Designing for Trust in Autonomous Systems

How to Use Design Methodology to Increase Peoples Trust in Autonomous Systems

Stine Marie Hjetland
Department of Product Design
Norwegian University of Science and Technology

ABSTRACT

The prevalence of autonomous systems in the world is increasing. However, several people are afraid of the change from manually, to automatic and later to autonomous solutions. The mistrust in machines and the feeling of losing control can be frightening to many people. When users of an autonomous system are not taken into account during the developing process, important human-factors to create the necessary trust in the system can be forgotten. This article addresses trust-related challenges around the development of autonomous cars. Trust in autonomous systems will be defined and discussed, as well as how a system can be developed to gain the necessary trust from the user. Furthermore, the article will explore how design methodology, especially human-centered design (HCD) and systems oriented design (SOD) can help achieve this trust. Keeping the humans in the developing loop adds necessary insight about the needed human-machine interface (HMI). SOD helps gain the necessary overview of the complexity in taking autonomous cars in use. A fully autonomous system is technical independent, but, as in the case of driverless cars, still depends on peoples trust to function properly. As people have different needs, and trust is personal, it is important to have an approach which can grasp the complexity. HCD and SOD embrace the diversity and combine user needs with technology.

KEYWORDS: Autonomous systems, Human-centered design, Systems oriented design, Trust

1. INTRODUCTION

Autonomous systems are included more and more in our everyday lives. People are getting used to robotic lawnmowers and vacuum cleaners, drones, and now cars are driving autonomous in California and Texas [1]. Over the last few decades an increasingly number of systems have been equipped with more automatic design elements such as elevators and automatic braking in cars [2]. Either people have to adapt to new technology, and new ways of living, or the technology has to adapt to humans and their new ways of living.

Today, the development of autonomous cars is increasing rapidly, and experts are assuring us that the truly driverless car will be reality in immediate future [3, 4]. 14. October 2015, Tesla launched their new autopilot system [3], making it possible to drive autonomously in certain areas, under supervision of the human driver. While some people praise this innovation, others condemn it [4]. One can argue that a machine conduct many chores better than a human being.

As an example the Tesla autopilot reaction time is 16 milliseconds, compared to human's 150 milliseconds [4]. Even though it is mostly human error that leads to car crashes [2], people seem to be less forgiving if an automated system fails than if a human causes the same error [5]. Following, when humans have the wrong amount of trust in a system the human-machine interaction fails and wrong choices are made.

1.1 Previous Work

Bainbridge [6] describes how developers often look at the human as the weakest link in a system. A correct programmed machine could perform a task more precise than a human. It does not get tired and the chance of human errors are almost excluded. Then again, systems can break down of or fail in other ways [7]. Dzindolet et al. [8] describe how people initially can trust automated systems, but experiencing the system making errors leads to distrust even in reliable systems. Bainbridge [6] states that there is no perfect system, and that the chance of human error will always be present. Parasuraman and Riley [9] stresses how under-trust or over-trust in a system leads to bad human-machine interaction and disuse or misuse of the system. There have been several severe accidents due to wrong amount of trust in automatic or autonomous systems [10, 11].

Trust in automatic and autonomous systems are to some extent comparable. Autonomization of systems can lead to many of the same results as automation of systems, including lack of control and replacement of humans [6, 7, 12]Some research has been done on automated systems, mainly since the early 1980's [6, 13]. The levels of automation increases as new technology develops, from automatic to fully autonomous. The focus has been on systems in general, their technical aspects and people's reactions to such systems. Throughout this article, the term autonomous system is used to refer to fully autonomous systems.

A big part of the follow-up research is aimed at automation in planes [14-17]. Research have been done on why and how accidents happens when coping with an autopilot in planes, and how this danger can be excluded. There are many similarities between autonomization of planes and cars, especially human factors. One have to keep in mind that pilots are trained to interact with the system e.g. autopilot, while private drivers probably are not. Still, it is likely to believe that the same questions as Wiener [16] used to describe pilots' frustration towards the autopilot – "What is it doing?", "Why is it doing that?", and "What is it going to do next?"- also are relevant for people's early interactions with autonomous cars. Dzindolet et al. [8], Parasuraman and Riley [9], Johnson et al. [18], and Madhavan and Wiegmann [19] all stresses how trust affects human-machine interaction. Additionally, Wiener and Curry [17] emphasize psychosocial problems as one of the main issues in automation.

Wiener [16] strived to make autopilot in planes into a team player, empathizing the human-machine relationship to create the necessary trust in the system. Trust affects the human-machine interaction [8, 9, 17-19], the human-machine interaction can affect the outcome of autonomous systems [6, 7, 9], and the outcome affects the trust [5, 8].

