# Towards trusted autonomous vehicles from vulnerable road users perspective

# Towards Trusted Autonomous Vehicles from Vulnerable Road Users Perspective

Khaled Saleh, Mohammed Hossny and Saeid Nahavandi
Institute for Intelligent Systems Research and Innovation (IISRI)
Deakin University, Australia
Email: {kaboufar, mhossny, nahavand}@deakin.edu.au

***Abstract*—A number of recent research projects in human-vehicle interaction field are addressing the problem of human trust in autonomous vehicles. Almost all of these work are focusing on investigating the attributes and the factors that influence the human drivers' trust of these vehicles. However, a little research has been done on the bystander humans' trust of autonomous vehicles. Bystander humans in the context of autonomous vehicles, are humans that does not explicitly interact with the automated vehicle but still affect how the vehicle accomplishes its task by observing or interfering with the actions of the vehicle. Vulnerable road users (VRU) are considered one example of the bystander humans interfering with the autonomous vehicle. According to a recent research study, intent understanding between vulnerable road users and autonomous vehicles was one of the most critical signs that accounted for a trusted interaction between the two entities. In this paper we are proposing a computation framework for modeling trust between vulnerable road users and autonomous vehicles based on a shared intent understanding between the two of them.**

***Index Terms*—Human-on-the-loop, trust modeling, vulnerable road users, autonomous vehicles and intent understanding.**

## I. Introduction

Autonomous Vehicles (AVs) development have got some momentum over the recent years. The first appearance of AVs was in early 2000, during the grand challenges organized by the US DARPA (Defense Advanced Research Projects Agency) [1]–[3]. In 2009, specifically after the success that have been witnessed during the DARPA urban challenge in 2007, Google initiated its self-driving car project by testing it on the freeways in California. Following Google, other car manufacturers started to catch-up such as Toyota, Mercedes Benz, GM, Ford, Audi, and more. Ultimately, AVs have a strong potential in reducing traffic accidents, increasing road capacity and providing critical mobility to the elderly and handicapped [4], [5]. However, one of the most critical barriers against the wide spread of AVs in the coming few years is, the human trust of their capabilities.

According to a recent survey done by the American Automobile Association (AAA), only one-in-five Americans say they would trust an autonomous vehicle to drive itself [6]. Human trust in automation in general, and in autonomous vehicles specifically, is rather complex and multidimensional construct, due to the large number of factors and attributes influence it. Therefore, the issue of human trust in automation has got an increased interest over the past 20 years from a number of research communities such as human factors [7]–[9], cognitive sciences [10]–[12] and human robot interactions (HRI) communities [13], [14].

Since there is a little research done on the human trust of AVs, there is not a unified definition of human trust in the context of AVs so far. Though, most of the technologies involved in developing AVs are actually based on research work done in the robotics field, we found that the trust definition used in Human-Robot interactions (HRI) frameworks in [14] and [15] would be more suitable for our proposed framework. In fact, the human-robot trust definition used in the aforementioned HRI frameworks is not entirely different from the human-automation trust definition introduced by Lee and See [16]. Lee and See defined human-automation trust as "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability.

According to the previous definition when it is applied in HRI frameworks specially the mobile robots, the human involved in the interaction loop with the mobile robot is the individual, and the agent is the mobile robot that is trying to accomplish a certain task determined by goals set by another human. Since these goals are highly dependent on the role of the human involved in the HRI loop, Schlotz [17] defined the five different roles that the humans can assume in any HRI framework, as the following:

- Human can become a supervisor who monitors the mobile robot and can intervene once it is a necessity;
- Human can act as an operator who is closely in a direct interaction with the mobile robot manipulating the robot's actions;
- Human can be a programmer or a mechanic who have an access to modify the software or the hardware capabilities of the mobile robot;
- Human can be a team-mate or a peer to the mobile robot working together to achieve a mutual goal; and
- Human can be a bystander who is not directly part of the HRI loop but can affect how a mobile robot is accomplishing a certain task.

Similarly, the previous formulation of human-robot trust can be extended to the special case of human-AVs trust, however instead of the agent being the mobile robot, in this case it will be the AV and instead of the five roles that human can assume in HRI frameworks, it will be only restricted to three roles in human-AVs trust frameworks. Whereas, the human can assume one of the following three roles:

- Driver/Passenger of the host AV (depending on the level of autonomy of the AV);
- Vulnerable road user (VRU) sharing the road with the AV such as pedestrians and cyclists; or
- Driver of other vehicles sharing the same road with the AV.

