International Journal of Human–Computer Interaction

ISSN: 1044-7318 (Print) 1532-7590 (Online) Journal homepage: https://www.tandfonline.com/loi/hihc20

Human–Vehicle Cooperation in Automated Driving: A Multidisciplinary Review and Appraisal

Francesco Biondi, Ignacio Alvarez & Kyeong-Ah Jeong

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
https://doi.org/10.1080/10447318.2018.1561792

Taylor & Francis
Taylor & Francis Group

Check for updates

Human–Vehicle Cooperation in Automated Driving: A Multidisciplinary Review and Appraisal

Francesco Biondi[a,b], Ignacio Alvarez[c], and Kyeong-Ah Jeong[d]

[a]University of Windsor, Windsor, Ontario, Canada; [b]University of Utah, Salt Lake City, Utah, USA; [c]Intel Labs, Portland, Ontario, USA; [d]Intel Corp., Portland, Ontario, USA

ABSTRACT

To draw a comprehensive and cohesive understanding of human–vehicle cooperation in automated driving, a review is made on key studies in human–robot interaction and human factors. Throughout this article, insight is provided into how human drivers and vehicle systems interplay and influence each other. The limitations of technology-centered taxonomies of automation are discussed and the benefits of accounting for human agents are examined. The contributions of machine learning to automated driving and how critical models in human-system cooperation can inform the design of a more symbiotic relationship between driver and vehicle are investigated. Challenges in the human element to enable the safe introduction of road automation are also discussed. Particularly, the unintended consequences of vehicle automation on driver's workload, situation awareness and trust are examined, and the social interactions between driver, vehicle, and other road users are investigated. This review will help professionals shape future directions for safer and more efficient and effective human–vehicle cooperation.

1. Introduction

Vehicle automation holds the promise of improved transportation efficiency and increased traffic safety. Indeed, recent estimates suggest that adopting assistance systems and automated vehicle technology would reduce the number of road crashes and fatalities by 25–90% (European Commisssion, 2016; Litman, 2017). However, studies in user automation as well as recent investigations conducted on road accidents involving semi-automated systems indicate the potential for automation to escalate, rather than mitigate, the effect of human error on safety (Biondi et al., 2018, under review; Endsley, 2017; National Transportation Safety Board, 2017, 2018).

The research community is warning about the unintended consequences that drastic shifts in human–vehicle interaction will have on safety, efficiency, and acceptance of this automated driving technology (Noy, Shinar, & Horrey, 2018). At the same time, society is at a technological inflexion point where advances in artificial intelligence and vehicle automation hold the promise for a near-future decline in road accidents. We believe the time is right for revising our understanding of human–system cooperation in the context of automated driving.

The rest of the article is structured as following. First, we focus on the vehicle element involved in automated driving. We examine the implications that adopting a human-centered taxonomy (rather than a technology-centered classification) of vehicle automation has on establishing a more cooperative interaction between the user and the vehicle. In discussing human–vehicle team dynamics, we borrow from the literature in human–robot interaction to analyze how conceptual models in this field can help understanding user–vehicle interactions in multi-agent scenarios (more users, more vehicles). Lastly, we focus on the contribution that machine learning and training can have in crafting a more cooperative and less hierarchical relationship between the human and vehicle agents.

In the second part of this review, we look at the human element in automated driving. We focus on the main challenges that transportation researchers need to address to enable a safe, accident-free introduction of vehicle automation. We borrow from the literature in other fields of transportation, to discuss the implications that interacting with automation has on the driver state and workload. We also discuss the research examining the detrimental role that *clumsy automation* has on the driver monitoring of the vehicle systems, and the resulting reduction in awareness and arousal. In the last portion of this section, we also examine the social qualities of human–vehicle interaction, and how social automated systems will need to compensate for the now absent human-like communication among road users.

2. The vehicle element in human–vehicle cooperation

Automation in driving has expanded from aiding and easing driver's vehicle control to completely taking over control tasks under certain conditions. In the last decade, in the attempt to develop a taxonomy of vehicle automation, governmental and

CONTACT Ignacio Alvarez ignacio.j.alvarez@intel.com Intel Labs, Portland, Oregon, USA.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hihc.

industry organizations like the National Highway Traffic Safety Administration (NHTSA), the Society for Automotive Engineers (SAE), and the German Federal Highway Research Institute have put forward their respective proposals to standardize different levels of automated driving systems. After years of debating, the community reached a consensus in 2017, when, with a formal endorsement by NHTSA, it agreed on the use of the SAE classification as the official taxonomy on automated driving systems (NHTSA, 2017).

2.1. SAE taxonomy of automated driving systems

According to SAE J3014 (SAE, 2018), there are six levels of automated driving systems, from level-0 (or fully manual) to level-5 (or fully automated) (see Figure 1). Level-1 requires a substantial amount of human control, with the human operator controlling the majority of vehicle operations while being assisted by systems (Adaptive Cruise Control or Lake Keeping Assist) controlling single operations (steering or accelerating). In levels 2 and 3, also referred to as partial automation or semi-automated driving, advanced driver assistance systems (ADAS; see, Biondi et al., 2018) take charge of multiple driving operations (steering + accelerating), while the human driver is still responsible for monitoring the functioning of the vehicle or taking back manual control when required. Levels 4 and 5 are defined as highly and fully automated driving systems, respectively. They require minimal or no input from the human driver.

Despite the considerable amount of effort gone into developing the SAE taxonomy and accepting it as a standard, the transportation community has highlighted a number of limitations. Templeton (2015), for instance, noted that any numerical classification of levels of automated systems suggests a hierarchy or ordering to technology development, which, most likely, will not evolve as expected (also, see Liu, 2016). In support of his argument, he mentions the example of the French company Navya, which has already commercialized its version of a fully automated shuttle for low-speed public transportation, while, at the time this manuscript was written, no fully developed level-3 automated system has yet been introduced to the market. Criticism on the numerical classification on levels of autonomy also comes from the Defense Science Board (2012). It suggests that, while such taxonomies attempt to group functions needed for generalized scenarios, they put too much emphasis on what the computer can do, rather than on the collaboration between the computer and the operator. In their critique to the SAE classification for level-3 driving, Inagaki and Sheridan (2018) also suggest that defining rigid levels of autonomy does not inform the design process needed to engineer efficient transitions of control between the human and the automated system, and vice versa. Lastly, these taxonomies imply that there exist well-defined levels of "system intelligence" that can reliably control the vehicle at a specific level for the entire duration of the drive (Department of Defense, 2012), which, as investigations on recent accidents involving level-2 vehicles have clearly shown (NTSB, 2017, 2018), is far from being true.