Human-centered design take the user needs and aspect into account [20, 21]. Kolko [22] empathizes design thinking (DT) as a tool to make systems and products more user-friendly and understandable. Sevaldson [23] stresses how the problems of today gets increasingly complex and how designers are especially good in coping with such situations, balancing all aspects of the problem. System oriented design (SOD) offers a systematic design approach to the problem solving [23, 24].

Some researchers have focused on trust in autonomous cars. Parasuraman and Riley [9], Lee and Moray [25], Lee and See [26], Muir [27], Muir and Moray [28] and Riley [29] show that trust is a critical psychological factor in interactions with automated systems. Moreover, Marsh and Meech [30], and Hoff and Bashir [10] argue that personality, education and experience strongly affects adaption to automated systems. Still, the concept of trust develops along with technological developments. Based on several news articles and peoples shared experiences with partly autonomous vehicles [3, 4, 31], trust is one of the big issues.

1.2 Contribution of this Paper

Few have, to the author's knowledge, tried to find a way to develop for trust, only pointed out the need for it. As autonomous systems have many influencers, the approach to gain trust gets more complex. This article will focus on trust-related challenges within the development of autonomous systems:

- What the user needs in order to develop the necessary trust in an autonomous system.
- How design methodology can be used to establish trust.
- Which methods that can be included in the developing process.

The question around whether or not the autonomous system is safe is important, but will not be discussed in this article.

1.3 Outline of this Paper

This paper is based on a literature review that has been conducted, including research on automatic and autonomous systems, human factors, trust and psychological challenges in human-machine interfaces, and design methodology. The actuality of the development of autonomous cars with its corresponding problems has been investigated through newspapers and relevant social forums such as blogs. An illustration example has been conducted in order to give a rough impression of potential situations that passengers are in when they are in an autonomous car. This paper is organized as follows: Section 2 explains and defines what an autonomous system is, in Section 3 trust is defined. Both Section 2 and 3 includes examples around autonomous cars. Further, relevant design methodology are described in Section 4. Then the findings from the illustration example are presented in Section 5. In Section 6 the article will discuss how and why designers and design methodology should be included in the developing process of autonomous cars. Finally, the paper is concluded in Section 7.

2. AUTONOMOUS SYSTEMS

This chapter should give the reader an understanding of what an autonomous system is. Further, it will explain the different levels of automation and how automation relates to autonomy.

2.1 Defining Autonomy

Autonomous means self-ruled [32]. An autonomous system is not to be mistaken as an automatic system, even though they are both automated systems. As illustrated below in figure 1, an automatic system is less automated than an autonomous system. An automatic system is a system that can replace humans in small, repeating tasks [33]. Compared to an autonomous system the automatic system is not self-ruled, it simply conducts the tasks it is programmed to execute. Elevators and escalators are examples of automatic systems.

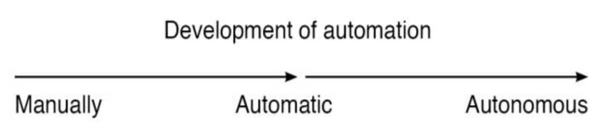

Figure 1: The development of automation.

The definition of automation varies across different sectors and the amount of human-machine interaction. As stated by Parasuraman, Sheridan and Wickens [34] automation is a system or device that adopts functions humans earlier had or could have had, partially or fully. Thus, it is possible for automation to replace functions of a human operator both partly or fully. An autonomous system is fully automated and functions independently, approximately without interference from another part such as a human. A truly self-driving car is an example of such a system.

Systems and devices can be automated to different extent. It can help and simplify an operator's work, or fully take over the operator's tasks. Automation can vary in comprehensiveness. Parasurman, Sheridan et al. [34] describe ten different levels of automation, as seen in Table 1, level one being low automation and level ten being high, and is based on the table from Sheridan [35]. That is, Table 1 shows the steps from no automation, to automatic systems and further on to fully autonomous systems.

Levels of Automation of Decision and Action Selection

HIGH 10.	The computer decides everything, acts autonomously, ignoring the human.
9.	Informs the human only if it, the computer, decides to
8.	Informs the human only if asked, or
7.	Executes automatically, then necessarly informs the human, and
6.	Allows the human a restricted time to veto before automatic execution, or
5.	Executes that suggestion if the human approves, or
4.	Suggest one alternative
3.	Narrows the selection down to a few, or
2.	The computer offers a complete set of decision/action, or
1. LOW	The computer offers no assistance: human must take all decisions and actions.