Recently, a number of research projects started to consider the issue of human trust in AVs. However, they were focusing mainly on the driver/passenger of the host AV [18]–[20]. There is a little research done so far on the issue of trust between the AV and the VRU or the drivers sharing the road with the AV. Since VRU are considered the weakest road users due to their vulnerability to traffic accidents, we believe they deserve a special interest over the other drivers sharing the road with the AVs.

In this paper we will be focusing on VRUs' trust of AVs. The rest of this paper is organized as follows, a review of the existing quantitative models developed for human-AVs trust in general will be first discussed, then we will draw from these models the specific factors that are applicable for the VRUs-AVs trust case. Finally, we will present our proposed computation framework for modeling the trust between VRUs and AVs based on these factors.

## II. Human-AVs Interaction Loop

One of the key elements in establishing a trusted interaction loop between humans and automated systems, begin with the humans' assessment of the trustworthiness of these automated systems [16]. The research in [21], [22], outlined that the characteristics of an automated system are among the major factors that impact the trustworthiness of any automated systems. Furthermore, they indicated that characteristics of an automated system can be described in terms of the level of autonomy of this system and how well its function is perceived by humans. Since the definition of automated systems tend to vary across different sectors and industries. Parasurman et al. [23] came up with a taxonomy for the definition of automated systems based on a ten levels of decision making and/or control done by the human or the machine in these systems.

Similarly, in the arena of AVs, and in a reaction to the uprise of the self-driving cars development, another taxonomies of autonomy were proposed for AVs. One of the taxonomies was proposed in early 2013 by the USA's National Highway Transportation Safety Administration (NHTSA). NHTSA provided a preliminary policy statement in order to regulate the development and testing of AVs across the US, which had a taxonomy of the five levels of autonomy of vehicles as the following:

- **Level 0**: No automation
- **Level 1**: Function-specific automation
- **Level 2**: Combined function automation
- **Level 3**: Limited self-driving automation
- **Level 4**: Full self-driving automation

Table I, presents a comparison between NHTSA's taxonomy of levels of autonomy with respect to the human-vehicle decision-making and the human-vehicle interaction loop. Human-vehicle decision-making loop is concerned with the actual responsibilities that the human driver can perform in each level of the autonomy.

On the other hand, Human-vehicle interaction loop is more of a higher abstraction level of the degree of the involvement that the human driver can take within the vehicle. Whereas, the human driver can be totally engaged and in control of the vehicle most of the time (levels 0 and 1), and we refer to the driver here as "Human-in-the-loop" as it is specified in human-machine collaboration research community [24]. Alternatively, the human driver can have an intermittent control of the vehicle and the automated system of the vehicle is responsible for the decision-making most of the time (levels 2 and 3), and we refer to here as "Human-on-the-loop" or "Human supervisory control" as they call it in process control research community [25]. Lastly, the human driver can be just a passenger in the vehicle and has no control of the actual driving task of the vehicle (level 4) and it is commonly referred to the driver here as "Human-out-of-the-loop" [26].

From this taxonomy we can draw out that it is completely cantered around the human driver, and does not take into account any other human road users who are sharing the road with the vehicle. Specifically, when it comes to VRUs, they did not have so much attention in human-AV interaction frameworks in contrast to the bystanders in HRI frameworks. Given that, most of the interactions between VRUs and vehicles right now are based on an implicit social rules with the human driver inside the vehicle. Thus, we are arguing that VRUs will be considered as the new Human-on-the-loop when vehicles are operating in the upper level of autonomy (level 4) of NHTSA's taxonomy.

The motive for our believe that VRUs will be acting as Human-on-the-loop when human drivers will be out of the loop (level 4), is that the degree of influence VRUs will impose on the decision making loop of AVs will be much higher than the one of the human passenger inside the AV. For instance, when AV encounter a pedestrian trying to cross the road from a non-crosswalk stop, the AV will change its motion planning to avoid collision with the pedestrian which in returns impact the decision-making loop of the AV.