2.2. A necessary human-centered perspective on automation

One of the main criticisms advanced by the user automation research community toward technology-centered taxonomies is their limited consideration for the human element involved in automation. For semi-automated driving systems, for instance, the human driver is required to monitor the functioning of the automated system (level-2), and promptly regain control of the vehicle (level-2 and level-3), regardless of the specific circumstances, road or traffic scenarios. In other words, the human driver is expected to act as instructed, regardless of the possibility of them being distracted, fatigued, or, in any way, becoming disengaged from the monitoring task. An alternative, human-centered perspective on user interaction with automation is provided in the seminal study by Sheridan and Verplank (1978). Their model defines 10 distinct levels of automation, whereby the human and the system take charge of different functions. At level-1, the system "offers no assistance: human must take all decision and actions," while, at level-10, the system "decides everything, acts autonomously, ignoring the human" (Sheridan & Verplank, 1978). For level-5 or lower, the human retains authority over the final decision. Vice versa, for level-6 or higher, the system would automatically execute its own decision, unless the human intervenes.

This model was later revised by Parasuraman, Sheridan and Wickens in 2000, to account for different stages of information processing at which the sharing of responsibilities between the human and the system takes place. While the original ten-level model focused on response selection and execution, it did not account for preliminary stages where the information is being acquired and integrated. A total of four stages of processing are described by Parasuraman, Sheridan, and Wickens (2000): information acquisition, information analysis, decision selection, action implementation. The four stages were developed upon the phases of perception, response selection, and response execution described in models of human information processing (Kahneman, 1973; Pashler, 1994; Tombu & Jolicoeur, 2003; Wickens, 2008).

These four stages can be automated, on a control continuum from low automation to high automation (as in Sheridan & Verplank, 1978). Information acquisition applies to sensing of information and registration of environmental inputs. At a low level of automation, this may entail having sensors that scan the road environment. A highly automated sensing system, instead, may classify road objects depending on their characteristics and dimensions (pedestrians, bicyclists, passenger vehicles, trucks, etc.). System information analysis refers to cognitive functions such as working memory and inferential processing (Parasuraman et al., 2000). While a low-automated system may provide information about the trajectory of road objects, a system with a higher level of automation in this domain may inform the human operator about the potential for a collision between the vehicle and other road agents. Decision selection is the ability to make decisions and select actions. Within the realm of vehicle automation, this describes systems recommending the human driver to execute certain

SAE level	Name	Definition
0	No Automation / Fully-Manual	The human driver is in charge of all aspects of driving.
1	Driver Assistance	The human driver is assisted in the driving task by a driver assistance system that controls either steering or acceleration/deceleration. The driver stays in charge of monitoring the driving task.
2	Semi-automated	Driver assistance systems are in charge of steering and acceleration/deceleration and the human driver is expected to monitor the driving task and respond to a request by the system to intervene.
3		The automated driving system is responsible of driving and monitoring the driving environment. The human driver in expected to respond appropriately to a request to intervene.
4	Highly-Automated	The automated driving system is in charge of the driving task, monitoring the driving environment, even if a human driver does not respond appropriately to a request to intervene.
5	Fully-Automated	The automated system is in charge of all aspects of driving under all roadway and environmental conditions that can be managed by a human driver

Figure 1. SAE levels of vehicle automation adapted from SAE (2018).

safety-relevant maneuvers ("BRAKE" when the following distance is shorter than a safety threshold). Action implementation refers to the ability to execute specific actions that may or may not have been selected with the assistance of the automated systems. Cruise Control represents an example of a system with low automation, while Adaptive Cruise Control, with its ability to automate both distance keeping and speed maintenance, is a system with a higher level of automation.

One advantage of using a more human-centered design approach to automation is its ability to inform the design of the automated system, and its cooperation with the human. According to this model, the first step in the design process is identifying what level of automation applies to the system at each of the four stages of information processing. The next step is to evaluate the effect that each automation level may have on human-related elements like: workload, trust, situation awareness, and skill degradation (a more comprehensive discussion of these Human Factors aspects is provided in Section 3 of the manuscript). To highlight the importance of the human element in automation, Parasuraman et al. define these as Primary Evaluation Criteria. The authors also describe Secondary Evaluation Criteria, which include: automation reliability and costs of decision/action outcomes.

With this approach in mind, it is possible to understand the root causes of recent accidents involving automated vehicles, and relative design oversights. The investigation conducted by the National Transportation Safety Board on the fatal crash involving a Tesla Model S in 2016, suggests that "contributing to the car driver's overreliance on the vehicle automation was its operational design, which permitted his prolonged disengagement from the driving task and his use of the automation in ways inconsistent with guidance and warnings from the manufacturer" (NTSB, 2017). Within the theoretical framework offered by Parasuraman et al. (2000), this is an example of *clumsy automation* at the action execution stage of processing. Prior to the crash, the system retained full control of vehicle operation, yet it did not reduce the vehicle's speed nor the assisted emergency brake activated. With the human driver relying on the system capabilities, this facilitated the driver disengagement over time, impairing his awareness of the surrounding traffic conditions. A more in-depth discussion of human-related aspects associated with vehicle automation is presented in a later section of the manuscript.

2.3. Human–vehicle team dynamics

The human–vehicle interaction models discussed so far (also see Goodrich & Schultz, 2018) entail a one-to-one relationship between one vehicle and one operator. Research in human–robot interaction (HRI) have explored more flexible situations in which a single human must coordinate with multiple automated robots (grounded or aerial), and multiple humans share control of one or multiple robots.

Within the transportation domain, one example of shared control between humans and automated vehicles is truck platooning. Truck platooning is the linking of two or more trucks in convoy. The vehicles maintain a set, close distance between each other when they are connected for certain parts of the journey (ACEA, 2017). While, in its initial iteration, the lead vehicle will be in charge of setting the speed for the entire convoy with other drivers required to monitor their respective trucks, it is envisioned that, in later stages, the lead driver only will monitor the functioning of the automated system, giving other drivers the option to fully disengage from the driving task.