Table 1: Table from Parasurman, Sheridan et al. [34].

The advantage with automation is that it can increase productivity and exclude mortal danger in some situations [12, 33]. System and devices can replace humans if they work in a risky environment. A considerable disadvantage is, as discussed by Bainbridge [6], that automation can lead to less training for humans in manual control, which can lead to difficulties in a potential necessary manual take-over.

2.2 The Increase of Autonomous Systems

Throughout history, we have developed devices and systems with different levels of autonomy. Planes are approximately steering themselves, and the driver assistance in cars gets rapidly more advanced. Cruise control, parking assistance and now Tesla's upgraded autopilot system. Tesla's upgrade makes it possible for the car to switch lanes and drive on its own, but the driver is still recommended to keep his or her hands on the steering wheel [4]. The autopilot will let the driver know when there is need for a manual take over, in situations or environment it cannot handle. All necessary hardware have existed in Teslas Model S for a long time, but it is the software update, Tesla Version 7.0, which gives the car autonomous features [4, 31]. The necessary hardware includes a forward radar, a forward-looking camera, 12 long-range ultrasonic sensors positioned to sense 16 feet around the car, and a high-precision digitally-controlled electric assist braking system [3]. Inside, the car communicates the autonomous take over on the already existing screen at the dashboard[4], shown in figure 2.

Figure 2: Photo of the Tesla screen. Foto: Torgeir Strandberg [4].

All Tesla cars share their experiences. This way the cars can learn, just like humans, based on each other's earlier mistakes [4]. Ironically, it is easier to drive and control a car in areas of higher speed such as highways and country roads. These roads are adapted to high speed and to decrease disturbances from the surroundings. In more urban environments, the infrastructure is more complex, and there are more factors interfering with the partly autonomous systems, such as traffic, people, and traffic lights. However, there are still chances of meeting unplanned challenges at autonomous friendly roads. Snow and ice can for example be difficult to predict precise despite online access to weather forecasts. Such unexpected events is a reason for keeping the human aware and present while driving.

2.3 Open System

Up until today, car developers have tried to maintain the system around cars as closed as possible. Thus, as controlled and with few external influences. This can be done by only using input already existing in the car e.g. speedometer, and not include functions that is dependent of other influences in the cars' environment. An example is automated features such as cruise control that lock the car to a steady speed set by the driver and include the driver as the main operator. Another small feature in many cars is automatic braking, if the cars sense possible collision they apply the brakes [2]. These automatic features does not influence the overall infrastructure as much as autonomous system does. In comparison to autonomous systems, automatic systems are small and isolated and maintain the human as an overall operator.

What makes autonomous systems so challenging to implement and difficult to trust, is that the system, unlike automatic systems, are open. A closed system like for instance an elevator, functions without interacting with the outside. Autonomous systems, especially autonomous cars, are affected by the surroundings and therefore more like an open system. An open system gets disturbances from the environment [36], which makes it harder to predict possible dangerous situations. An autonomous car must always be aware of its surroundings. It is dependent on feedback to drive safely. The car's environment includes other cars, road marking, pedestrians, all kinds of weather, animals, amongst others. This is why it is easier to introduce autonomous cars on e.g. highways, where there is less of such disturbances. If the autonomous system fails to register any environmental factors it can have fatal consequences. Thus, getting the infrastructure around autonomous cars to be as controlled as possible would make the system more closed, and further easier to predict upcoming events.

2.4 Society and Culture

Culture differences play an important role in the implementation of an automated system [12]. This also applies to the implementation of an autonomous system. An example is the considerations one would have to make when designing electronic displays containing Arabic or Hebrew words. These written languages are read from right to left, and therefore need another layout than displays containing other languages [12]. This is just one very general example of how culture can affect the design of a system. As for an open system e.g. autonomous cars, there are high demands to for example infrastructure. Thus, such a system will probably function better in a place with well-developed roads and traffic shielding, compared to a place with unorganized roads. In some cases it might be easier to adapt the surroundings to new technology, instead of the other way around. There are still some things autonomous cars can not yet register. Pedestrians' intuition, and their glance-based awareness is two examples of this. Keeping the human as an controller could help receive and process all the feedback the car needs.

2.5 People and Systems

Billings [14] and Wiener [16] describes the problem of the human-machine communication. In aviation does, not only, the pilots struggle to understand what the autopilot is doing, but the system often fails to inform them about it. When people do not understand why the system does something, they might not agree on the action [8], and try to overrule the system.