TABLE I. NHTSA Level of Autonomy Taxonomy.

| NHTSA Levels of Autonomy | Human-Vehicle Decision-making Loop | Human-Vehicle Interaction Loop |
|---|---|---|
| Level 0 | Driver is in full charge of the vehicle. Includes warning-only systems (like Forward Collision Warning, Lane Departure Warning) | Human-in-the-loop |
| Level 1 | Driver is also in full control. Only minor control functions automated (Adaptive Cruise Control, Electronic Stability Control) | Human-in-the-loop |
| Level 2 | At least two primary control functions are automated, driver responsible for monitoring safe operation and is available for control on short notice | Human-on-the-loop |
| Level 3 | Driver cedes full control to automation under certain conditions, driver is available for occasional control, but does not have to constantly monitor safe operation | Human-on-the-loop |
| Level 4 | Driver provides destination or navigation support, but is not expected to be available for control at any time during the trip | Human-out-of-the-loop |

Additionally, since AVs will be prioritizing the safety of its passengers over VRUs [27], so it will need to model the human factors and social behaviors of VRUs in order not to risk that objective. Which in fact is similar to the case of level 3, where the vehicle will have to model the human factors and behaviors of its human driver in order to cede control to him in a safer manner [28], [29].

Unlike human drivers, VRUs are not governed or constrained by the existing regulations of traffic, which that makes their behaviors unpredictable most of the time [30]. Thus, AVs need to have a deeper understanding of VRUs' behaviors and motives in order to increase their trust of AVs. Because by doing so, it will help in achieving a safer cooperative interaction between the two of them [31].

## III. Objective Measures for Bystanders and VRUs' Trust

A number of the factors that affect the bystanders' trust in autonomous systems have been studied extensively in the HRI field over the past 10 years. However, just a few have addressed this issue in the context of AVs "i.e, VRUs". One of the most commonly adopted methods in assessing human trust in HRI frameworks, is the objective measures [32]. Objective measures can be viewed as a quantitative metric of a behavioral data produced unconsciously by humans. In the following, we will review a number of the work that have been done on the objective measures of trust for the bystanders in HRI field in general and in its specific case of human-AV interaction.

In HRI, Tsui et al. [33], investigated the level of trust the bystanders have of a mobile robot in a specific scenario of a corridor passing based on social behavior cues. In their experiments, they found out that the bystanders tend to have a higher levels of trust in the mobile robot when it accommodates some degree of adhering to social protocols. Social protocols can be such as complex as behaviors they would expect from other humans like yielding for them when in cases the corridor can only fit for one entity to pass "i.e, either human or robot". Another work by DeSteno et al. [34] investigated a number of behavioral cues that accounted for a trusted interaction between human and a tele-operated robot. They relied on facial verbal cues done by a bystander human to identify a set of non-verbal cues, that when performed in a specific sequence, they can be used as an indicator of a trustworthy or untrustworthy behavior perception of the robot.

In [35], Boerkoel et al. utilized a number of social cues between bystanders and a mobile robot deployed in an industrial manufacturing environment for a trustworthy interaction between the two of them. Whereas, the mobile robot used to infer the intentions of the bystander human (such as another human coworker in a factory) by recognizing the coworker's activities as well as feed-back him with its intended route that it will take next while it is navigating.

In the context of AVs, Florentine et al. [36] proposed a method to increase the VRUs' trust in autonomous golf-cart by incorporating a human-like social behaviors. They substituted the human eye contact that happens between human-driven vehicles and pedestrians by a Light-Emitting Diode (LED) strips as shown in Fig. 1. The LED strips were used to convey a perception information to nearby pedestrians passing by the autonomous golf-cart. They placed a number of the LED strips around the golf-cart's front side to notify pedestrians that the golf-cart can see them. Whenever a pedestrian is in a close proximity to the golf-cart, the LED strips will change its colors, and as the pedestrian walks around it, the LED lights will start following him.

In [37], Matthews et al. connected the intent understanding as an indicator for a trusted interaction between the AVs and VRUs. Similar to [36], they also used an autonomous golf-cart provided with LED strips as well as LED word display for communicating its intent for the pedestrians. They used the LED word display for explicitly communicating the intent of the golf-cart using pictures, words, or a combination of the two. For identifying the pedestrians' intent they surveyed a number of people about the most probable actions they


Fig. 1. The autonomous golf cart equipped with LED strip [36]. The blue color indicates no obstacle within close range, whereas red shows presence of a nearby obstacle.

