikmen & Burns, 2017; Kircher, Larsson, & Hultgren, 2014; Choi & Ji, 2015; Hergeth et al., 2017; Weinstock, Oron-Gilad, & Parmet, 2012; Filip, Meng, Burnett, & Harvey, 2016). Furthermore, questionnaires were often used right before the driving task and right after. Some questionnaires were even distributed without an actual driving task, but relied on people's prior experiences (Lazanyai & Maraczi, 2017; Dikmen & Burns, 2017; Dixon et al., 2018). Despite being a commonly used measurement, questionnaires do not measure continuously (Walker, Martens, & Verwey, 2018). Therefore, it cannot capture any real-time changes in trust, whilst trust is a dynamic construct. That is why some researchers have turned to another measure: eye-tracking (Hergeth et al., 2016; Kircher, Larsson, & Hultgren, 2014; Gold et al., 2015; Walker, Martens, & Verwey, 2018). Gaze behavior is said to be an indicator for attention and situation awareness (Kircher, Larsson, & Hultgren, 2014; Gold et al., 2015). Attention and situation awareness are used to indicate the frequency and duration of monitoring behavior of the driver. It is theorized that if a driver monitors the system and the road less, trust is higher than when a driver monitors these more often (Gold et al., 2015; Walker, Martens, & Verwey, 2018). That is why monitoring behavior could potentially be an effective measure of trust (Walker, Martens, & Verwey, 2018).

Trust can also be objectively measured through psychophysiological measures (Akash, Hu, Jain, & Reid, 2018; Wang, Hussein, Rojas, Shafi, & Abbass, 2018; Hirshfield et al., 2014; Filip et al., 2016; Bui, Verhoeven, Lukkien, & Kocielnik, 2013; Vecchiato et al., 2014; Khawaji, Zhou, Chen, & Marcus, 2015). The two most commonly used psychophysiological measures are galvanic skin response (GSR) and electroencephalography (EEG). These measures are non-invasive and it allows the measurement of participant's states in real-time (Akash et al., 2018; Hirshfield et al., 2014). GSR, or electrodermal activity, indicates the amount of arousal a person feels by measuring the conductivity of the skin. The level of arousal was already used to indicate other states of mind, such as stress and anxiety. Now, it is also used to measure trust. The second measurement, EEG, is an instrument that measures brain activity, more specifically the cortical activity (Akash et al., 2018). This activity is analyzed by measuring the electromagnetic field of the brain by collecting signals from electrodes (Artinis, 2018). These signals can indicate changes in thoughts, emotions, and actions. Four brain regions were discovered as being responsible of trust: the left frontal region, the right frontal region, the fronto-central region, and the occipital area (Wang et al., 2018). A part of EEG measurements are event-related potentials (ERPs) (Akash et al., 2018). ERPs are used to measure brain activity that occurs as a response to a certain event. This has been seen as an impractical form of using EEG as a measurement, due to the difficulty of pointing out the specific triggers for the measured brain activity. Despite this, EEG measurements could help in the understanding of trust (Wang et al., 2018). More recently, fMRI has been used in addition to GSR and EEG measures (Hirshfield et al., 2014). However, it is not deemed very useful in a setting in which the participant has to interact with the system. This is because it is necessary for the participant to lie still while undergoing fMRI scans. Therefore, a new tool called functional near infrared spectroscopy (fNIRS) has been developed. This is a wearable headset that can measure brain activity in real time. Unlike an EEG measure, fNIRS measures the changes that occur in the level of oxygen in the blood in a specific region of the brain when it becomes active (Artinis, 2018). This is different from EEG, as EEG measures the electromagnetic field in the brain that occurs with firing neurons. Finally, a lesser used psychophysiological measurement is the heart rate (HR) measure, or electrocardiography (ECG) (Bui, Verhoeven, Lukkien, & Kocielnik, 2013; Vecchiato et al., 2014). Measuring heart rate, or heart rate variability (HRV), can indicate the presence of activity in the parasympathetic nervous system. Especially this activity seems to be of importance when one person is judging another person on trustworthiness (Vecchiato et al., 2014).

The aim of this study

This study follows the PRISMA guidelines, which refers to a systematic review of the literature relevant to the subject, in this case trust in automated vehicles (Liberati et al., 2009). The PRISMA statement itself consists of a checklist and a flow diagram that is intended to be a guideline for reviews such as this one. The goal is to optimize the quality of systematic review reports. Therefore, the objective of this study is to review the literature on trust in automated vehicles in an appropriate manner and to identify all the important factors and processes that contribute to said trust.

This study is a systematic review, because the topic of trust in automated vehicles is gaining interest. However, it is important to establish a baseline of the knowledge that exists so far. This is because trust in automation is a specific kind of trust of which it is not known if and how similar it is to interpersonal trust. Moreover, researchers are still testing for a dependable, objective way of measuring trust in automation. This study may provide a direction in which future research can move in order to find a more dependable measure for trust.

In the remainder of this study, three questions will be addressed. Firstly, the interest in the subject, as well as the different levels of automation will be researched. Furthermore, the various experiment materials and measuring methods will be investigated. Finally, this study will attempt to determine the most successful experiment and measuring methods.

Methods

This systematic review was conducted by using articles from a variety of sources. Scopus and EBSCO PsycINFO were the databases used to gather articles. Furthermore, Google Scholar was used for additional information. Finally, a Google Drive folder was used for the exchanging of articles that were found by others.

The keywords that were used across all databases were ‘trust’, and ‘automated vehicles’ (Table 1). These were selected, because these were thought of as the core keywords of the subject. Other keywords were added whenever a specific part of the subject needed more clarification. The term ‘self-driving car’ was not included after the initial search in Scopus. The reason for this was that by only looking into cars, and not vehicles, the search would yield too many articles that were too specific. Additionally, Google Scholar was used to find background information on the definition of the levels of automation. Articles found in the databases did not discuss all the levels of automation. Most only explained the specific level that was tested. Only the first page of the search results was used, as just a definition of the levels of automation was required. Finally, Google Scholar was also used to look for a handful of additional articles that were not found in the database search. This search was also limited to the first page of the search results. The reason for this being that there were simply too many results to review all of them.

The limitation to the search field that were used during the literature search, was the limitation of publishing years, namely only the years 2000 to 2019 were included. in the case of Google Scholar and Web of Science an additional limitation was used, namely the limitation of using the first page only.

Table 1

Keywords used for each database

Database	Keywords	Limitations
Scopus	Trust AND 'automated vehicles' OR self-driving car	Years 2000-2019
	Trust OR 'trust calibration' AND 'automated vehicles'	Years 2000-2019
	"Electrodermal activity" OR "galvanic skin response" AND trust	Years 2000-2019

EBSCO PsycINFO	Trust AND ‘automated vehicles’	Years 2000-2019
Google Scholar	Levels of automation	Years 2000-2019, first page only
	Trust in automated vehicles	Years 2000-2019, first page only

At the beginning of the collecting of articles, a number of criteria were set up. Using these criteria, articles were either included or excluded from further examination. However, it was deemed necessary to change the criteria towards the end of the article collection process. The reason for this choice was the possibility that too many valuable articles were overlooked, resulting in too many excluded articles. By altering the criteria, more previously discarded articles were included after all. Those articles could give a more complete view of the subject. These changes in the criteria can be seen in Table 2.

Table 2

Criteria for excluding or including articles

Initial criteria	Criteria after reconsideration
1. The article should be longer than 2 pages	1. The article should be longer than 2 pages
2. The article should be on the topic of both trust and automated vehicles, or background information on only trust or only automated vehicles.	2. The article should be on the topic of both trust and automated vehicles, or background information on only trust or only automated vehicles.
3. The article should discuss the influence of trust on automated vehicles, and not the influence of automated vehicles on trust.	3. <i>Alteration:</i> The article can discuss both directions of influence on trust or on automated vehicles.
4. The article should not be about one specific function of an automated	4. <i>Alteration:</i> The article should not be about one specific function of an

vehicle or on one specific factor of trust. The subject of automated vehicles or trust should also not only be a small part of the article, but has to be the main focus

automated vehicle or on one specific factor of trust. The subject of automated vehicles or trust should also not only be a small part of the article, but has to be the main focus. The study can discuss other domains to investigate the different kinds of measurements that are being used.