Even inside the vehicle, given the dynamic environment of an automated system, researchers should not expect static roles among its passengers, such as they exist today. Instead, they should plan for efficient transitions in the team structure affected with changes in responsibility, authority, and roles. These dynamic changes already take place in transportation-as-a-service use cases, in particular, shared drives, where multiple passengers are involved in route negotiation. These dynamic negotiations where no passenger is in direct control of the vehicle bring the issue of how could the automated vehicle resolve conflicts between passengers. Scholtz (2003) warned that peer human-automation interactions, such as those derived from natural language interactions with an automated vehicle, might not have time to solve conflicts, and must swiftly retract to supervisory relations. In automated driving, this could be applied to last mile navigation, where one of the passengers might have to take a lead role to direct the vehicle to an appropriate parking location. Another, more collaborative approach to resolve dynamic control scenarios in automation is "sliding autonomy" in which any of the passengers can actively decide to transfer, or share control with other passenger or with the automated vehicle. Sellner, Heger, and Hiatt (2006) indicated that sliding autonomy was very effective for dynamic task completion. We foresee automated vehicle passengers will require peer-to-peer collaboration with their vehicles to have them adapt to their preferences in route navigation, driving style, safety, and guided maneuvers. Dias, Kannan, and Browning (2008) highlighted that, in order to achieve successful peer-to-peer coordination, automated systems must have the ability to request help, maintain coordination, establish user situational awareness, enable interactions at different levels of granularity, and learn from those interactions.

2.4. Learning, training, and adaptation in human–vehicle interactions

Machine learning has been applied to autonomous vehicles as a solution to multiple tasks, including low-level environment perception tasks, such as signal detection or pedestrian recognition, and high-level cognitive processes, such as path planning or conversational dialog managers in in-vehicle infotainment systems. The goal of peer-to-peer coordination between automated systems and vehicle occupants requires the successful application of artificial intelligence learning techniques, such as supervised learning, imitation learning, and reinforcement learning. Supervised Learning makes use of labeled data to train a system to accomplish a particular task, the labeling being traditionally done via human annotation. This method has been used to estimate the cognitive level of the human operator across time (Mahi, Atkins, & Crick, 2017). This capability allows the vehicle to detect the emergence of cognitive stress in drivers, increasing their level of autonomy, and reducing demands on the driver's attention. Supervised learning has also been used to teach a robot social affordances in daily human interactions through labeled videos, and transfer the knowledge to human–robot interactions in unseen scenarios (Shu, Gao, Ryoo, & Zhu, 2017). This could enable vehicles to learn user interactions preferences and social behavior from outside the vehicle that could be applied to in-cabin scenarios.

Imitation Learning enables an automated system to learn to perform a particular task from examples or demonstrations. The demonstrations are typically carried by a person and then encoded in a sequence of state-action pairs that the system uses as dataset to derive a policy that reproduces the demonstrated behavior. This method has been used to enable human–computer interactions (HCIs) with "teach and show" interfaces, that allow users to train complex behaviors in robots (Saunders, Syrdal, & Koay, 2016). Applied to automated driving, this approach can be used to personalize automated driving behaviors to maximize user comfort. Reducing the complexity of custom driving behavior generation makes it more approachable to non-technical users. However, there are limitations on the application of Imitation Learning techniques. The automated system performance on the driving task will be intrinsically limited to the performance of the teacher, and the system might be unable to execute undemonstrated driving tasks (Argall, Chernova, Veloso, & Browning, 2009). These drawbacks limit the generalization of the taught behaviors, and might put a heavy burden on the teacher to monitor the vehicle. A possible solution to these Imitation Learning shortcomings is the application of constant feedback from the teacher (vehicle user), or to allow self-exploration in the shape of reinforcement learning.

Reinforcement Learning enables an autonomous system to independently discover optimal behaviors through trial-and-error interactions with its environment. To apply this technique, the designer/engineer typically selects a state model or representation for the task, and designs a reward function that provides performance feedback as a scalar number. Reinforcement Learning applications to human–system interactions have been successfully explored in the form of behavior adaptation mechanisms. These could read small discomfort signals in the human partner, and, in turn, adjust interaction parameters in a robot, such as gaze meeting, motion speed, and timing (Mitsunaga, Smith, Kanda, & Ishiguro, 2006). The same technique can be applied to observe the vehicle passengers discomfort reactions to driving behaviors, and optimize driving behavior for comfort. This method is also useful to guide system learning on conversations between passengers and their autonomous vehicle assistant (Hinds et al., 2004). High-level user behaviors can be mapped to low-level communicative expressions as demonstrated by Hemminghaus and Kopp (2017).

While exploratory results highlight reinforcement learning as a very promising field to autonomous driving, we note that applications of HCI and, more specifically, human–vehicle interaction are still in a nascent phase. One issue that autonomous vehicle designers experience is the difficulty in selecting appropriate behavior representations for the interactions. Researchers must also manually craft reward functions that balance short-term rewards versus longer term rewards (Kober, Bagnell, & Peters, 2013). Another drawback is that field testing of Reinforcement Learning methods in automated driving must be trained first in safe environments, such as realistic driving simulators, to allow behavior exploration without endangering passengers (the Car Learning to Act driving simulator or CARLA shows particular promise for these tasks; Dosovitskiy, Ros, Codevilla, & López, 2017).

2.5. Cooperative models for human–system interaction in automated driving

The above described development of artificial intelligence methods applied to automated driving will enable interactions where human passengers and automated vehicles are able to cooperate as peers in the driving task. But, as we learned, both automated system and humans must learn how to cooperate safely and efficiently in the dynamic environment of a moving vehicle, and thus, the need to create models that define the nature and laws governing human and system interactions in driving conditions.

One of the most popular models of vehicle automation interaction in the literature is the H-Metaphor (Flemisch et al., 2003; see Figure 2). This model takes inspiration from the coordinated interactions between a horse and its rider, and translates them to automated driving systems. The automated vehicle is capable of operating in general conditions avoiding obstacles and other vehicles. The human driver, in a supervisory role, can perform other activities as long as they share a physical monitoring interface through which the driver is constantly aware of what the automated vehicle is doing, and is alerted when the vehicle senses danger, or is unsure about where to go. The vehicle might also be aware of how engaged the driver is, and adjust its interaction behavior accordingly.

Rather than an operational framework for the development of cooperative human–system interactions, this model presents a simplification of the point of view of the cooperation. The H-metaphor also makes a series of assumptions on the automated system intelligence and the user mental model, such as that the vehicle must have predictable, situation-appropriate, and comprehensive behaviors that allow the operator to reliably divert attention elsewhere. The application of artificial intelligence is showing promise to develop these behaviors but as we learned generalization to different scenarios is a challenging task.