Another aspect of automation is that when humans start monitoring instead of conducting tasks the operator-skills fade [6, 12]. In a situation where it is need for a manual takeover, this lack of competence can lead to a crisis. The paradox of automation is, as Bainbridge [6] states, that there always will be a human somewhere in the process. It will never be fully autonomous, but always a human-machine system. A human will always be there to produce, maintain, use or improve the system, and there will always be a chance of human errors. Thus a *perfect autonomous system* is considered impossible [6]. This means that autonomy might fix the human errors that we know of, but it does not eliminate them, just moves and changes them.

As humans always will be somewhere in automated systems, even in autonomous systems, they should be taken into account in the development process. To increase safety it might even be necessary to keep people as a more active part in the system to maintain awareness and training, and avoid them turning passive. Activation of the passengers keeps them familiar with the HMI, and capable of possible manual takeovers. By including people more the demands to e.g. infrastructure around the partly autonomous system decreases. People can register what the cars still struggles with, like invisible communication such as intuition and eye-contact. When humans are more involved with the system it sets higher demands to the human-machine interaction. It is important that the system is understandable and informative to the person interacting with it.

3. THE HUMAN FACTOR – TRUST

This chapter explains what trust is. How the level of trust increases and decreases, as well as how it affects peoples interaction with autonomous systems.

3.1 Defining Trust

Hoff and Bashir [10] define trust as a mental state. Blomqvist [37] describes two possible extremities of trust. In the first situation there is no information about a product/system, and in the second all information is known. In case number one [37] claims there is no trust, only faith or gambling. In the second case there is only rational calculations. Thus, trust is somewhere in between. One extreme state is not necessarily better than the other. Little information increases uncertainty and suspense. A lot of information increases the knowledge on what it takes for the system to fail, which might not create more trust. Marsh and Meech [30] describe trust as situational and dynamic. Trust grows and shrinks based on experience and time. For instance, the experience you have with a certain brand will affect how you trust a new product from the same brand. A well-known, and recognised brand will probably have more trust from people, than a unknown or less acknowledge brand. This will not be further discussed in this paper, but it has a remarkable influence on peoples' trust and should be investigated further.

Hoff and Bashir [10] claim trust varies in three layers, adding dispositional trust and learned trust to situational trust. In other words trust varies based on who you are, what background you have and what situation you are in. Thus, trust is individual and different from person to person. Trust is also affected by earlier experience with a product or system. Your own or other people's experience influence your trust in something. People who are cultural alike might have more similar conditions for trust, but there will probably still be individual differences amongst them.

3.2 Trust in Systems

Parasuraman and Riley [9] define disuse and misuse of autonomous systems as under-trust or over-trust the system, respectively, e.g. manually interfering with the system due to wrong amount of trust. Even though there are many examples in history of situations where peoples disuse or misuse of systems have led to fatal accidents, people are less likely to forgive when an *automated* system causes the fail [5]. The situations caused by wrong amount of trust from the human to the machine show the necessity of right reliance in automated systems. In 2012 the cruise ship Costa Concordia sank, killing 32 passengers, after what seems to be the captain under-trusting the ship's navigation system. The captain took manual control over the ship causing it to crash [10].

The Air France Flight 447 in 2009 [11] is another example of a fatal accident assumed caused by errors in the human-machine interaction. According to CNN the accident was most likely caused by a variety of factors; failed communication and interpretation between the pilots and the system, amongst other reasons [11]. Langewiesche [38] states that the accident would most likely not have happened if the pilots had not interfered with the system. The pilots did not understand what the autopilot was doing, and ended up stalling the plane and shortly after, crashing into the ocean, killing all 228 people onboard [38].

Even though there are examples of human-error in airplane accidents, many pilots feel that replacing them with an autopilot is not a good solution. The automation is not adequate to insure safety in an emergency situation, and therefore there is need for three pilots [12]. Few pilots seem to trust the autopilot with the full responsibility.

3.3 Trust and Psychology

Autonomous cars is a hot topic, but people are skeptical. Who would send their child off to soccer practice in a self-driving car? When it comes to automated systems, Parasuraman and Miller [39](p.52) describe trust as the "users willingness to believe information from a system and make use of its capabilities." Thus, the ways in which the user receives, e.g. understands, and interpreters, information forms the trust between the human and the system. The way a person receives information is affected by the human-machine relationship, its evolvement, as well as what and how information is presented [26]. Wiener and Curry [17] were early on pointing out psychosocial aspects of automation as possibly "the most important of all" problem areas in automation, because it affects the operator's attitude towards the system. There is broad consensus in earlier research, that trust affects the human-machine interaction and therefore the use in automation [8, 9, 18, 19]. The phenomenon is important in so many relationships elsewhere and should be emphasized in HMIs as well.