5. The article can discuss background information necessary for understanding the whole picture. For example, the subjects trust calibration, levels of automation, and different kinds of measurements can be included.

5. The article can discuss background information necessary for understanding the whole picture. For example, the subjects trust calibration, levels of automation, and different kinds of measurements can be included.

6. The article should only be about trust in automated vehicles, or trust in automation in general. It should not be about interpersonal trust as this is too different from trust in automated vehicles.

6. *Alteration*: The article should only be about trust in automated vehicles, or trust in automation in general. It should not be about interpersonal trust as this is too different from trust in automated vehicles. The exception on this rule is an article about interpersonal trust that used psychophysiological measures.

The first criterion was chosen to ensure that the article would provide enough in-depth information. The second criterion was chosen to ensure enough specificity in the articles, as well as to provide a complete basis in the background information. The third criterion was first chosen to ensure even more specificity in the articles. However, after reconsideration, it was concluded that this third criterion was in fact too specific. The direction of influence is important to keep in mind. Still, the

direction of influence can change at any time. So to exclude one direction would mean that an important part of the subject is missing, which might provide an incomplete view. Therefore, this criterion was altered to express that the direction of influence was not a reason for exclusion any longer. The next and fourth criterion was initially chosen to ensure the optimal amount of specificity and in-depth information. This criterion was changed, because it did not yet take into account that broadening the search field could be beneficiary. Therefore, the fourth criterion now also includes the notion that other domains are allowed to be used in order to compare the measurements used. The possibility exists that measurements that were proven to work in other domains can be used to measure trust in automated vehicles as well. The fifth criterion includes a similar notion as the second criterion, namely the background information. The difference between the two is that the second criterion includes background information on the broader subjects of trust and automated vehicles. The fifth criterion specifies some more specific kinds of background information that were needed for a more in-depth view of the general topics. Finally, the sixth criterion was chosen at first due to the important difference between interpersonal trust and trust in automation. It was speculated that an article on interpersonal trust would not be useful. However, it could also be said that interpersonal trust is the factor that has the closest relation to trust in automation. Therefore, a study that tested interpersonal trust could in fact say something about the possibilities for measuring trust in automation. Especially psychophysiological measurements are the most promising at the moment. If psychophysiological measurements were deemed reliable to measure interpersonal trust, it might just be a reliable start to measure trust in automation as well.

These changes were made quite late in the process of literature collection, as can be seen in Figure 1. At the beginning of the process, both the databases and the Google Drive folder were used to identify promising articles. At first, articles that were found in more than one source were removed. Next, articles were screened based on the title, the keywords, and the abstract. The initial criteria were already chosen at this time. An article was excluded when it was not on topic, or when any of the criteria were not met. After this, the articles were read in full to determine if the criteria were met. If it did not, it was excluded and marked with the specific criteria it did not meet. After having collected articles thus far, the criteria were revised and altered where necessary. Then, the already discarded articles were reviewed again to see if the criteria were met this time. When that was the case, the article was included in the systematic review after all.

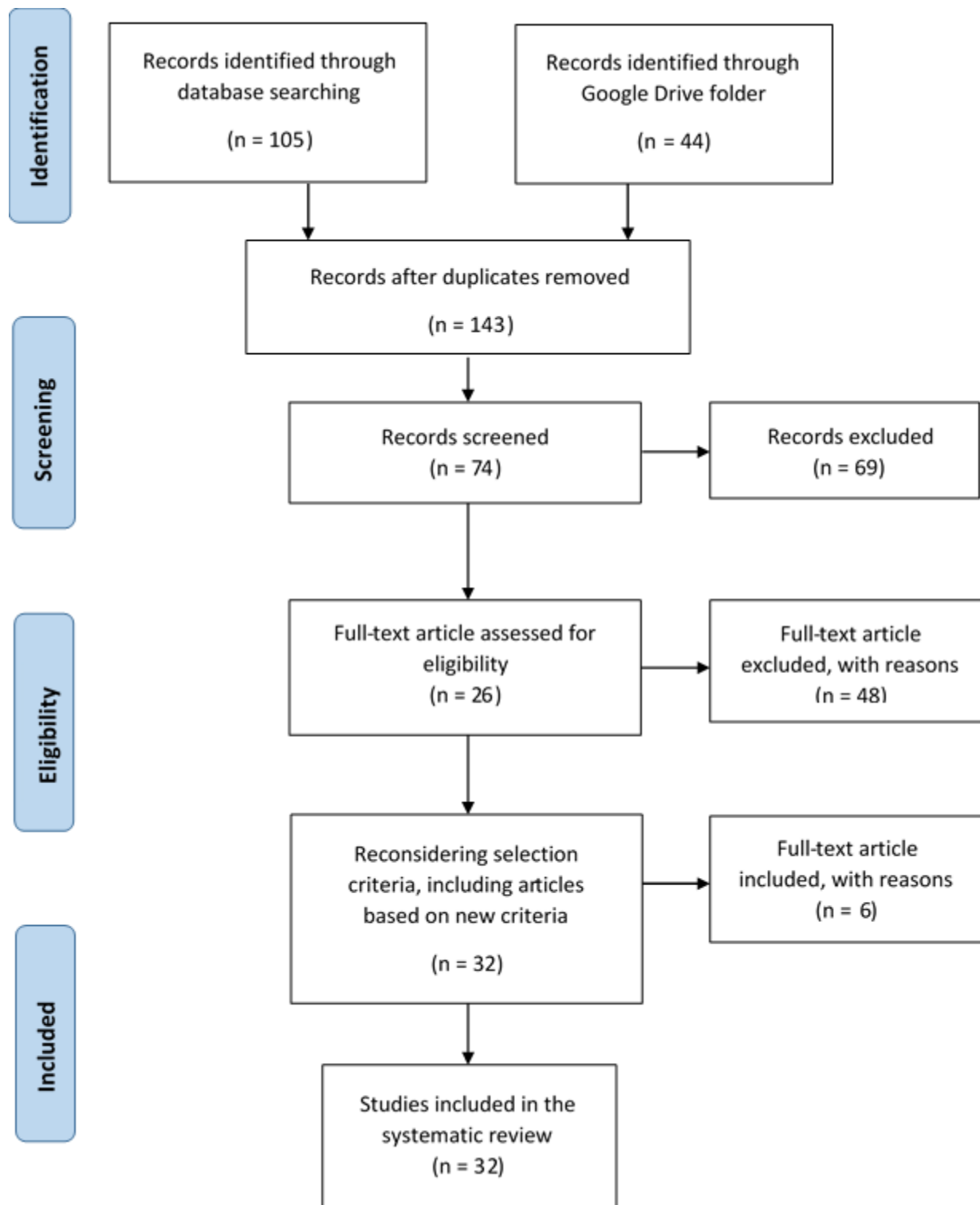


Figure 1. PRISMA flowchart: the process of literature collection

Results

In the following analysis, only 29 out of the total of 32 studies will be used. This is because three studies were only useful as background information in the introduction and did not provide useful data for the analysis (Artinis, 2018; Kelechava, 2018; SAE, 2014). In Table 3, all the studies used for the further analysis can be seen, as well as a short description of the subject(s) that were investigated.