A more useful model under today's automation restrictions was proposed by Hoc and Blosseville in a theory that categorizes cooperation in automated driving according to the type of cooperation activities. They define four cooperation modes: perception, mutual control, function delegation, and fully automatic (Hoc & Blosseville, 2003). These modes of cooperation have been studied for automatic cruise control (ACC) and lateral safety control functions, such as Lane Keeping Assist (LKA). The findings suggest that cooperation at the perception level, such as in current ACC and LKA systems (SAE Levels 1 and 2), tends to act at a symbolic level (i.e., warning messages are provided with no action being taken). Similarly, in fully automatic cooperation modes (SAE Levels 4 and 5), interactions might take place mainly at the symbolic level, with the system providing sporadic updates to the environmental factors or control decisions.

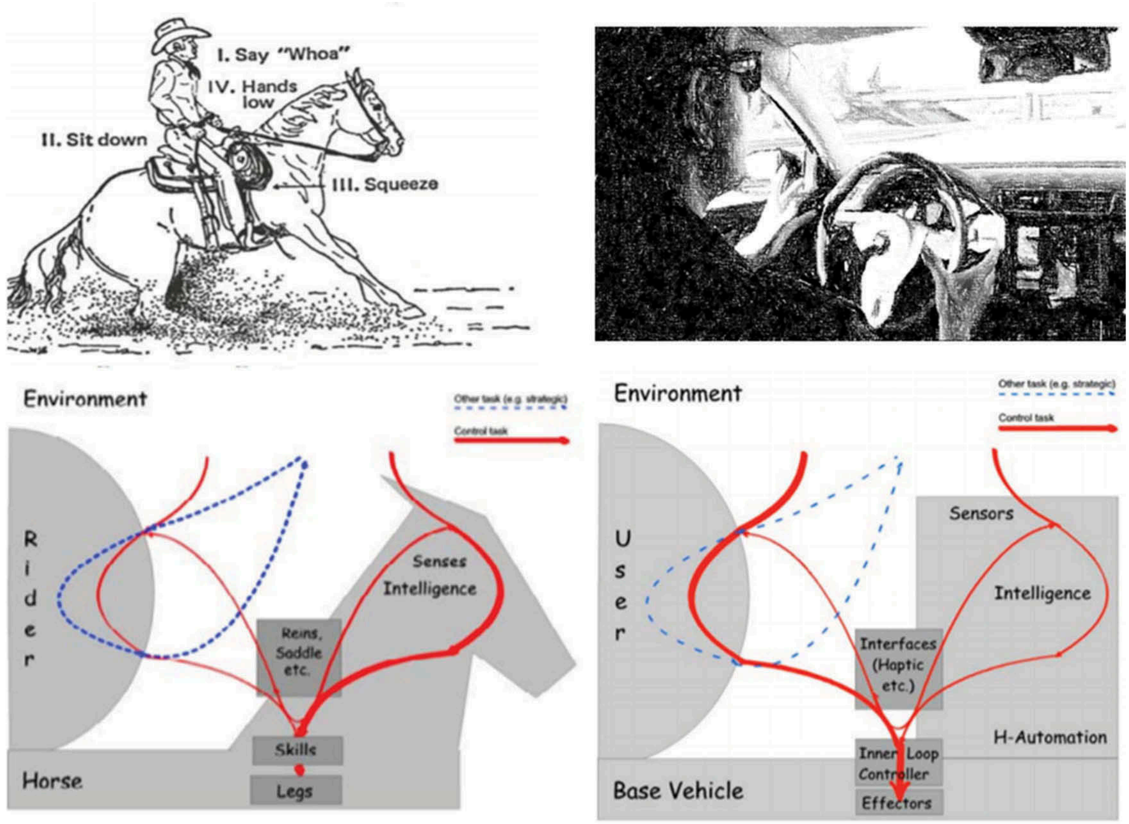

Figure 2. The H-Metaphor model for automated vehicle and user interactions.

The research performed in ACC and LKA showed that symbolic level cooperation is less effective than acting at both the drivers' symbolic, perceptual, and motor control processing levels (Hoc et al., 2006, Jordan, Franck, & Hoc, 2007). This is in line with the H-Metaphor requirement of a shared monitoring interface. Using Hoc and Boseville's framework, automotive interaction researchers can look at how to develop effective cooperation strategies across the perception, mutual control, function delegation, and full automation modes.

Effective cooperation is usually measured by the efficiency of the collaboration. Efficiency is extremely critical in the automated driving context for the purpose of safety and timely decision-making. This requires models that support the formation of a common plan between human and automated driving system. Such a model was presented as "intent-based cooperation" by Bauer, Wollherr, and Buss (2008). A key construct for intent-based cooperation is intention modeling. Based on knowledge from the human partner's intention and current activity, the automated system can plan its own actions to reach a common goal. Therefore, the system needs the abilities for perceiving and comprehending their environment, decision-making, planning, learning, and reflection.

Bauer also specifies that humans need to communicate their intentions, either deliberately by explicit communication (through speech, gesture, or haptic input), or implicitly by actions (via manipulative gestures, proactive task execution, and physiological signals). The recent development on intent inference using machine learning methods, such as reinforcement learning, show promise toward these type of intelligent cooperative systems. However, intent definition is largely still tailored for particular tasks. Automated driving systems need to develop more elaborate and scalable techniques for implicit and explicit communication handling with passengers. These include decision-making in instances in which explicit and implicit communication does not correlate, e.g., when the vehicle passenger points to the right but request verbally a left turn, when social politeness hides explicit expression of discomfort with driving behavior, or when the passengers delegate certain control functions to the vehicle but remain wary of the system's capability.

This last issue where control authority is shared between the vehicle and the human driver is the specific focus of many cooperation models in the automated driving literature. Abbink and colleagues presented the concept of control authority in human–machine cooperation via the shared haptic interface of force feedback on the steering-wheel (Abbink, Mulder, & Boer, 2012). In this system, an additional torque activated by the controller informs the driver that the current steering wheel position differs from the one that the machine estimates to be optimal, and at the same time the driver can force their own will by applying an increased torque to correct the ADAS input. We can see this model in commercial application of modern vehicles (SAE Level-2). However it is limited, because the ADAS system does not take into consideration visibility conditions, thus prompting the user to deactivate the automated system under rain, snow, or extreme light conditions for fear of erratic behavior (see

Mars, Deroo, & Hoc, 2014, for an extension of Abbink et al.'s model).

More recent work has evolved the control authority concept into a "shared authority" model that includes viewpoints of initiative holder and intention consistency (Nishimura, Wada, & Sugiyama, 2015). The Shared Authority model proposes a gain-tuning control method for lane-keeping assistance to enable the driver to change lanes smoothly.