3.4 Creating Trust

Experiencing an automated system failing can lead to distrust, according to Dzindolet et al. [8]. Further, it is described how trust can be gained again by informing about why a system fails. Knowing the reason why the system failed can increase the trust [8], it might also create trust. As trust varies based on situation, time and experience it varies based on who the user is, what kind of system it is and where it is. This makes facilitating for trust more complex [10]. Including focus on trust in the design process will create better interactions for the user [30]. When the human-machine interaction improves the chance of misuse and disuse of automation decreases [10].

In the development of the autopilots in planes, the team of developers have learned that they have to make the system more understandable for the pilots interacting with it. It is evenly important not to lose the communication and teamwork between the pilots [16], especially when the team consist of pilots and an autopilot. When the system becomes a team player and is understandable to the person involved the interaction will be better. The frustration around what and why a system is doing something [16] will decrease when the user feels included and informed. Intuitive systems remove the uncertainty of the human, and fatal errors can be avoided. Especially when something unexpected happens or in stressful situations, such as the Air France accident.

Trust is a decisive factor in automated systems. Including people in the system can make it safer, but it also opens for potential disuse and misuse of the system. A well developed HMI can prevent this. The system should be understandable and humanized, as a team player for the person interacting with it. By communicating, resenting and disseminating information in an intuitive way, trust can be established. As people have different needs, and trust is personal it is important to have an approach which can grasp this complexity. The person have to be seen as an individual with needs, but also as a part of the whole system to create a coherent and functional system. Knowing what the different users need of information, and how to present it creates better human-machine interaction and prevents disuse and misuse. Thus, intuitive HMI increase the chances of taking the right decision, and gaining trust. How can autonomous systems be made more intuitive for the user?

4. USING HUMAN-CENTERED DESIGN

This chapter presents different design methods that can help develop and increase trust in autonomous systems.

4.1 Including the Human in the Process

To make an automated system function well the human has to be included and taken into account in the development process. Involvement of those who have to cope with the future system in the design process creates better user-interactions [30], and helps optimize the HMI and minimize the human errors [21]. To create trust it is important to know when and why people trust automation. Human-machine systems combine engineering and psychology [12]. Designers focus on the human in a technological system and play an important part as an intermediator between technology and psychology.

Human-centered design (HCD) focuses on the user and is based on the user needs, not only on the technical possibilities. User-centered design is a subset of HCD and applies different tools to map out user needs, described by Abras et al. [20]. What does the system include to make the user interactions better, and to get the user to understand and acknowledge the system quicker? By focusing on the nature and needs of the user, applying HCD will lead to better interaction, learning and faster understanding and decrease the human errors [21].

4.2 Design Thinking

IDEO [40] describes design thinking (DT) as a human-centered approach to design and innovation. Taking into account the desire from people, the technological feasibility and the business' viability as shown in figure 3.

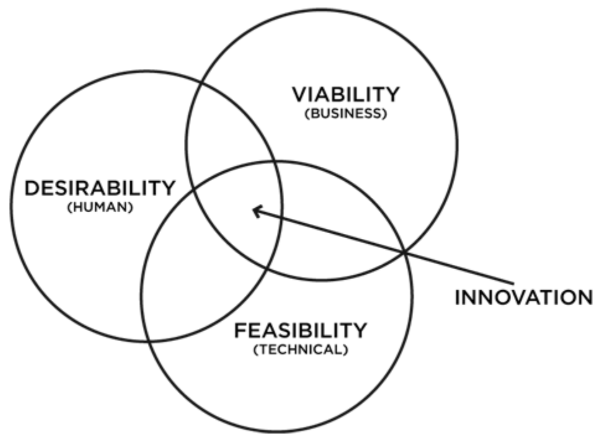

Figure 3: IDEO's visualization of design thinking [40].

DT is an iterative process of five steps; empathize, define, ideate, prototype, and test [41]. The whole DT process has to adapt to the problem area. As for testing, in the case of autonomous cars, one has to keep in mind that testing out in real life can have big consequences. DT can also be seen as different analytic, creative and user-centered method of problem solving and innovation [42].

Brown [43] mentions the complexity of the systems around autonomous cars as an example of a design thinking-problem. New radical innovations require changes in human behavior and lifestyle. This makes the problems more complex and comprehensive. Design is about improving user experiences. Good user experiences are intuitive, trustworthy and encourage reuse of the system. Kolko [22](p.70) describes DT as "an essential tool for simplifying and humanizing".

4.3 System Oriented Design

System oriented design (SOD) combines system thinking with design thinking [23]. It is a systematic design method for more complex problems [23]. In open and complex systems such as autonomous cars it is important to have the total overview of stakeholders, possible influences and outcomes. SOD helps map out the whole system while maintaining the human in center.