Table 3

Articles included in the analysis

N.	Article	Year	Subject
1	A classification model for sensing human trust in machines using EEG and GSR	2018	Measuring trust using EEG and GSR
2	A meta-analysis of factors influencing the development of trust in automation: implications for understanding autonomy in future systems	2016	Analysis of factors that influence trust in automation
3	A trust evaluation framework for sensor readings in body area sensor networks	2013	Evaluating trustworthiness of sensors
4	Are we ready for self-driving cars – a case of principal-agent theory	2018	Trust between principal and agent
5	Calibrating trust through knowledge: Introducing the concept of informed safety for automation in vehicles	2018	The effect of knowledge on trust
6	Calibrating trust to increase the use of automated systems in a vehicle	2017	Trust as influence on usage. Investigation into the factors that influence trust and calibration of trust
7	Changes in trust after driving level 2 automated cars	2018	Trust calibration and measurement of trust when using a level 2 automated car

8	Designing and calibrating trust through situational awareness of the vehicles (SAV) feedback	2016	Situational awareness
9	Dispositional trust – Do we trust autonomous cars?	2017	Dispositional trust
10	EEG-based neural correlates of trust in human-autonomy interaction	2018	Measuring trust through decision-making
11	First workshop on trust in the age of automated driving	2017	Calibration of trust
12	Fully automated driving: Impact of trust and practice on manual control recovery	2016	Fully automated driving and manual control recovery
13	Gaze behavior as a measure of trust in automated vehicles	2018	Research for effective trust measurements testing gaze behavior as a measure
14	Investigating the importance of trust on adopting an autonomous vehicle	2015	Testing a trust model
15	Keep your scanners peeled: Gaze behavior as a measure of automation trust during highly automated driving	2016	Gaze behavior as measurement of trust
16	Neuroelectrical correlates of trustworthiness and dominance judgments related to the observation of political candidates	2014	Measuring trust with EEG, GSR, and HR
17	Prior familiarization with takeover requests affects drivers' takeover performance and automation trust	2016	An analysis into the effect of prior familiarization on performance when a take-over request is made
18	Quantifying compliance and reliance trust behaviors to influence trust in human-automation teams	2017	Compliance and reliance
19	Tactical driving behavior with different levels of automation	2014	Tactical driving behavior, attention

20	The effect of system aesthetics on trust, cooperation, satisfaction and annoyance in an imperfect automated system	2012	Effect of aesthetics on trust, cooperation, satisfaction and annoyance
21	Trust in automation –Before and after the experience of take-over scenarios in a highly automated vehicle	2015	Takeover scenarios, situation awareness and monitoring
22	Trust in automation: Designing for appropriate reliance	2004	What to keep in mind when creating automation that can be trusted
23	Trust in automation: integrating empirical evidence on factors that influence trust	2015	An analysis into the different layers of trust
24	Trust in autonomous vehicles: The case of tesla autopilot and summon	2017	Trust and confidence in Tesla vehicles
25	Trust in driverless cars: Investigating key factors influencing the adoption of driverless cars	2018	Testing driverless cars in a closed environment
26	Understanding trust and acceptance of automated vehicles: An exploratory simulator study of transfer of control between automated and manual driving	2018	Trust and acceptance in relation to transfer of control
27	Using galvanic skin response (GSR) to measure trust and cognitive load in the text-chat environment	2015	Measuring trust and cognitive load with GSR
28	Using noninvasive brain measurement to explore the psychological effects of computer malfunctions on users during human-computer interactions	2014	Measuring the real-time state of a person with fNIRS
29	What drives support for self-driving car technology in the United States?	2018	Predictors of support of automated vehicles

To get a more tangible idea of the existing literature on automated vehicles, various factors were analyzed. Firstly, the amount of articles per year was investigated (Figure 2).

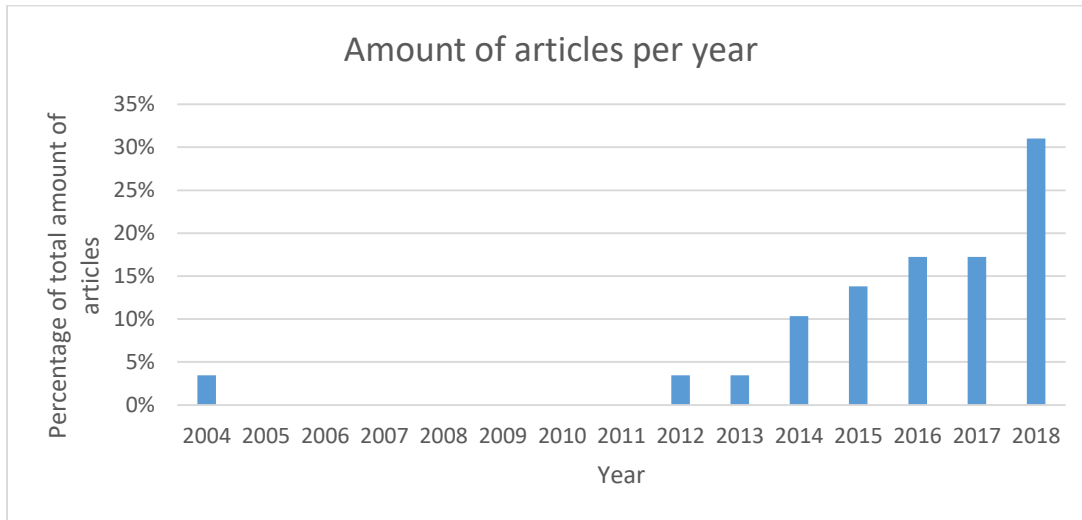


Figure 2. Amount of articles per year

In Figure 2, only the articles that have the specific focus on automated vehicles and/or trust are included. The reason for this being that it is important to look at the development of focus on this specific field over the years. As shown in the graph, the number of articles published increases in the past years. No articles that met the criteria were found in the years between 2004 and 2014. To see which specific studies were published per year, see Table 3. Furthermore, the levels of automation that were investigated were compared (Figure 3).

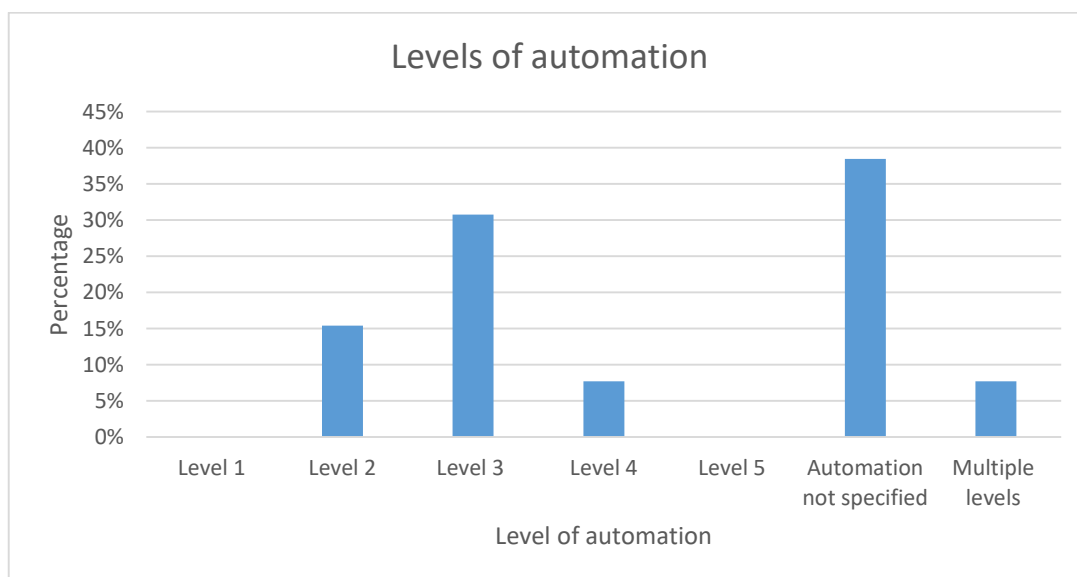


Figure 3. Amount of experiments per level of automation

Not all articles are clear on which specific levels of automation were used in the experiments (Lazanyi & Maraczi, 2017; Hergeth et al., 2016; Boubin et al., 2017; Kircher, Larsson, & Hultgren, 2014; Walker, Martens, & Verwey, 2018). Therefore, when the level of automation is not specifically mentioned, the study is placed in the category ‘automation not specified’. Furthermore, not all studies investigated levels of automation. That is why only 13 studies are included in this analysis.

As is shown in Figure 3, level 3 is the most researched level of automation (Payre et al., 2016; Hergeth et al., 2017; Gold et al. 2015; Molnar et al., 2018). The unspecified level of automation and the level 2 of automation are the second most researched levels (Walker, Boelhouwer, et al., 2018; Lazanyi & Maraczi, 2017; Walker, Martens, & Verwey, 2018; Hergeth et al., 2016; Boubin et al., 2017; Dikmen & Burns, 2017). The least researched levels are level 4 and a combination of multiple levels of automation (Lazányi, 2018; Khastgir et al., 2018). No studies were found that investigated the lowest and highest level of automation. For an overview of which studies researched which level of automation, see Appendix A.