There are, however, still issues present in the shared authority model. Human drivers may become confused if the degree of shared control changes with the context. We need longitudinal studies to understand how drivers adapt to shared control systems, and to deconstruct the human adaptive mappings of automation levels and capabilities, in order to find the most effective methods to learn the preferences of individual drivers, and match their needs and goals.

A step in that direction was proposed with the general framework for transitions in automated driving presented by Lu, Happee, Cabrall, Kyriakidis, and de Winter (2016; see Figure 3). In this model, they define the static states for primary driving tasks between automation and driver with specific longitudinal and lateral control, driver monitoring, as well as the transitions from one driving state to another.

In this model, the vehicle actuators are controlled by switches that activate automation control, and add parameters to account for dynamic increase/reduction of input, during transitions and shared control periods. The driver monitoring level can also modulate the human driver input, and even account for cases such as distracted driving. While the parameterization in this model provides certain flexibility to changes in automation levels, it does not account for the learning process that collaboration between vehicle and individual driver might undergo.

A learning-based model, where the automated driving system learns to increase autonomy by increasing its competences and managing the system limits, was proposed by Vanderhaegen (2012). This AI-based model (see Figure 4) focuses on two capabilities, the capacity of human and system to cooperate and the capacity of the system to self-learn from these collaborative actions. Vanderhaegen proposes an adjustment of the system learning parameters to accelerate the convergence of the correct prediction rate in a similar process to imitation learning. The model pre-defines the cooperation between human and AI and the adjustment of the learning parameters until it can be validated by the driver after several iterations.

The two limitations on this model are how the user can validate the learning as correct, based on his input, and how the automated driving systems can integrate the learning competences while maintaining the validation, i.e., ensure

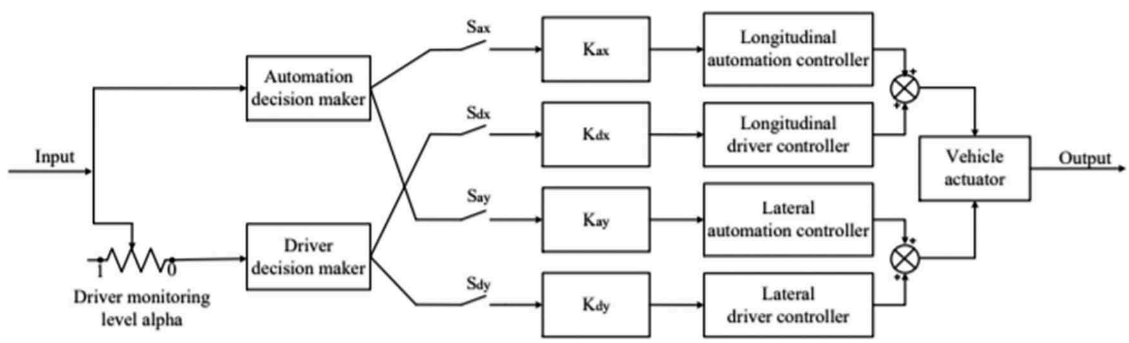

Figure 3. General framework for state transition in automated driving by (Lu et al., 2016).

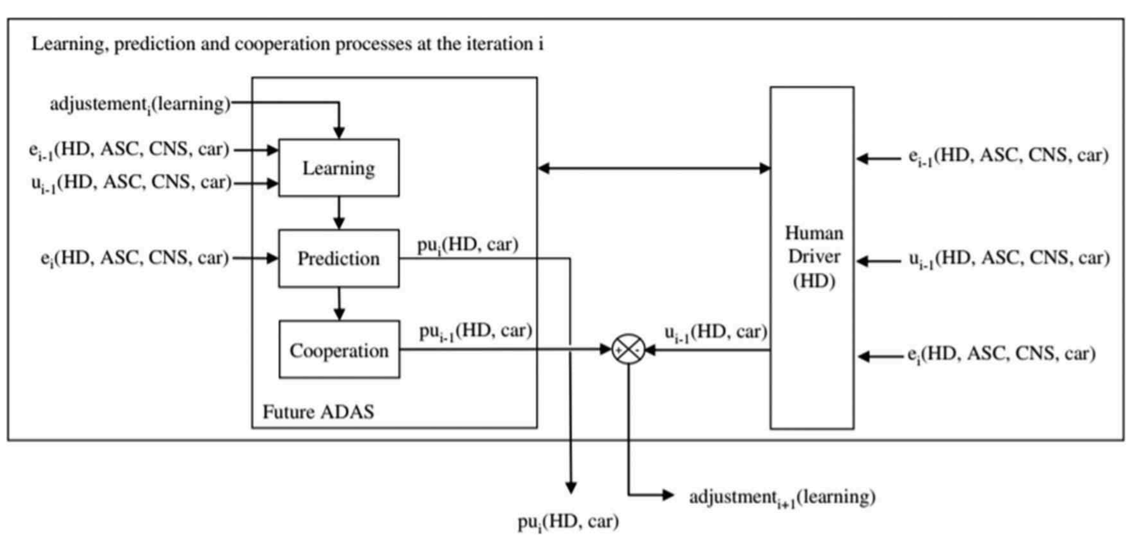

Figure 4. Learning cooperation process for automated driving from Vanderhaegen (2012).

the predictability of correct behavior under all driving conditions. Those are critical drawbacks that require deep research and engineering focus. However, these novel models have the ability to produce highly personalized collaborative interactions adapting to both the dynamic driving conditions, and the learning process of both human and automated vehicle to work with each other.

The latest development of learning models of collaboration brings us closer to a human–system symbiosis that was proposed in the 60s by Licklider (1960). In his vision, computers would facilitate "formulative thinking," rather than the solution of formulated problems. In essence augment the human through collaboration rather than substitute him.

3. The human element in human–vehicle cooperation

The human element will determine the success or demise of automated driving. In the previous section, we discussed how models of human–robot interaction can inform the development of cooperative driving systems. We have also hinted at the effect that automation may have on driver behavior, and the relevance that workload, trust, and other human-related aspects will have in determining the adoption of vehicle automation. In this section, we will examine these aspects in more detail. We will also discuss the social element associated with automated vehicles and personal in-vehicle assistants. In a connected traffic environment, automated vehicles will also need to interact with road users outside the vehicle. We will provide a critical review of the human challenges associated with automated driving, and offer a discussion on how the literature on user automation in fields outside driving can inform the development of a safe human–vehicle cooperation.

3.1. Mental workload

The human mind is limited in its ability to process information (Kahneman, 1973; Pashler, 1994). Studies on dual-task performance show human performance in a single-task scenario to worsen following the introduction of additional resource-demanding activities (Pashler, 1994; Tombu & Jolicœur, 2003). In the driving domain, distraction represents a manifestation of the driver inability to process multiple sources of information at once or for a sustained amount of time.