In SOD designers use mapping and visualization as tools to organize and understand systems. Giga mapping is a comprehensive map embracing the complexity of problems. It is used for the sake of finding blind spots and new intersections, but it is also a great way to communicate every aspect of the system that is considered [24]. Thus, the system becomes more transparent for the user.

‘Designers are especially well suited to cope with the complexity of the real world because of three reasons: they are trained to synthesise from complex and fuzzy material and they are able to visualise which is an enormous advantage for thinking in complexity. (…) Finally designers are creative people trained to come up with new solutions.’ Birger Sevaldson, Sevaldson [23]

According to Minsky [13], knowing what causes errors is almost as important as knowing what causes success, in an attempt to make a more understandable machine. As mentioned in section 2.5, autonomous systems can lead to new, unknown problems therefore it is important to minimize the possibilities of failure that we know. Humans being one side of it. According to Sevaldson [23] designers are especially good in coping with such fuzzy and unknown problems. Designers are trained to see the person but also the bigger picture, including them in the development process can make technology understandable and into a team player. Autonomous cars is a system with many stakeholders, and with SOD designers can keep a holistic perspective to optimize the system’s structure. Further, designers are trained to look for the unknown, and therefore well fitted to develop radical innovations.

5. ILLUSTRATION EXAMPLE

The illustration example was an early stage prototype of an autonomous car in a student project at The Norwegian University of Science and Technology (NTNU) in 2015. The purpose of the project was to trigger some of the same emotions and needs of the user as in an autonomous car, e.g. lack of communication, and control. The passengers was driven around in various environments and traffic in approximately one hour. They had no relations to the driver, and there was no communication between them. This was to make the driver as close to a machine as possible. In the backseat of the car there was one interviewer and one observer taking notes. The passenger was asked to describe his or her feelings and thoughts throughout the trip. During the trip the passenger was set to drawing and blindfolded to distract him or her from paying attention to the surroundings. Some unpleasant situations where staged, such as suddenly braking heavily, to create possible dangerous situations and provoke user reactions.

The illustration example showed how differently people react, even though the test persons where approximately the same age and living in the same city. They all wanted to some extent to be given more information in extreme situations, such as when the driver was heavily braking. Other than this, the test persons reactions and feelings varied greatly.

Given the variety of personal reactions during the illustration example the necessity of a human-centered approach are visible. Focusing on humans in the developing process can help discover and understand what people needs in terms of e.g. feedback, and information from the car to be able to trust it. In this brief illustration example the diversity of human reactions is evident. To grasp such a dynamic and diverse phenomenon as trust the focus has to be on human needs and at the system as a whole.

One has to keep in mind that this illustration example was a rough prototype with a few participants within the same age group. Many of the situations where staged and emotions provoked. The driver was also visible, though not communicating, but quite obviously a human. The results from the example is not to recon as a final result or answer to how people might react in autonomous cars, but more as a warning of how different people are and react in such situations. More work needs to be done to map the diversity in user reactions to different possible situation in autonomous cars.

6. DICUSSION

Trust is a human factor that is hard to grasp. It varies by situation, personality and time. As it differs between people, designing for autonomous cars can be challenging because cars are for all people. Or, at least have the potential to be, including children and disabled people. The great diversity in users and their needs require a HCD-approach. By focusing on the user, when, why and how they trust automation, one can try to implement these features in autonomous cars. Well-adapted HMIs create better information flow and receiver interpretations. Thus, it leads to trust in human-automation relationships. Making the autonomous system into a team player by providing the human with information on why and when the system is expected to react, creates trust.

Embracing the complexity in the overall system creates a more holistic and coherent user experience. SOD is a way of getting the overview of all connections, seeing what is known and unknown. New technology can cause new unknown problems, so new solutions might also introduce never previous experienced challenges. In radical innovations such as fully autonomous cars it is hard, even impossible, to predict future problems. This is why the development process for autonomous cars should be an iterative HCD process. Testing and then adjusting. This way the system will keep improving. The question is when the system is good enough for launching on the road. Again, whether or not the car is safe enough are not discussed here. Safety has a major impact on peoples trust and more research should be done around this. To minimize system failures, one must attempt to uncover as many of the unknown problems as possible.

One thing is to create the necessary trust between a human and a machine when interacting, the other is to implement it even before the interaction begins. How does HCD get a person to enter the car in the first place? Branding might be a way to establish the trust in a system before experiencing it. As mentioned more research has to be done on this, and the feelings people connect to earlier experience from the same brand. This can lead to the system being perceived as trustworthy before people have tested it. HCD helps uncover what people need, in addition to brand awareness, in terms of creating the necessary trust in advance. SOD can be a good way to map and connect other tools that can help establish trust in a system.