As for the different methods used to measure the amount of trust, a distinction was made between the experiment materials and the actual measurement itself. Firstly, the experiment materials will be discussed (Figure 4). Not all studies conducted an experiment, which is why this analysis includes 24 studies only.

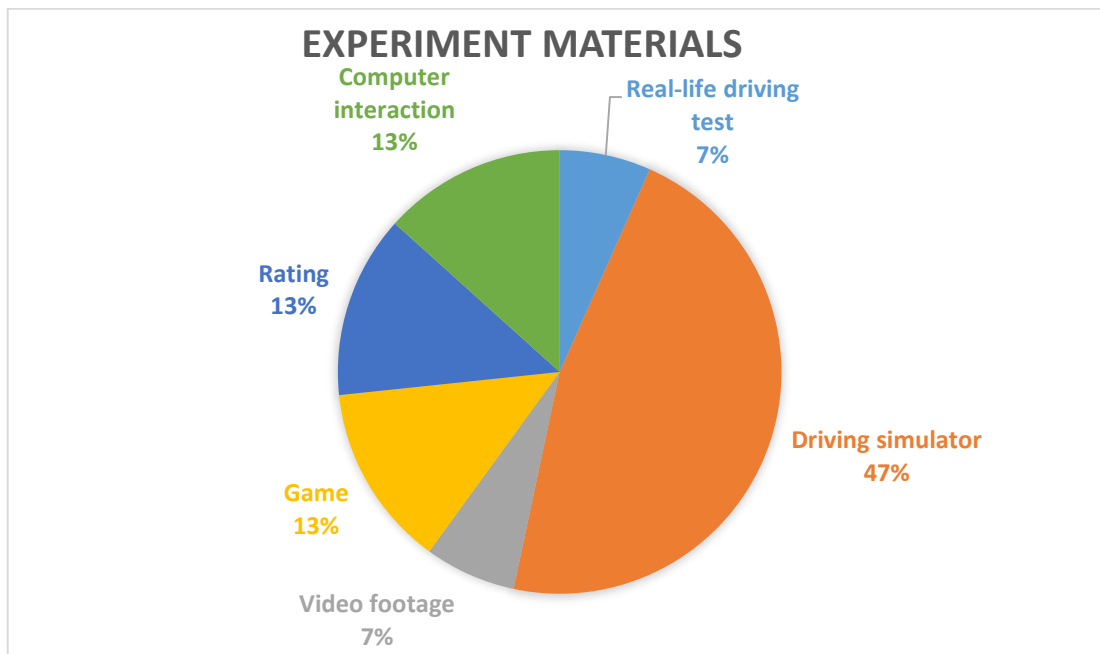


Figure 4. Experiment methods

In the pie chart in Figure 4, the different testing methods are visualized. These include all articles, both on the specific topic of trust and/or automation, and off-topic articles including articles about interpersonal trust and articles that are not specifically automation related. The reason for this is that the articles not specifically focused on automated vehicles did test some potentially useful methods for measuring trust in general. The most widely used method is the driving simulator (Khastgir et al., 2018; Payre et al., 2016; Hergeth et al., 2016; Hergeth et al., 2017; Kircher, Larsson, & Hultgren, 2014; Gold et al., 2015; Molnar et al., 2018). Almost half of all the articles investigated trust by using this method. Games were not used as often, but a couple of studies did try this method (Wang et al., 2018; Boubin et al., 2017). The games that were used were an air traffic control game and an investment game. The category ‘rating’ includes rating the aesthetics of maps and rating the faces of politicians on the perceived trustworthiness (Vecchiato et al., 2014; Weinstock, Oron-Gilad, & Parmet, 2012). This category is, along with the use of games and computer interaction (Akash et al., 2018; Hirshfield et al., 2014) the method that is used most frequently after the driving simulator. Lesser researched methods are the use of video footage and a real-life driving test (Walker, Boelhouwer, et al., 2018; Walker, Martens, & Verwey, 2018) Next, the methods to measure trust should also be taken into account (Figure 5). Again only 24 studies were included in this analysis, as other studies did not conduct an experiment.

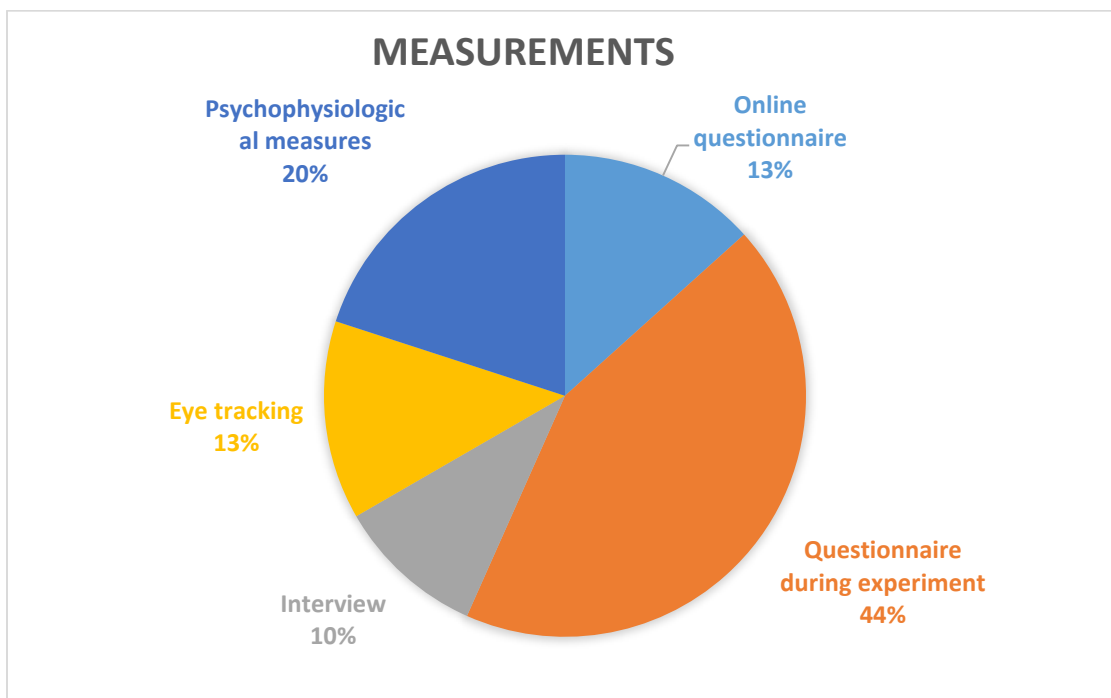


Figure 5. Measurements used to indicate trust

Questionnaires are the most commonly used measurement. Sometimes only one questionnaire was used, other times multiple. Sometimes, multiple measurements are used in combination. Often, a questionnaire in combination with an interview, eye tracking, or a psychophysiological measure is used. A more in-depth investigation into these different types of questionnaires can be seen in Appendix B. The different combinations of measurements are shown in Appendix C. Aside from questionnaires, psychophysiological measures (Akash et al., 2018; Bui et al., 2013; Wang et al., 2018; Vecchiato et al., 2014; Khawaji, Zhou, Chen, & Marcus, 2015; Hirshfield et al., 2014) as well as eye tracking are the second most used measurements (Walker, Martens, & Verwey, 2018; Hergeth et al., 2016; Kircher, Larsson, & Hultgren, 2014; Gold et al., 2015). Interviews are not used as often (Filip et al., 2016; Molnar et al., 2018). To establish an even more comprehensive view, the usage of these measurements over time was investigated (Figure 6).