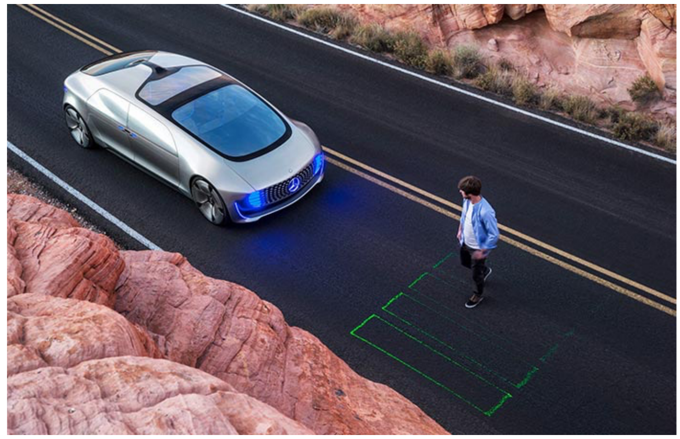

Fig. 2. Concept autonomous vehicle from Mercedes-Benz [31] projecting laser beams on the road as a virtual crosswalk for pedestrians to walk.

would take when encountered with an AV. Then, based on the results of their survey they used a Belief-State Markov Decision Process to recognize a five pedestrians' actions such as pedestrians moving out of the golf-cart's way.

Similarly, in [38], Pennycooke et al. outfitted a prototype vehicle with a number of sensors to resemble an actual AV, to convey a trusted interaction based on intent understanding between the prototype AV and pedestrians in a simulated indoor environment. When the mounted Xbox Kinect on the prototype AV, sense the presence of a nearby pedestrian, an eye-shaped blue LED mounted on the prototype AV flashes to notify the pedestrian that it is ok to cross.

## IV. Modeling the Trust of VRUs

From the previous review, the intent understanding between the AV and the VRUs has been proven to be a really strong objective measure of a trusted interaction between the two entities. Generally speaking, the intent understanding cycle between AVs and VRUs can be viewed as a combination of two steps: 1) recognizing and predicting the intent of the VRU 2) conveying the AVs' acknowledgment of this intent to the VRU through an intent communication interface. Most of the work that has been done so far on the intent understanding between VRUs and AVs, tend to focus mainly on the second part, the intent communication interface.

The methods of recognizing and predicting the VRUs' intent seems to be rather underestimated. Despite the fact of the importance of the work on the intent communication interface for AVs to convey their intent to the VRUs. However it diminishes the task of predicting the intent of VRUs to a simple actions recognition in a contained environment ignoring the uncertainty existed in the human actions. In order for AVs to convey a reliable intent that can be trusted by the VRUs, it needs to have a shared deep intent understanding between it and the VRUs. Whereas, the AV could be able to predict and communicate the intent back and forth between it and the VRU.

### A. *Intent Understanding Framework*

The first and the most critical step for a shared intent understanding between AVs and VRUs for a trusted interaction between them, is the intent capturing of VRUs by the AV [39]. However, VRUs' intent is not such an observable quantity that can be sensed or captured directly by AVs. Historically, the intent is commonly inferred using other observable cues and behaviors done by the VRUs such as gestures (i.e, nodding or waving) [40], body/head pose combined with motion patterns [41], [42].

Since, the behaviors of VRUs are unpredictable and has some degree of uncertainty that can change over small periods of time, the process of capturing their intent are done on a two steps; one for a short-term prediction (less than 1 second), and the other step rely on the first one to predict the intent over a long-term period (up to 10 seconds ahead). Once the intent prediction cycle done, it can be safely communicated to the VRU.

In Fig. 3, we are presenting our proposed framework for intent understanding, whereas we are breaking down the problem into three main modules:

- Short-term Intent Prediction;
- Long-term Intent Prediction; and
- Intent Communication Interface.

The short-term intent prediction component will be concerned with a momentary prediction of VRUs based on a number of behavioral and social cues. The behavior cues could be such as the activity recognition of the VRU. For example, a pedestrian might be walking on the curbside and stop to cross the road, or a cyclist is signaling to take a left-turn. The social cues could be such as the social interactions between VRUs and