In their tri-pronged approach to driver distraction, Strayer, Watson, and Drews (2011) described three sources of distraction. Manual, visual, and cognitive distractions occur when drivers' hands are not on the steering wheel, their eyes are not on the forward roadway, and their attention is away from the driving task, respectively. Research on driver distraction shows inattention to slow drivers' responses and impair their voluntary control of the vehicle (Hosking, Young, & Regan, 2009; Ranney, Harbluk, & Noy, 2005; Rossi, Gastaldi, Biondi, & Mulatti, 2012). When talking on a hands-free cell phone, for instance, cognitive distraction causes drivers to maintain a more variable speed and distance to the lead car (Haigney, Taylor, & Westerman, 2000; Strayer & Drews, 2007), thus making their behavior less predictable to other road users.

Cognitive distraction is also associated with impaired visual scanning of the forward roadway. Biondi, Turrill, Coleman, Cooper, and Strayer (2015) had participants drive an on-road vehicle while executing one of six possible driving-unrelated tasks. The route selected for the study presented a number of hazard locations, ranging from un-signalized and stop sign-controlled intersections, to school zones and crosswalks. The increasing amount of cognitive workload associated with the secondary-task deteriorated driver visual scanning of the traffic environment, impairing their ability to predict the occurrence of road hazards (Strayer & Fisher, 2016). Taylor et al. (2013) found similar results, with drivers' accuracy in anticipating potential hazards decreasing from 77% in the non-distracted condition, to 35% when distracted.

Given the potential for mental overload and distraction to impair the human ability to attend to relevant information within the road environment, recent research has begun to examine the effect that automated driving, and partial automation in particular, may have on the driver workload, and their ability to stay engaged in the monitoring task.

The Yerkes-Dodson law (Yerkes & Dodson, 1908) defines an inverted-U shape relationship between workload and human performance, whereby optimal performance in a given task is observed under conditions of intermediate workload, while lower accuracy and slower responses are expected in situations of either under-load or overload (see Figure 5). During manual driving, this approach has been adopted to account for phenomena like distraction and fatigue. In the context of partial automation, it is hypothesized that relinquishing control of vehicle operations to the system will bring the driver into a phase of mental lull, in which the system controls acceleration and steering, and the human operator has nothing to attend to. The research conducted by Biondi et al. (2018) had tested this hypothesis with participants driving a semi-automated vehicle on the open road. Driver state was monitored using a combination of self-reported, behavioral, and physiological measures. To investigate the effect of mode on driver workload and arousal, participants drove in both manual and semi-automated mode. Consistently with what is hypothesized by the Yerkes-Dodson law, operating a partially automated vehicle decreased driver levels of physiological activation, and their performance in a detection task dropped during semi-automated driving.

Self-regulation is observed when drivers adopt resource allocation strategies as a result of changes in task demand. While *reactive* self-regulation occurs when the driver adjusts their secondary-task activities following a spike in driving demand (e.g., a call is terminated upon entering a school zone), *proactive* self-regulation occurs when levels of mental and physical workload plummet, and drivers attempt to return to optimal levels of functioning by regulating their own behavior. The drop in physical and mental workload resulting from partial automation is expected to trigger a phase of self-regulation, with drivers becoming engaged in distracting activities, and further reducing the level of attention directed toward the monitoring of the automated system.

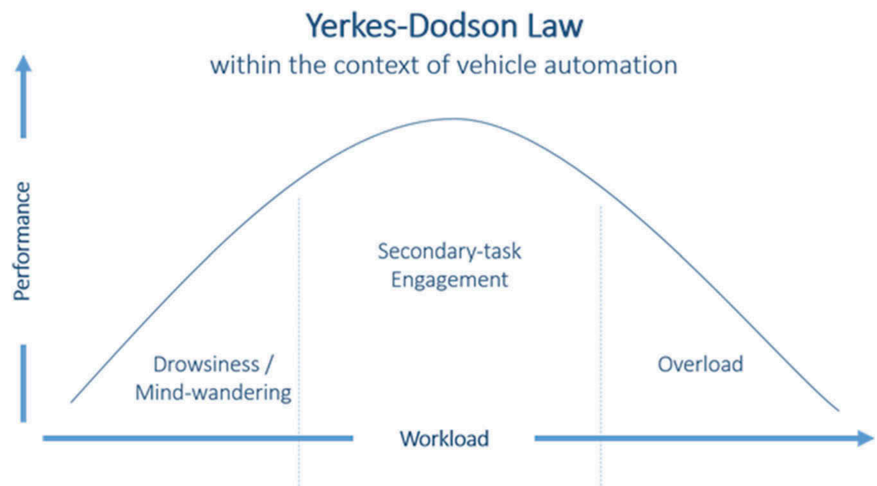

Figure 5. Yerkes-Dodson law applied to automated driving.

Observational data collected in the studies by Endsley (2017), Biondi, Goethe, Cooper, and Strayer (2017), and Banks, Eriksson, O'Donoghue, and Stanton (2018) confirm this. Following the activation of the semi-automated systems, drivers were indeed observed to become completely "hands and feet free" and only placed their hands back on the steering wheel as a result of system warnings.

3.2. Transitions of control

One safety consequence of driver disengagement during semi-automated driving is for the human operator to be unprepared to take back control during system-to-driver transitions of control. A distinction can be made between anticipated vs. unanticipated transitions of control. Anticipated transitions occur when the driver is informed about the system inability to stay in control of the vehicle under certain road or traffic scenarios – e.g., upcoming roadworks, traffic jams. Unanticipated transitions, instead, occur when the system unexpectedly fails or the driver is not given sufficient notice about the upcoming request for transition – e.g., the lane markings are worn out and the system can no longer stay in control of the system.

In a study by Stanton, Young, and McCaulder (1997), authors noted that during level-1 automation, drivers failed to regain control of the vehicle following automation failure more than one third of the times. Recent studies have examined system-to-human transitions of control, and measured the time needed for drivers to regain control of the vehicle following a system failure. Research conducted by Gold, Korber, Lechner, and Bengler (2016), Louw, Merat, and Jamson (2015), and Walch, Lange, Baumann, and Weber (2015) indicate that warning messages should be presented with at least a 2-to-3-second notice, before the system completely disengages. Belderbos (2015), and Naujoks and Nekum (2014) argue differently, showing driver take-over times of around 6 s – up to 3 times longer than what suggested by Louw et al. and Walch et al. Eriksson and Stanton (2017) advocate in favor of a less deterministic, more adaptive approach, suggesting that factors like the road scenario, the characteristics of the driver and workload, among other things, will have to be accounted for when designing transition maneuvers. Results from their study with participants being instructed to take over control of the vehicle in increasingly demanding scenarios, indeed show take-over response times to increase as a factor of the demand associated with the secondary task.