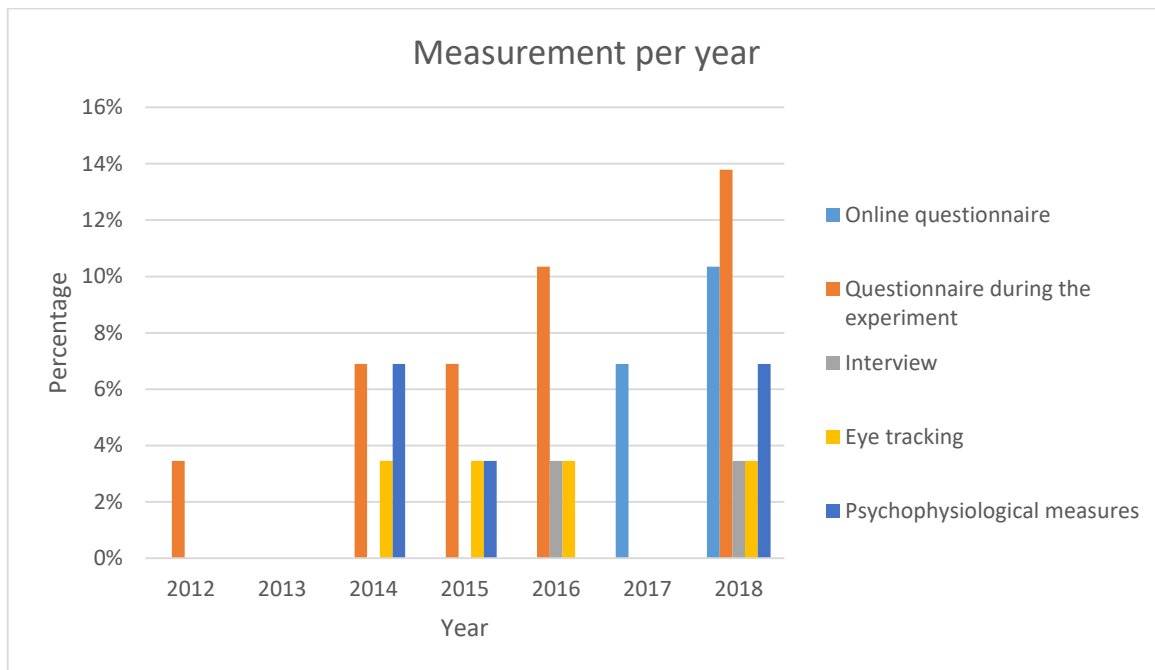


Figure 6. Measurements used per year

As shown in Figure 6, the amount of usage of questionnaires increases over time (See Appendix C). The same increase applies for other measurements as well. Eye tracking has become a relatively stable measurement ever since the year 2014. Psychophysiological measurements are being included as well. For example, EEG, GSR and heart rate are used to measure trust. Online

questionnaires are used more often in recent years as well. Finally, interviews are also explored as indicators for trust. For a more in-depth overview of the measurements, see Appendix C. The results of experiments, when indicated clearly, were investigated as well (Figure 7). This includes the 24 studies investigated in Figure 4, 5, and 6.

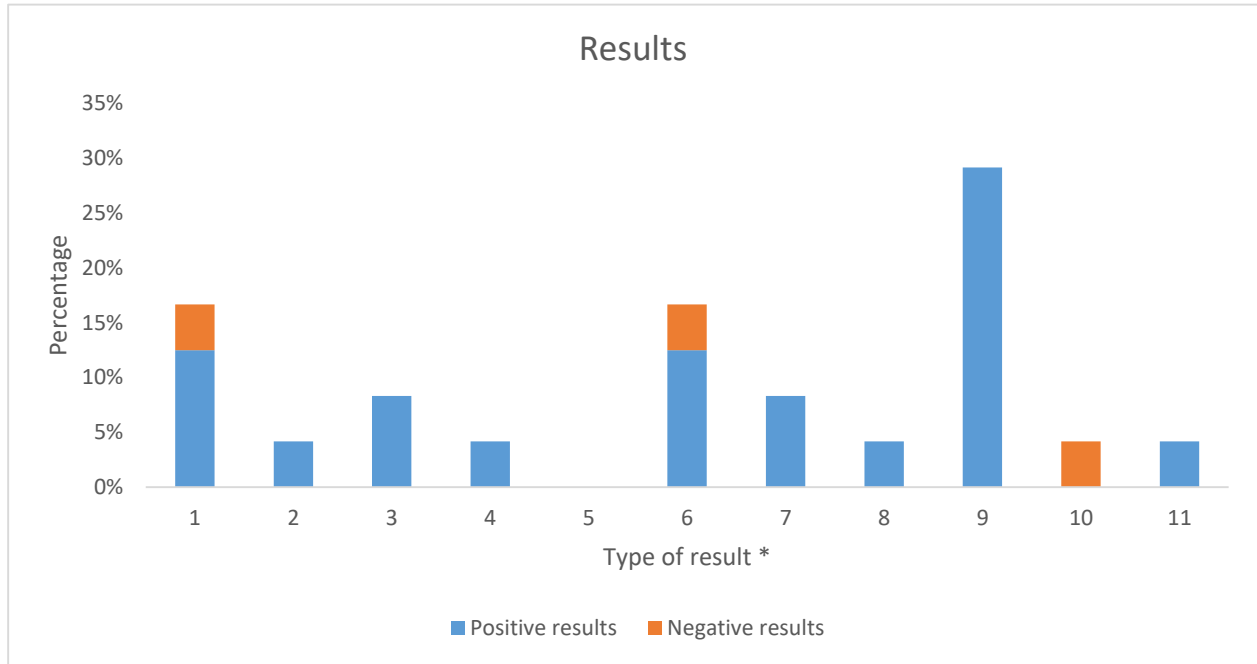


Figure 7. Results

**Note: The x-axis is defined as follows: 1.) influence of age/gender; 2.) reaction time as indicator; 3.) monitoring and control as indicator; 4.) increased trust over time; 5.) decreased trust over time; 6.) gaze behavior tested as measurement; 7.) influence of experience and familiarization; 8.) influence of knowledge; 9.) psychophysiological measures tested as measurement; 10.) influence of system aesthetics, 11.) influence of situational awareness of the vehicle (SAV).*

In this study, a positive result is defined as a significant result. For example, a positive result would indicate that the influence of one factor on another was found to be significant. This does not take into account what kind of influence was measured, only that there was one. When a negative result occurs, it is defined as a non-significant result.

Figure 7 shows all the different results, both positive and negative. All items were chosen following the literature that was used in this study. Item one is defined as the influence of age and gender ((Lazányi, 2018); (Lazanyi & Maraczi, 2017); (Gold, Körber, Hohenberger, Lechner, & Bengler, 2015); Walker, Boelhouwer, et al., 2018). A positive result indicated that age and/or gender has an influence on trust. A negative result means that no effect was found. The second

item, reaction time as indicator, means that trust was measured by using reaction time to indicate higher or lower trust (Payre et al., 2016). A positive result indicates that reaction time was a successful measure of trust, whereas a negative result means that it was not found to be a successful measure. A study in the category of item three researched monitoring and control as an indicator to measure trust ((Hergeth et al., 2016); (Molnar et al., 2018)). A positive result indicates that the measurement was successful, but a negative result means that it was not. The items four and five are about the increase or decrease of trust over time ((Dikmen & Burns, 2017)). A positive result means that trust was found to increase or decrease, whereas a negative result would not indicate such a change. Item six entails studies that researched using gaze behavior, or eye tracking, as a measurement ((Walker, Verwey, et al., 2018); (Hergeth et al., 2016); (Kircher, Larsson, & Hultgren, 2014); (Gold et al., 2015)). A positive result means that the measurement was successful, but an unsuccessful measurement falls under a negative result. The seventh item shows the results for studies that looked at the influence of experience and familiarization on the amount of trust ((Hergeth et al., 2017); (Kircher, Larsson, & Hultgren, 2014)). When this influence was found, it was marked as a positive result. Otherwise it was marked as a negative result. Item nine includes studies that tested psychophysiological measurements ((Akash et al., 2018); (Walker, Verwey, et al., 2018); (Hergeth et al., 2016); (Kircher, Larsson, & Hultgren, 2014); (Khawaji, Zhou, Chen, & Marcus, 2015); (Hirshfield et al., 2014)). A positive result occurred when this measurement was successful. When the measurement was performed without success, it fell under a negative result. The influence of system aesthetics on trust, or item ten, was also investigated (Weinstock, Oron-Gilad, & Parmet, 2012). When an influence was found, it was included in the positive results. Otherwise, it was shown as a negative result. Finally, item 11 was the influence of situational awareness of the vehicle (SAV) on trust (Filip et al., 2016). SAV entails the vehicle's ability to sense the surrounding environment. A positive result means that the SAV did have an influence on trust, but a negative result means that it did not have an influence.