Trust is dynamic, and autonomous systems contains of new technology that continuously develops. When working within these areas it is important to be flexible and effective. Implementing new design methodology can be difficult and time consuming. Designers can conduct this more efficient. They are also trained to grasp the unthinkable, using SOD and DT as tools, and therefor well-fitted coping with future unknown problems.

7. CONCLUSION

This article has described how design methodology can help establish trust, and improve human-machine interaction in autonomous systems. The importance of well-developed HMIs has been elucidated, as well as how trust affects the interaction with it. The advantage of using designers in such interdisciplinary and complex innovations are explained, and use of design methodology proposed. As autonomous systems are based on new technology, there is still a lot of research to be done. More research should also be done on how trust can be established even before a system is launched.

HMI affect peoples' trust in a system. By communicating, resenting and disseminating information in an intuitive way, trust can be established. The user needs the right amount of information from the system to develop the appropriate reliance, and prevent disuse and misuse of the system. Knowing why the system failed can increase trust. Given the necessary information, and by making the autonomous system a team player, trust increases. Trust is personal and varied, thus, culture differences can affect the way people interact with systems, and what they need in order to develop trust.

To create trust it is important to know when and why people trust automation. User insights help reveal what to take into consideration in the HMI, and how to optimize the communication and facilitate for trust across space, culture and time. Well-developed HMI improves the human-machine interactions, makes the system understandable and decreases human errors.

The presented design methodologies include both technology and the human factors in the development process. HCD creates better HMIs by empathizing the user in the developing process. User-centered design uncovers user needs, and DT gives a human-centered approach to design and innovation. An autonomous car is an open system affected by its surroundings, it needs to be designed with a holistic approach. SOD creates an overview. With its tools SOD help uncover blind spots and connections. This way a more coherent solution is created. The design methodologies maintain the user's interests and trust development, to optimize the human-machine interaction. Further, HCD and SOD can together play an important part as an intermediator between technology and psychology.

REFERENCES

1. Google. *Google Self-Driving Car Project - Where we've been*. Available from: https://www.google.com/selfdrivingcar/where/.
2. Valdes-Dapena, P. *10 car companies make autmatic brakes standard*. 2015 [cited 2015 30.11.]; Available from: http://money.cnn.com/2015/09/11/autos/automatic-braking-nhtsa-iihs/.
3. Team, T.T.M. *Your Autopilot has arrived*. 2015 14.10.2015; Available from: http://www.teslamotors.com/blog/your-autopilot-has-arrived.
4. Bjørkeng, P.K. *Uten hendene på rattet - i 80 km/t!* [Online Newspaper] 2015 [cited 2015 17.11]; Available from: http://www.aftenposten.no/fakta/innsikt/Uten-hendene-pa-rattet--i-80-kmt-8245664.html.
5. Wickens, C.D., et al., *An introduction to human factors engineering, 2004*. Pearson Education International, Upper Saddle River, NJ.
6. Bainbridge, L., *Ironies of automation*. Automatica, 1983. **19**(6): p. 775-779.
7. Inagaki, T., *Automation and the cost of authority*. International Journal of Industrial Ergonomics, 2003. **31**(3): p. 169-174.
8. Dzindolet, M.T., et al., *The role of trust in automation reliance*. International Journal of Human-Computer Studies, 2003. **58**(6): p. 697-718.
9. Parasuraman, R. and V. Riley, *Humans and automation: Use, misuse, disuse, abuse*. Human Factors: The Journal of the Human Factors and Ergonomics Society, 1997. **39**(2): p. 230-253.
10. Hoff, K.A. and M. Bashir, *Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust*. Human Factors: The Journal of the Human Factors and Ergonomics Society, 2015. **57**(3): p. 407-434.
11. Jonathan. *Air France crash pilots lost vital speed data - report*. 2011; Available from: http://www.airfrance447.com/.
12. Sheridan, T.B., T. Vámos, and S. Aida, *Adapting automation to man, culture and society*. Automatica, 1983. **19**(6): p. 605-612.