Despite such findings offering important indications for the design of take-over maneuvers, all these studies were conducted in simulated environments, with participants driving in highly controlled scenarios. Thus, we argue that take-over response times collected in these studies only provide conservative estimates that cannot be fully reflective of what would happen in real-world situations.

To our knowledge, there are no studies in literature examining driver performance during take-over maneuvers with on-road highly automated vehicles. However, empirical studies have investigated the time it takes drivers to redirect their full attention to the primary driving task, following the disengagement from driving-unrelated activities. In the study by Strayer, Cooper, Turrill, Coleman, and Hopman (2015), for example, participants drove a manual car while switching between on-task and off-task scenarios. When off-task, driving was their only task. When on-task, drivers also completed infotainment activities using the in-vehicle voice command system. Dynamics fluctuations in driver workload were measured using the ISO version of the peripheral detection task (Detection Response Task or DRT; ISO, 2016; see also Harbluk, Burns, Hernandez, Tam, & Glazduri, 2013; Strayer, Biondi, & Cooper, 2016). To respond to intermittent visual stimuli, the driver had to press a micro-switch as fast and accurately as possible. The faster and more accurate their responses, the lower the cognitive demand of the secondary task. Switching from off-task to on-task led to longer DRT response times and lower accuracy, as expected. Vice versa, when the secondary task was not required, higher DRT performance was observed. However, a more fine-grained analysis showed that, to resume the DRT performance recorded during off-task driving, drivers took an average of 27 s. In other words, the process of completely disengaging from the secondary task and fully recovering the resources consumed when *multitasking* was not immediate or cost-free, but required at least 27 s, on average. This phenomenon, also known in experimental research as switch cost or residual cost (Monsell, 2003), is indicative of the human inability to multitask, and, within the context of highly automated driving, it demonstrates the complexity of transition-of-control maneuvers and their potential to require longer times than what showed in the driving simulator literature.

3.3. Situation awareness and trust

Situation awareness. As the user becomes less attentive toward the automated system, this is often found to decrease their levels of monitoring of the system and awareness of its functioning. Endsley (1995) defines situation awareness as "the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" (p. 36). As drivers become disengaged from the driving task, it is expected for them to become even less aware not only of the correct functioning of the automation, but also of potential safety hazards occurring outside the vehicle.

In the study on spacecraft operators, Reichenbach, Onnasch, and Manzey (2010) had participants interact with an apparently reliable system. As participants became more familiar with automation over time, this decreased their performance in monitoring the system and detecting system failures (for similar results, see Bailey & Scerbo, 2007; Hoc, 2001). In the driving field, studies conducted by Stanton and Young (2005) and Vollrath, Schleicher, and Gelau (2011) on driver interaction with level-1 automated systems show similar findings. Stanton and Young had drivers drive a simulated vehicle in two conditions: Adaptive Cruise Control (ACC) ON or OFF. Information about the functioning of the ACC was presented by either one, two, or three warning messages (low, medium, and high feedback, respectively). Driving with ACC ON under the medium feedback condition helped drivers maintain a safe speed, with no negative impact on their monitoring ability. However, when driving in the high feedback condition, higher levels of mental workload imposed by the more complex warning system impoverished driver awareness of the surrounding traffic. Similar findings were obtained by Vollrath et al. (2011). Drivers drove a simulated vehicle in three conditions: manual driving, regular Cruise Control, and ACC. Fewer speed violations were recorded with increasing automation. However, relative to manual driving or driving with Cruise Control activated, having ACC ON deteriorated driver situation awareness, and their ability to readily respond to upcoming hazards.

Strayer, Getty, Biondi, and Cooper (in press) speculate on the effect that partial automation will have on driver situation awareness. In the context of the theoretical model developed by Strayer and Fisher (Fisher & Strayer, 2014; Strayer & Fisher, 2016), situation awareness is affected by five intertwined mental processes that include: scanning specific areas, predicting where the threats might materialize, identifying target objects in the scenario, deciding whether an action is necessary, and executing appropriate responses – SPIDER for short. In the SPIDER model, situation awareness is represented as a cylinder, with a larger circumference representing higher moment-to-moment levels of situation awareness (see Figure 6). Applied to partial automation, the initial engagement of the semi-automated system, and the resulting transfer of vehicle control

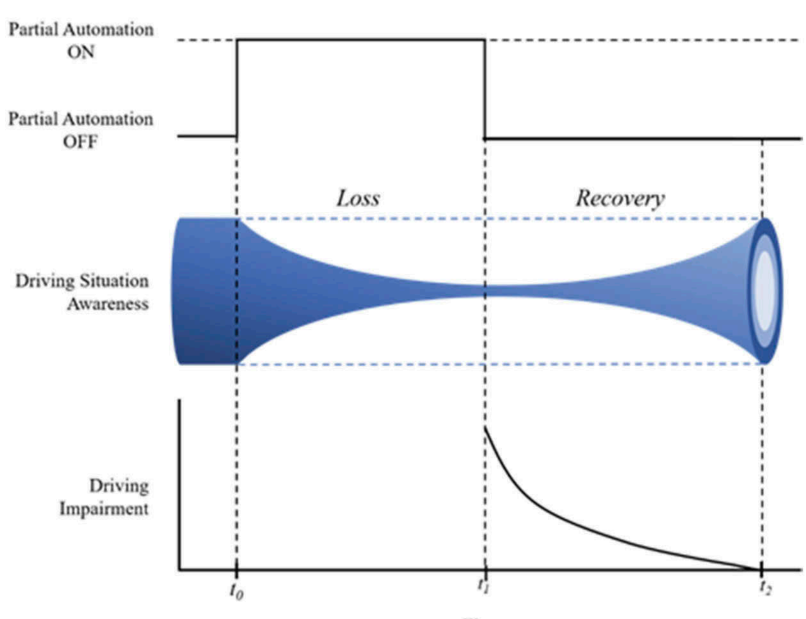

Figure 6. SPIDER model applied to partial automation. Partial automation is engaged at t_0, and disengaged at t_1. Full situation awareness is recovered at t_2. Adapted from Strayer et al. (in press).

from the user to the automation, is expected to result in a steady loss in driver situation awareness. With the semi-automated system requiring periodic inputs by the human driver (especially in unstable driving scenarios), this is expected to trigger a period of recovery. The length of this period and the driver ability to regain control of the vehicle are dictated by the loss in situation awareness occurring prior to the system disengagement.