The most positive results were yielded in the case of articles that used psychophysiological measures (Akash et al., 2018; Walker, Martens, & Verwey, 2018; Hergeth et al., 2016; Kircher, Larsson, & Hultgren, 2014; Khawaji, Zhou, Chen, & Marcus, 2015; Hirshfield et al., 2014). Furthermore, age and gender were found to be mostly of influence on the amount of trust (Lazányi, 2018; Lazanyi & Maraczi, 2017; Gold et al., 2015), aside from negative findings in one of the studies (Walker, Boelhouwer, et al., 2018). Aside from these results, the items monitoring and

control (Hergeth et al., 2016; Molnar et al., 2018) and experience and familiarization also had exclusively positive results (Hergeth et al., 2017; Kircher, Larsson, & Hultgren, 2014). Gaze behavior as a measurement provided mostly positive results (Walker, Martens, & Verwey, 2018; Hergeth et al., 2016; Kircher, Larsson, & Hultgren, 2014), with the exception of one negative result (Gold et al., 2015). Lesser researched items were reaction time as indicator of trust (Payre et al., 2016), increased trust over time (Dikmen & Burns, 2017), the influence of knowledge (Hergeth et al., 2017), as well as the influence of system aesthetics (Weinstock, Oron-Gilad, & Parmet, 2012) and the influence of SAV (Filip et al., 2016). All items, except for item ten, did have positive results. Finally, the item of decreased trust over time did not show up in any results.

Discussion

In recent years, the interest in the subject of trust in automated vehicles has increased. This could be an indicator of a general interest in automated vehicles, or even an increased focus on the impact of trust. It can also mean that there has been a higher demand for the development automated vehicles. The gap in articles found between the years 2004 and 2014 might hold no meaning at all. The criteria based on which the articles were selected could have been too strict to show the general interest in the topic. However, the specific subject is being researched more and more. Furthermore, the interest in the specific levels of automation is highest for the unspecified level of automation, as well as for level 3 automation. Some studies in the category of unspecified level of automation did indicate that a 'high level of automation' was being tested (Payre et al., 2016; Hergeth et al., 2016; Walker, Martens, & Verwey, 2018). Therefore, it could be said that either level 3 of automation and above, or the less specified 'high automation', are the most researched levels of automation. It could also be hypothesized that the interest in the higher levels of automation has been growing. Possibly, there exists an idea that the highest level of automation would be the most beneficial level to incorporate in vehicles. However, these claims needs to be investigated further, as this is more speculation than fact. This issue will also be investigated further later in this section.

Subsequently, the experiment materials and measuring methods were compared. The most used experiment material was the driving simulator. Questionnaires are the most used measuring method. However, most articles did indicated that neither of these were the most dependable environment or measure (Walker, Boelhouwer, et al., 2018; Payre et al., 2016; Hergeth et al., 2016; Khastgir et al., 2018; Molnar et al., 2018; Gold et al., 2015; Walker, Martens, & Verwey, 2018). For example, a driving simulator might cause bias that a real-life setting can account for (Walker, Boelhouwer, et al., 2018; Payre et al., 2016; Molnar et al., 2018; Gold et al., 2015; Walker, Martens, & Verwey, 2018). Some studies indicated that participants trusted the driving simulator too much, because there was no real danger if something went wrong. However, this does imply that a real driving test might be too dangerous. Not only that, it can be speculated that it is simply too costly as well. Moreover, perhaps driving simulators can be improved upon to make it more life-like (tactical driving behavior). For example, one study explains that the driving simulator that was used can mimic the movement and vibration that can be felt in a real car. Additions like these could solve the problem of over-trust in participants. If a real-life driving test is being performed

after all, a suggestion made by Kaur and Rampersad (2018) is that it could be tested in a closed-off environment. This could improve the safety of the driving test as there is less traffic.

Furthermore, other researchers suggested to replace questionnaires with a more objective and continuous measure (Noah et al., 2017; Molnar et al., 2018; Walker, Martens, & Verwey, 2018). Many studies agreed that a questionnaire was not sufficient to measure something as fluid as trust. Therefore, a timeline with the most used methods per year was established to look at possible developments over time. This did show a greater variety in the methods that were used, but it did not yet provide a concluding answer to which methods were the most effective or promising. The only speculation that could be made, was that this greater variety was caused by the search for a better, alternative method.

The last analysis was the investigation into the results of each study. In this analysis, age and gender were shown to have some influence on trust. Not all studies were conclusive on this, but it could be something to investigate further. Also, the articles that were used in this study did not find a decrease in trust over time. Furthermore, promising results were shown in studies that investigated psychophysiological measurements. This is an interesting find, as these kinds of measurements were not used very often so far. More research is necessary to determine the reliability of psychophysiological measurements. However, it is probably not recommended to simply replace questionnaires with psychophysiological measurement. Questionnaires do provide benefits in regards to structuring the data and creating transparency (Kaur & Rampersad, 2018). It can be very helpful to include a measurement that creates quantitative data, even if it does not grasp the fluidity of trust as well as psychophysiological measurements. A combination of both, as some studies already tried, could be a step in the right direction (Walker, Marten & Verwey, 2018; Kircher, Larsson, & Hultgren, 2014; Gold et al., 2015; Hirshfield et al., 2014).

A difficulty that was encountered in this study, was that some studies did not always provide the necessary information for the analysis that were done in this study. Especially the level of automation was not always specifically mentioned (Lazanyi & Maraczi, 2017); Walker, Marten & Verwey, 2018; Hergeth et al., 2016; Boubin et al., 2017; Kircher, Larsson, & Hultgren, 2014). This made for inconclusive results in the analysis on which level of automation was the most analyzed. For example, some studies would indicate that a 'high level of automation' was being researched (Walker, Marten & Verwey, 2018; Hergeth et al., 2016). Perhaps it was meant to indicate level three or four of automation, as this level corresponds with a high level of automation

according to the SAE criteria (Kelechava, 2018; (Hergeth et al., 2017). However, because it is not reliable to assume what the authors meant exactly, it cannot be said with certainty. Another example is a study in which ‘multiple levels of automation’ were being researched (Kircher, Larsson, & Hultgren, 2014). Yet, the description of the level(s) that were being researched seemed to only indicate a level three of automation. All in all, it might be that there was more interest in the higher levels of automation than the results indicate. On the other side of the scale, there seemed to be no interest in the lowest levels of automation. This could be explained by the fact that these kinds of vehicles have been around for some time already. However, it can still be interesting to compare the different levels of automation to see if there is any difference in trust. Not only the levels of automation were not consistently reported in a clear manner. When investigating the different types of questionnaires, it was not always clear what type had been used (dispositional trust; fully; understanding trust). These two findings make the replicability of these studies questionable. Perhaps the researchers were not aware of the importance of specifically indicating the level of automation, or that the difference between the levels was not clear to them. It could also be the case that it was assumed that the information was already clear enough. In the case of questionnaires, it might be that questionnaires are so widely used that it was not deemed necessary to fully explain the process or the questions that were asked. No matter the possible explanations, it does seem to indicate a problem that should be looked into.

Limitations

There were several limitations to this study. Firstly, the criteria for the inclusion of articles may have been too strict. This showed in the analysis of the interest in the topic over the years. There was a period in between 2004 and 2014 in which no articles were found. It is possible that a more insightful analysis could have been conducted if more general criteria had been used. This way, the general interest in the topic could have been displayed in a more inclusive manner. Secondly, the keywords that were used might have contributed to the first limitation. The keyword ‘self-driving cars’ was not used after the first search in Scopus. However, the assumption was made that this term was so similar to ‘automated vehicles’ that it would not provide major different search results. It would have been more thorough if an analysis on the two keywords was performed to indicate if it would indeed make no difference. By having prematurely excluded one keyword, some important articles might be missing from this study.