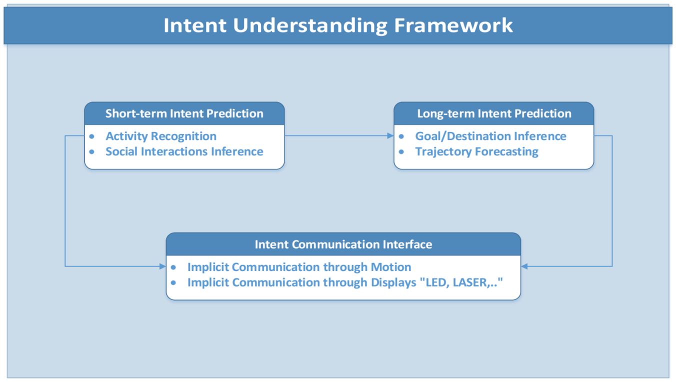


Fig. 3. Intent understanding framework for modeling the trust of VRUs.

each other's. For instance, a pedestrian with his family could be trying to cross the road, the AV needs to have knowledge about that kind of social interactions and provide the proper communication accordingly in order to gain the trust of VRUs.

Conversely, long-term intent prediction will be concerned with prediction of the intent of VRUs over a longer period. The prediction will be based on the prediction output of the short-term module as well as other variables such as the forecasting of his/her trajectory a number of seconds ahead in the future as well as the inference of his/her destination/goal. In scenarios such as the AV has sensed a pedestrian from a distance trying to take on a cross-walk in urban environment where is no traffic lights, and slowed down in order to make him pass the cross-walk on his ease, that would be similar to the action of "after you" done by the human driver in the same situation.

Finally, the intent communication interface, and it will be responsible for giving the implicit communication between VRUs and the AV based on the information that flows to it from the short-term and/or the long-term intent prediction modules. A two common types of implicit communications can be broadcasted from the intent communication interface to the VRU either through the motion of the AV itself or through light-based displays such as LED and LASER projectors. Implicit communication through motion can be such as when the AV try to communicate a hesitation like the human drivers to alert a pedestrian by oscillating its motion between accelerating and breaking. On the other hand, implicit communication through displays can be either using LED strips [36], [37] or laser projectors [31] as shown in Figures 1 and 2, respectively.

### B. Integrated Trust Modeling Framework

In our proposed framework for intent understanding outlined in Fig. 3, the details of the inputs or the outputs that our proposed framework assume and generate were abstracted. However, in order to explicitly indicate them, we need to integrate our framework with one of the system architectures for AVs. Whereas, that will clearly demonstrate how things fit together within these architectures and to make sure that our proposed framework can easily interfaces with them. Thus, we chose the system architecture of the AV, Boss that achieved the first place in the DARPA urban challenge in 2007 to demonstrate that [3]. The system architecture used in Boss is considered one of the standard system architectures used in AVs and almost the other architectures like the ones in [43], [44] are considered just a derivatives of it.

The main software components of Boss's architecture are organized as the following:

- **Perception:** where in this component, all the sensor data coming from the sensors mounted on the vehicle are fused together to form a world model that can provide the vehicle with information such as: the pose of the vehicle, map of the static obstacles around it, map of the road itself, locations of dynamic obstacles such as other vehicles, pedestrians and cyclists with respect to the vehicle.
- **Mission planning:** this component responsible for computing the possible routes that the vehicle can take given the information it takes from the perception component, and prioritize the optimal paths based on the safest shortest time.
- **Behavioral executive:** it executes the plan determined by the mission planner, it can be considered as a finite state machine that converts the mission plan into a temporal-sequence of motion planning goals that are taking the current condition of the road and the driving context into account.
- **Motion planning:** it converts the motion goal generated by the behavioral executive to an actual trajectory that can take the vehicle safely to its goal and sends this trajectory through a steering and acceleration/braking commands to the control system of the vehicle.

In Fig. 4, the integrated framework for trust modeling of VRUs is proposed with emphasis on the input and output signals of the intent understanding framework.

The input signals of the intent understanding sub-framework are: dynamic obstacles, the road map, the static obstacle map, and vehicle intent. Dynamic obstacles signal from the perception component contains the locations of the moving objects around the vehicle in general, whereas inside the short-term intent prediction module it is processed to get the poses, activities and the social interactions of VRUs specifically which are then passed to the long-term intent prediction module as it was discussed previously. The road map and the static obstacle map signals respectively are passed to the long-term intent prediction module, where they are essential for the inference of the destination and the trajectory forecasting of VRUs as discussed in [45].