13. Minsky, M., *Why people think computers can't*. AI Magazine 1982. **3**(4): p. 3-15.
14. Billings, C.E., *Aviation automation : the search for a human-centered approach*. 1997, Mahwah, N.J.: Lawrence Erlbaum Associates Publishers.
15. Sarter, N.B., D.D. Woods, and C.E. Billings, *Automation surprises*. Handbook of human factors and ergonomics, 1997. **2**: p. 1926-1943.
16. Wiener, E.L., *Human factors of advanced technology (" glass cockpit") transport aircraft*. Vol. 177528. 1989: NASA Ames Research Center.
17. Wiener, E.L. and R.E. Curry, *Flight-deck automation: promises and problems*. Ergonomics, 1980. **23**(10): p. 995-1011.
18. Johnson, N., P. Patrón, and D. Lane. *The importance of trust between operator and auv: Crossing the human/computer language barrier*. in *OCEANS 2007-Europe*. 2007. IEEE.
19. Madhavan, P. and D. Wiegmann, *Similarities and differences between human–human and human–automation trust: an integrative review*. Theoretical Issues in Ergonomics Science, 2007. **8**(4): p. 277-301.
20. Abras, C., D. Maloney-Krichmar, and J. Preece, *User-centered design*. Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications, 2004. **37**(4): p. 445-456.
21. Oviatt, S., *Human-centered design meets cognitive load theory: designing interfaces that help people think*, in *Proceedings of the 14th ACM international conference on Multimedia*. 2006, ACM: Santa Barbara, CA, USA. p. 871-880.
22. Kolko, J., *Design Thinking Comes of Age*, in *Harvard Business Review*. 2015. p. 66-71.
23. Birger Sevaldson, M.H., Peter Jones, Harold Nelson, Alex Ryan. *About Systems Oriented Design*. 2009? [cited 2015 18.11.]; Available from: http://www.systemsorienteddesign.net/index.php/information/systems-oriented-design.
24. Sevaldson, B., *Giga-mapping: visualisation for complexity and systems thinking in design*. Nordes, 2011(4).
25. Lee, J. and N. Moray, *Trust, control strategies and allocation of function in human-machine systems*. Ergonomics, 1992. **35**(10): p. 1243-1270.
26. Lee, J.D. and K.A. See, *Trust in Automation: Designing for Appropriate Reliance*. Human Factors: The Journal of the Human Factors and Ergonomics Society, 2004. **46**(1): p. 50-80.
27. Muir, B.M., *Trust in automation: Part I. Theoretical issues in the study of trust and human intervention in automated systems*. Ergonomics, 1994. **37**(11): p. 1905-1922.
28. Muir, B.M. and N. Moray, *Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation*. Ergonomics, 1996. **39**(3): p. 429-460.
29. Riley, V., *Operator reliance on automation: Theory and data*. 1996.
30. Marsh, S. and J. Meech, *Trust in design*, in *CHI '00 Extended Abstracts on Human Factors in Computing Systems*. 2000, ACM: The Hague, The Netherlands. p. 45-46.
31. Detroit, A.P.i. *Tesla's new autopilot system lets electric car change lanes by itself*. [Online newspaper] 2015 15.10.2015; Available from: http://www.theguardian.com/technology/2015/oct/15/teslas-new-autopilot-system-lets-electric-car-change-lanes-by-itself.
32. Mathias Sagdahl, S.n.l. *Autonomi*. 27.11.2014; Available from: https://snl.no/autonomi.
33. Olav Skjeggdal, S.n.l. *Automat*. 2015 16.03.2015; Available from: https://snl.no/automat.
34. Parasuraman, R., T.B. Sheridan, and C.D. Wickens, *A model for types and levels of human interaction with automation*. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 2000. **30**(3): p. 286-297.
35. Sheridan, T.B., *Telerobotics, automation, and human supervisory control*. 1992: MIT press.
36. Lindsay, J., *Avoiding environmental fallacy with systems thinking*, in *Progrium*. 2012.
37. Blomqvist, K., *The many faces of trust*. Scandinavian Journal of Management, 1997. **13**(3): p. 271-286.
38. Langewiesche, W., *The Human Factor*, in *Vanity Fair*. 2014.
39. Parasuraman, R. and C.A. Miller, *Trust and etiquette in high-criticality automated systems*.

Communications of the ACM, 2004. **47**(4): p. 51-55.
40. IDEO. *About IDEO*. [cited 2015 23.11.]; Available from: https://www.ideo.com/about/.
41. Design, S.U.I.o. *Welcom to the virtuel crash course in design thinking*. [cited 2015 23.11.]; Available from: http://dschool.stanford.edu/dgift/.
42. Cooper, R., S. Junginger, and T. Lockwood, *Design thinking and design management: A research and practice perspective.* Design Management Review, 2009. **20**(2): p. 46-55.
43. Tim Brown, R.L.M., *Design for Action*, in *Harvard Business Review*. 2015. p. 56-64.