Conclusion

Despite these limitations, this study does provide useful insights that can be used in other research. Unfortunately, the question arose whether all studies included in this study were completely replicable. This was due to some missing information, especially about the level of automation and in some cases the questionnaires that were used. Despite this finding being all but desirable, it does provide the opportunity for future studies to improve the manner of reporting. This could be improved by specifically naming the level(s) of automation that are being tested, as well as indicating which questionnaire is used and how it is developed.

Other improvements that can be looked into are the experiment materials and the measurements. It could be a step in the good direction to investigate if and how a driver simulator can be built to make it more life-like. That way, participants could be more likely to have the appropriate amount of trust during an experiment. Otherwise, a real-life driving test could be tested in a closed environment as well. Aside from this, a combination of questionnaires and psychophysiological measurements could provide a good start in finding a reliable method in measuring trust.

In conclusion, this study might indicate the direction in which future research can go to understand trust better, to understand how to measure trust, and to create more replicable studies. This could eventually lead to not only a better understanding, but also the development of an automated vehicle that is better tailored to the user. Surely, the foreseen benefits of highly automated vehicles could become a reality in the future.

References

- Akash, K., Hu, W.-L., Jain, N., & Reid, T. (2018). A Classification Model for Sensing Human Trust in Machines Using EEG and GSR. *ACM Transactions on Interactive Intelligent Systems*, 8(4), 1-20. doi:10.1145/3132743
- Artinis (April 17, 2018). *Combining the world of NIRS and EEG*. Retrieved from: <https://www.artinis.com/blogpost-all/2018/4/11/combining-the-world-of-nirs-and-eeeg>
- Boubin, J. G., Rusnock, C. F., & Bindewald, J. M. (2017). Quantifying Compliance and Reliance Trust Behaviors to Influence Trust in Human-Automation Teams. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 61(1), 750-754. doi:10.1177/1541931213601672
- Bui, V., Verhoeven, R., Lukkien, J., & Kocielnik, R. (2013). *A trust evaluation framework for sensor readings in body area sensor networks*. Paper presented at the Proceedings of the 8th International Conference on Body Area Networks.
- Choi, J. K., & Ji, Y. G. (2015). Investigating the Importance of Trust on Adopting an Autonomous Vehicle. *International Journal of Human-Computer Interaction*, 31(10), 692-702. doi:10.1080/10447318.2015.1070549
- Dikmen, M., & Burns, C. (2017). *Trust in autonomous vehicles: The case of tesla autopilot and summon*. Paper presented at the 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC).
- Dixon, G., Hart, P. S., Clarke, C., O'Donnell, N. H., & Hmielowski, J. (2018). What drives support for self-driving car technology in the United States? *Journal of Risk Research*, 1-13. doi:10.1080/13669877.2018.1517384
- Filip, G., Meng, X., Burnett, G., & Harvey, C. (2016). Designing and calibrating trust through situational awareness of the vehicle (SAV) feedback.

- Gold, C., Körber, M., Hohenberger, C., Lechner, D., & Bengler, K. (2015). Trust in Automation – Before and After the Experience of Take-over Scenarios in a Highly Automated Vehicle. *Procedia Manufacturing*, 3, 3025-3032. doi:10.1016/j.promfg.2015.07.847
- Hergeth, S., Lorenz, L., & Krems, J. F. (2017). Prior Familiarization With Takeover Requests Affects Drivers' Takeover Performance and Automation Trust. *Hum Factors*, 59(3), 457-470. doi:10.1177/0018720816678714
- Hergeth, S., Lorenz, L., Vilimek, R., & Krems, J. F. (2016). Keep Your Scanners Peeled: Gaze Behavior as a Measure of Automation Trust During Highly Automated Driving. *Hum Factors*, 58(3), 509-519. doi:10.1177/0018720815625744
- Hirshfield, L. M., Bobko, P., Barelka, A., Hirshfield, S. H., Farrington, M. T., Gulbranson, S., & Paverman, D. (2014). Using Noninvasive Brain Measurement to Explore the Psychological Effects of Computer Malfunctions on Users during Human-Computer Interactions. *Advances in Human-Computer Interaction, 2014*, 1-13. doi:10.1155/2014/101038
- Hoff, K. A., & Bashir, M. (2015). Trust in automation: integrating empirical evidence on factors that influence trust. *Hum Factors*, 57(3), 407-434. doi:10.1177/0018720814547570
- Kaur, K., & Rampersad, G. (2018). Trust in driverless cars: Investigating key factors influencing the adoption of driverless cars. *Journal of Engineering and Technology Management*, 48, 87-96. doi:10.1016/j.jengtecman.2018.04.006
- Kelechava, B. (September 12, 2018) *SAE Levels of Driving Automation*. Retrieved from: <https://blog.ansi.org/2018/09/sae-levels-driving-automation-j-3016-2018/#gref>
- Khastgir, S., Birrell, S., Dhadyalla, G., & Jennings, P. (2017). Calibrating Trust to Increase the Use of Automated Systems in a Vehicle. In *Advances in Human Aspects of Transportation* (pp. 535-546).

- Khastgir, S., Birrell, S., Dhadyalla, G., & Jennings, P. (2018). Calibrating trust through knowledge: Introducing the concept of informed safety for automation in vehicles. *Transportation Research Part C: Emerging Technologies*, *96*, 290-303. doi:10.1016/j.trc.2018.07.001
- Khawaji, A., Zhou, J., Chen, F., & Marcus, N. (2015). *Using Galvanic Skin Response (GSR) to Measure Trust and Cognitive Load in the Text-Chat Environment*. Paper presented at the Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '15.
- Kircher, K., Larsson, A., & Hultgren, J. A. (2014). Tactical Driving Behavior With Different Levels of Automation. *IEEE Transactions on Intelligent Transportation Systems*, *15*(1), 158-167. doi:10.1109/tits.2013.2277725
- Lazányi, K. (2018). *Are we Ready for Self-Driving Cars-a Case of Principal-Agent Theory*. Paper presented at the 2018 IEEE 12th International Symposium on Applied Computational Intelligence and Informatics (SACI).
- Lazanyi, K., & Maraczi, G. (2017). *Dispositional trust—Do we trust autonomous cars?* Paper presented at the 2017 IEEE 15th International Symposium on Intelligent Systems and Informatics (SISY).
- Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human factors*, *46*(1), 50-80.
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., . . . Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS medicine*, *6*(7), e1000100.
- Molnar, L. J., Ryan, L. H., Pradhan, A. K., Eby, D. W., St. Louis, R. M., & Zakrajsek, J. S. (2018). Understanding trust and acceptance of automated vehicles: An exploratory simulator study

of transfer of control between automated and manual driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 58, 319-328. doi:10.1016/j.trf.2018.06.004

Noah, B. E., Wintersberger, P., Mirnig, A. G., Thakkar, S., Yan, F., Gable, T. M., . . . McCall, R. (2017). *First Workshop on Trust in the Age of Automated Driving*. Paper presented at the Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct - AutomotiveUI '17.

Payre, W., Cestac, J., & Delhomme, P. (2016). Fully Automated Driving: Impact of Trust and Practice on Manual Control Recovery. *Hum Factors*, 58(2), 229-241. doi:10.1177/0018720815612319

SAE. (2014). Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems. *SAE International*.

Schaefer, K. E., Chen, J. Y., Szalma, J. L., & Hancock, P. A. (2016). A Meta-Analysis of Factors Influencing the Development of Trust in Automation: Implications for Understanding Autonomy in Future Systems. *Hum Factors*, 58(3), 377-400. doi:10.1177/0018720816634228

Vecchiato, G., Toppi, J., Maglione, A. G., Olejarczyk, E., Astolfi, L., Mattia, D., . . . Babiloni, F. (2014). Neuroelectrical correlates of trustworthiness and dominance judgments related to the observation of political candidates. *Comput Math Methods Med*, 2014, 434296. doi:10.1155/2014/434296

Walker, F., Boelhouwer, A., Alkim, T., Verwey, W. B., & Martens, M. H. (2018). Changes in Trust after Driving Level 2 Automated Cars. *Journal of Advanced Transportation*, 2018, 1-9. doi:10.1155/2018/1045186

Walker, F., Martens, M., & Verwey, W. V. (2018). *Gaze Behaviour as a Measure of Trust in Automated Vehicles*. Paper presented at the Proceedings of the the 6th Humanist Conference (June 2018).