Internally, the intent communication interface takes the output of the short-term and long-term intent prediction modules

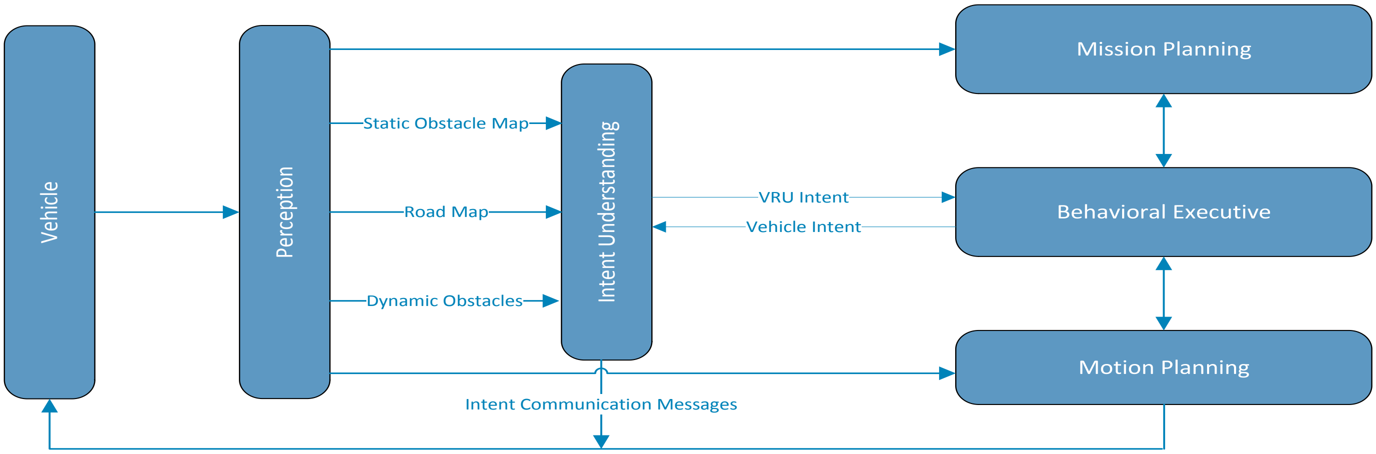


Fig. 4. Integrated framework for modeling trust of VRUs based on the system architecture of Boss [3]. Arrows represents the flow of information between the different components of the framework.

and send them to the behavioral executive component through the VRU intent signal, which in returns reasons about the given situation from the intent understanding. Then, based on that the behavioral executive component convey its situation awareness back to the intent interface and/or update its motion goal through the vehicle intent signal. Based on the vehicle intent signal, the intent communication interface of the intent understanding component decides on which suitable interface such as LED strips and LASER projection would be suitable for conveying the vehicle intent. Afterwards, it sends this specific intent to its corresponding hardware interface attached in the vehicle through the intent communication messages signal. Finally, the previous described cycle is repeated as soon as new observations are available from the perception component regarding VRUs.

## V. Conclusion

In this paper we proposed a framework for AVs to model a shared trusted interaction between them and VRUs based on a strong social cues such as the intent understanding. We also presented a review of the research gaps exits in the area of trust modeling for VRUs in the context of AVs. We also discussed a number of the objective measures used for modeling the VRUs' trust from the literature. Additionally, we presented and discussed an integrated framework of our proposed intent understanding model with one of the most successful system architectures used for AVs development. Whereas we showed how scalable our proposed framework is and how it can easily fit within any type of system architectures for AVs.

## Acknowledgment

This research was fully supported by the Institute for Intelligent Systems Research and Innovation (IISRI) at Deakin University.

## References

[1] C. Urmson, J. Anhalt, M. Clark, T. Galatali, J. P. Gonzalez, J. Gowdy , A. Gutierrez, S. Harbaugh, M. Johnson-Roberson, H. Kato, P. L. Koon, K. Peterson, B. K. Smith, S. Spiker, E. Tryzelaar, and W. R. L. Whittaker, "High speed navigation of unrehearsed terrain: Red team technology for grand challenge 2004," Robotics Institute, Pittsburgh, PA, Tech. Rep. CMU-RI-TR-04-37, 2004.

[2] S. Thrun, M. Montemerlo, H. Dahlkamp, D. Stavens, A. Aron, J. Diebel, P. Fong, J. Gale, M. Halpenny, G. Hoffmann *et al.*, "Stanley: The robot that won the darpa grand challenge," *Journal of field Robotics*, vol. 23, no. 9, pp. 661–692, 2006.

[3] C. Urmson, C. Baker, J. Dolan, P. Rybski, B. Salesky, W. Whittaker, D. Ferguson, and M. Darms, "Autonomous driving in traffic: Boss and the urban challenge," *AI magazine*, vol. 30, no. 2, p. 17, 2009.

[4] G. Silberg, R. Wallace, G. Matuszak, J. Plessers, C. Brower, and D. Subramanian, "Self-driving cars: The next revolution," *White paper, KPMG LLP & Center of Automotive Research*, 2012.

[5] D. J. Fagnant and K. Kockelman, "Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations," *Transportation Research Part A: Policy and Practice*, vol. 77, pp. 167–181, 2015.

[6] AAA, "Three-Quarters of Americans Afraid to Ride in a Self-Driving Vehicle," AAA, Tech. Rep., 2016.

[7] J. Lee and N. Moray, "Trust, control strategies and allocation of function in human-machine systems," *Ergonomics*, vol. 35, no. 10, pp. 1243–1270, 1992.

[8] R. Parasuraman and V. Riley, "Humans and automation: Use, misuse, disuse, abuse," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 39, no. 2, pp. 230–253, 1997.

[9] T. B. Sheridan and R. Parasuraman, "Human-automation interaction," *Reviews of human factors and ergonomics*, vol. 1, no. 1, pp. 89–129, 2005.

[10] C. Fallon, E. Bustamante, K. Ely, and J. Bliss, "Improving user trust with a likelihood alarm display," in *Proceedings of the 1st International Conference on Augmented Cognition, Las Vegas, NV*, 2005.

[11] M. Barnes, R. Parasuraman, and K. Cosenzo, "Adaptive automation for robotic military systems," *Human Factors Issues for Uninhabited Military Vehicle*, pp. 420–440, 2005.

[12] M. R. Endsley, "Direct measurement of situation awareness: Validity and use of sagat," *Situation awareness analysis and measurement*, vol. 10, 2000.

[13] R. Parasuraman, S. Galster, P. Squire, H. Furukawa, and C. Miller, "A flexible delegation-type interface enhances system performance in human supervision of multiple robots: Empirical studies with roboflag," *IEEE Transactions on systems, man, and cybernetics-part A: Systems and Humans*, vol. 35, no. 4, pp. 481–493, 2005.

[14] P. A. Hancock, D. R. Billings, K. E. Oleson, J. Y. Chen, E. De Visser, and R. Parasuraman, "A meta-analysis of factors influencing the development of human-robot trust," DTIC Document, Tech. Rep., 2011.

[15] A. Freedy, E. DeVisser, G. Weltman, and N. Coeyman, "Measurement of trust in human-robot collaboration," in *Collaborative Technologies and Systems. International Symposium on*. IEEE, 2007, pp. 106–114.

[16] J. D. Lee and K. A. See, "Trust in automation: Designing for appropriate reliance," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 46, no. 1, pp. 50–80, 2004.

[17] J. Scholtz, "Theory and evaluation of human robot interactions," in *System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on*. IEEE, 2003.

[18] F. M. Verberne, J. Ham, and C. J. Midden, "Trust in smart systems sharing driving goals and giving information to increase trustworthiness and acceptability of smart systems in cars," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 54, no. 5, pp. 799–810, 2012.

[19] C. Gold, M. Körber, C. Hohenberger, D. Lechner, and K. Bengler, "Trust in automation–before and after the experience of take-over scenarios in a highly automated vehicle," *Procedia Manufacturing*, vol. 3, pp. 3025–3032, 2015.

[20] N. Trübswetter and K. Bengler, "Why should i use adas? advanced driver assistance system and the elderly: Knowledge, experience and usage barriers," in *Proc. 7th Int. Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, 2013, pp. 495–501.

[21] B. M. Muir and N. Moray, "Trust in automation. part ii. experimental studies of trust and human intervention in a process control simulation," *Ergonomics*, vol. 39, no. 3, pp. 429–460, 1996.

[22] R. Parasuraman and C. A. Miller, "Trust and etiquette in high-criticality automated systems," *Communications of the ACM*, vol. 47, no. 4, pp. 51–55, 2004.

[23] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "A model for types and levels of human interaction with automation," *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans*, vol. 30, no. 3, pp. 286–297, 2000.

[24] M. R. Endsley, "Automation and situation awareness," *Automation and human performance: Theory and applications*, pp. 163–181, 1996.