Wang, M., Hussein, A., Rojas, R. F., Shafi, K., & Abbass, H. A. (2018). *EEG-Based Neural Correlates of Trust in Human-Autonomy Interaction*. Paper presented at the 2018 IEEE Symposium Series on Computational Intelligence (SSCI).

Weinstock, A., Oron-Gilad, T., & Parmet, Y. (2012). The effect of system aesthetics on trust, cooperation, satisfaction and annoyance in an imperfect automated system. *Work, 41 Suppl 1*, 258-265. doi:10.3233/WOR-2012-0166-258

Appendix A

N.	Title	Level of automation
4	Are we ready for self-driving cars – a case of principal-agent theory	All levels of automation
5	Calibrating trust through knowledge: Introducing the concept of informed safety for automation in vehicles	Level 4
7	Changes in trust after driving level 2 automated cars	Level 2
9	Dispositional trust – Do we trust autonomous cars?	Not specified
12	Fully automated driving: Impact of trust and practice on manual control recovery	Level 3 and above
13	Gaze behavior as a measure of trust in automated vehicles	High automated driving
15	Keep your scanners peeled: Gaze behavior as a measure of automation trust during highly automated driving	High automation
17	Prior familiarization with takeover requests affects drivers' takeover performance and automation trust	Level 3
18	Quantifying compliance and reliance trust behaviors to influence trust in human-automation teams	Not specified
19	Tactical driving behavior with different levels of automation	Multiple different levels
21	Trust in automation –Before and after the experience of take-over scenarios in a highly automated vehicle	Level 3
24	Trust in autonomous vehicles: The case of Tesla autopilot and summon	Level 2
26	Understanding trust and acceptance of automated vehicles: An exploratory simulator study of transfer of control between automated and manual driving	Level 3

Appendix B

N.	Title	Type of questionnaire
4	Are we ready for self-driving cars – a case of principal-agent theory	Online questionnaire developed by another person, not specified further.
5	Calibrating trust through knowledge: Introducing the concept of informed safety for automation in vehicles	Questionnaire after the experiment only. Three different questionnaires were used: simulator sickness questionnaire (SSQ), Van Der Laan's acceptance questionnaire, and a trust rating questionnaire. Only the results from the last questionnaire are shown in this study. This questionnaire used a rating in percentages from 0 to 100. 0 is low trust and 100 very high trust. Anything in between was interpreted as partial trust.
7	Changes in trust after driving level 2 automated cars	Questionnaire before, after, and 2 weeks after the experiment. The questionnaire was a 5-point likert scale questionnaire and it was self-made.
8	Designing and calibrating trust through situational awareness of the vehicles (SAV) feedback	Questionnaire on a 7-point likert scale. It consisted of modified questions that were already used in another study, and questions derived from literature.
9	Dispositional trust – Do we trust autonomous cars?	Online questionnaire, not specified further.
12	Fully automated driving: Impact of trust and practice on manual control recovery	Computerized questionnaire before and after the experiment, not specified further.

13	Gaze behavior as a measure of trust in automated vehicles	Questionnaire with a 7-point likert scale. It was a modified version of the ED (Emperically Derived) Trust Scale and was filled in online at home before the experiment, and in person after the experiment. The experiment was concluded with a questionnaire about the level of education and prior experience with automated vehicles.
14	Investigating the importance of trust on adopting an autonomous vehicle	Questionnaire with a 7-point likert scale, consisting of constructs extracted from literature. All these constructs were measured by using already existing methods of formulating questions.
17	Prior familiarization with takeover requests affects drivers' takeover performance and automation trust	Questionnaire with a 7-point likert scale before and after the experiment. It was an adapted version of an already existing scale which was not named specifically.
19	Tactical driving behavior with different levels of automation	Questionnaire on demographics and prior experience before the experiment. Questionnaire on experience, attitude and trust after the experiment.
20	The effect of system aesthetics on trust, cooperation, satisfaction and annoyance in an imperfect automated system	Questionnaire with a 7-point likert scale at the beginning and at the end of the experiment. It consisted of self-made questions and the Halden Human-Automation Cooperation Questionnaire.
21	Trust in automation – Before and after the experience of take-over scenarios in a highly automated vehicle	A questionnaire on demographics and driving experience was conducted before the experiment. After the experiment, a questionnaire on attitude was conducted.

24	Trust in autonomous vehicles: The case of Tesla autopilot and summon	A self-made online survey about attitudes and experiences towards Tesla vehicles with both a 5-point and 4-point likert scale.
25	Trust in driverless cars: Investigating key factors influencing the adoption of driverless cars	Online survey using a 5-point likert scale and multiple choice questions for demographics.
26	Understanding trust and acceptance of automated vehicles: An exploratory simulator study of transfer of control between automated and manual driving	A self-administered questionnaire, not specified further.
28	Using noninvasive brain measurement to explore the psychological effects of computer malfunctions on users during human-computer interactions	Questionnaire after each experiment session consisting of already existing questionnaires, namely the TLX (Task Load Index) and the SAM (Self-Assessment Manikin).
29	What drives support for self-driving car technology in the United States?	A self-made online survey, self-made with different types of questions such as likert scales (different points), statements, and demographic questions

Appendix C

Year	N.	Title	Measurement
2012	20	The effect of system aesthetics on trust, cooperation, satisfaction and annoyance in an imperfect automated system	- Questionnaire
2013	3	A trust evaluation framework for sensor readings in body area sensor networks	- GSR - ECG
2014	16	Neuroelectrical correlates of trustworthiness and dominance judgments related to the observation of political candidates	- EEG - GSR - HR
2014	19	Tactical driving behavior with different levels of automation	- Questionnaire - Eye tracking
2014	28	Using noninvasive brain measurement to explore the psychological effects of computer malfunctions on users during human-computer interactions	- Questionnaire - fNIRS - GSR
2015	14	Investigating the importance of trust on adopting an autonomous vehicle	- Questionnaire
2015	21	Trust in automation –Before and after the experience of take-over scenarios in a highly automated vehicle	- Questionnaire - Eye tracking
2015	27	Using galvanic skin response (GSR) to measure trust and cognitive load in the text-chat environment	- GSR
2016	8	Designing and calibrating trust through situational awareness of the vehicles (SAV) feedback	- Questionnaire - Open questions - Interviews
2016	12	Fully Automated Driving: Impact of Trust and Practice on Manual Control Recovery	- Questionnaire

2016	15	Keep your scanners peeled: Gaze behavior as a measure of automation trust during highly automated driving	- Eye tracking
2016	17	Prior familiarization with takeover requests affects drivers' takeover performance and automation trust	- Questionnaire
2017	9	Dispositional trust – Do we trust autonomous cars?	- Online questionnaire
2017	24	Trust in autonomous vehicles: The case of Tesla autopilot and summon	- Online questionnaire
2018	1	A classification model for sensing human trust in machines using EEG and GSR	- GSR - EEG
2018	4	Are we ready for self-driving cars – a case of principal-agent theory	- Online questionnaire
2018	5	Calibrating trust through knowledge: Introducing the concept of informed safety for automation in vehicles	- Questionnaire
2018	7	Changes in trust after driving level 2 automated cars	- Questionnaire
2018	10	EEG-based neural correlates of trust in human-autonomy interaction	- EEG
2018	13	Gaze behaviour as a measure of trust in automated vehicles	- Questionnaire - Eye tracking
2018	25	Trust in driverless cars: Investigating key factors influencing the adoption of driverless cars	- Questionnaire
2018	26	Understanding trust and acceptance of automated vehicles: An exploratory simulator study of transfer of control between automated and manual driving	- Questionnaire - Interview

2018	29	What drives support for self-driving car technology in the United States?	- Online questionnaire
------	----	---	------------------------