[25] M. Cummings, "Supervising automation: humans on the loop," pp. 7–11, 2007.

[26] D. B. Kaber and M. R. Endsley, "Out-of-the-loop performance problems and the use of intermediate levels of automation for improved control system functioning and safety," *Process Safety Progress*, vol. 16, no. 3, pp. 126–131, 1997.

[27] D. M. Morris, "Mercedes-benz's self-driving cars will choose passengers lives over bystanders," October 2016. [Online].

[28] N. Merat, A. H. Jamson, F. C. Lai, M. Daly, and O. M. Carsten, "Transition to manual: Driver behaviour when resuming control from a highly automated vehicle," *Transportation research part F: traffic psychology and behaviour*, vol. 27, pp. 274–282, 2014.

[29] A. H. Jamson, N. Merat, O. M. Carsten, and F. C. Lai, "Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions," *Transportation research part C: emerging technologies*, vol. 30, pp. 116–125, 2013.

[30] B. C. de Lavalette, C. Tijus, S. Poitrenaud, C. Leproux, J. Bergeron, and J.-P. Thouez, "Pedestrian crossing decision-making: A situational and behavioral approach," *Safety science*, vol. 47, no. 9, pp. 1248–1253, 2009.

[31] M. Mara and C. Lindinger, "Talking to the robocar - new research approaches to the interaction between human beings and mobility machines in the city of the future," pp. 86–91, 2015.

[32] I. Gaudiello, E. Zibetti, S. Lefort, M. Chetouani, and S. Ivaldi, "Trust as indicator of robot functional and social acceptance. an experimental study on user conformation to icub answers," *Computers in Human Behavior*, vol. 61, pp. 633–655, 2016.

[33] K. M. Tsui, M. Desai, and H. A. Yanco, "Considering the bystander's perspective for indirect human-robot interaction," in *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*. IEEE Press, 2010, pp. 129–130.

[34] D. DeSteno, C. Breazeal, R. H. Frank, D. Pizarro, J. Baumann, L. Dickens, and J. J. Lee, "Detecting the trustworthiness of novel partners in economic exchange," *Psychological science*, 2012.

[35] J. C. Boerkoel Jr and J. A. Shah, "Planning for flexible human-robot co-navigation in dynamic manufacturing environments," in *International Conference on Human-Robot Interaction (HRI), Proceedings of Human Robot Interaction Pioneers*, 2013.

[36] E. Florentine, H. Andersen, M. A. Ang, S. D. Pendleton, G. M. J. Fu, and M. H. Ang, "Self-driving vehicle acknowledgement of pedestrian presence conveyed via light-emitting diodes," in *Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), International Conference on*. IEEE, 2015, pp. 1–6.

[37] M. Matthews and G. V. Chowdhary, "Intent communication between autonomous vehicles and pedestrians."

[38] N. Pennycooke, "Aevita: designing biomimetic vehicle-to-pedestrian communication protocols for autonomously operating & parking on-road electric vehicles," Ph.D. dissertation, Massachusetts Institute of Technology, 2012.

[39] L. Müller, M. Risto, and C. Emmenegger, "The social behavior of autonomous vehicles," in *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. ACM, 2016, pp. 686–689.

[40] T. Lagström and V. M. Lundgren, "Avip-autonomous vehicles interaction with pedestrians."

[41] C. G. Keller and D. M. Gavrila, "Will the pedestrian cross? a study on pedestrian path prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 2, pp. 494–506, 2014.

[42] J. F. P. Kooij, N. Schneider, F. Flohr, and D. M. Gavrila, "Context-based pedestrian path prediction," in *European Conference on Computer Vision*. Springer, 2014, pp. 618–633.

[43] C. Berger and B. Rumpe, "Autonomous driving-5 years after the urban challenge: The anticipatory vehicle as a cyber-physical system," *arXiv preprint arXiv:1409.0413*, 2014.

[44] S. Behere and M. Trngren, "A functional reference architecture for autonomous driving," *Information and Software Technology*, vol. 73, pp. 136 – 150, 2016.

[45] B. D. Ziebart, N. Ratliff, G. Gallagher, C. Mertz, K. Peterson, J. A. Bagnell, M. Hebert, A. K. Dey, and S. Srinivasa, "Planning-based prediction for pedestrians," *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009*, pp. 3931–3936, 2009.