Trust. Researchers tend to agree that trust in automation is currently one of the biggest threats the automobile industry faces in terms of technology acceptance. Lee and See defined trust as “the attitude that an agent will help achieve an individual’s goal in a situation characterized by uncertainty and vulnerability” (p. 51) (Lee & See, 2004). We argue that a more useful definition of trust is a mental model that the user makes of the automated driving operational capabilities based on previous observations (or experience). As such, trust must be considered as a dynamic process that evolves as more experience is accumulated.

Research investigating trust in automation in participants with no prior experience of the automated system, report user complacency to increase over time (Waytz, Heafner, & Epley, 2014). Lee and Moray identified an inverse relationship between trust and monitoring, meaning that the more trust there was, the less the system was monitored (Lee & Moray, 1992). A review of empirical studies by De Winter et al. showed that drivers were more likely to engage in secondary tasks under automated driving conditions (de Winter, Happee, Martens, & Stanton, 2014). More recently these results have been corroborated by Payre et al. in simulator studies under fully automated conditions (Payre, Cestac, & Delhomme, 2016).

Current literature suggests that there are two dangerous trends with regard to trust in automated driving vehicles. Under-trust which poses the threat of technology rejection and over-trust which can cause over-reliance. As noted by Miller et al. (2016), the latter presents safety risks under unpredictable circumstances where the automation is not able to perform, or it cannot perform as expected. The continuous calibration of trust in the automated driving system requires a close cooperation between the system and the human operator, with the system providing accurate information about its modes of operation (in the case of automated driving, think of a system that can automate acceleration and steering only at speeds lower than 15 MPH). When this assumption is violated, this incurs in confusion or *mode error*. Sarter and Woods (1995) showed empirical evidence that mode error occurred primarily in the context of non-normal, time-critical situations, and that operator sometimes had problems anticipating system behavior and the associated mode.

Mode error is proven to be an egregious source of accidents, even in the relatively short history of automated driving. The 2017 investigation conducted by NTSB on the first-ever fatal accidents involving a semi-automated system, for example, suggests that the confusion regarding the limitation and capabilities of the Tesla Autopilot system is likely to have contributed to the use of the system in a way that was unintended. Studies by Beggiato, Pereira, Petzoldt, and Krems (2015) and Sullivan, Flannagan, Pradhan, and Bao (2016) also point at drivers’ fragmented knowledge of assistance systems’ capabilities and limitations as a determinant in accelerating the process leading to system misuse (Goddard, Roudsari, & Wyatt, 2012).

3.4. Social qualities of human–vehicle interaction

As technology advances and humans become less involved in the driving task, vehicles are expected to increasingly become places for work, leisure, or social interaction. Research in the field of human–robot interaction has long investigated the social element involved in the relationship between users and automatons.

Breazeal (2003) defines *social robots* as those that people anthropomorphize in order to interact with them. She also categorized social robots in terms of how well the robot can support the social model, and the complexity of the interaction scenario it can support. Fong, Nourbakhsh, and Dautenhahn (2003) extended this classification, by distinguishing between *socially situated* robots (that are surrounded by a social environment that they perceive and react to), *socially embedded* robots (that are also aware of human interactional structures), and *socially intelligent* robots (that can support a wide range of users and user interactions).

Automotive companies have attempted to introduce smart, social agents as part of the in-vehicle driving experience. At the 2005 and 2007 Tokyo Motor Show, for instance, Nissan presented PIVO, a two-seater pod on wheels with an anthropomorphic robot embedded in the dashboard (the robot was allegedly capable of improving the state of mind of the driver). More realistic solutions have recently been introduced in the form of personal assistants like Android Auto, Apple CarPlay, and Amazon Alexa, which, by linking the in-vehicle infotainment system with the user’s calendar, preferred navigation entries, etc., are designed to learn about the user’s personal preferences, and establish a human-like relationship with them (see Strayer et al., 2018, for an assessment of these systems).

Designing machines with anthropomorphic traits has been shown to enhance users’ understanding of the system, and increase the accuracy of their mental models. Research conducted in automotive HCI has demonstrated that automated driving assistants that displayed higher behavioral skills were in fact regarded as more capable and trustworthy, even when their capabilities were the same as more “robotic” assistants (Kiesler & Goetz, 2002; Nass et al., 2005; Waytz et al., 2014). There is also evidence suggesting that a mental model of collaboration is less linked to physical appearances, such as high-end graphics, natural sounding voices, or human-like faces. Instead, socio-behavioral skills like conversational turn-taking and lexicon-adaptation appear to facilitate more collaborative interactions, and reduce the risk of mistreatment (Carlson, Lemmon, Higgins, & Frank, 2017).

Aside from in-vehicle personal assistants, another relevant area in which the introduction of social agents will have a significant impact for the development and safe adoption of automated driving systems is that of the communication between the vehicle and road users. Nowadays, road users heavily rely on an efficient verbal and non-verbal communication to maintain safety, especially within urban environments. At a 4-way stop intersection, for instance, pedestrians often wait for motorists to acknowledge their presence (by establishing eye contact etc.), before attempting to cross the street. With automated driving systems, exterior human–machine interfaces will have to allow for a seamless, human-like interaction with road users, which needs to be tailored to the cultural characteristics of the end-user (Bartneck, Nomura, Kanda, & Suzuki, 2005; Bonnefon, Shariff, & Rahwan, 2016; Shinohara, Currano, Ju, & Nishizaki, 2017; Złotowski, Yogeeswaran, Human, & Bartneck, 2017). Recent research has begun to address this issue. The study by Rothenbücher, Li, Sirkin, Mok, and Ju (2016) with pedestrians and bicyclists interacting with a *fake* fully automated driving system, for instance, has shown some road users to be wary of the self-driving vehicle, and experience feelings of uncertainty, especially when trying to cross the street. Similar research has also attempted to tackle the issue of low user trust and acceptance of this technology, by developing proof-of-concept designs that will facilitate a more human-like communication between automated and human road users (Siripanich, 2017).

4. Discussion and conclusions

This literature review provides a perspective on the impact of automation on human–vehicle interaction. For decades, human ingenuity has been poured over the automobile, transforming transportation through faster, more efficient, more secure, and more comfortable machines. But we are now at an inflexion point where new technological horizons are open. There is little disbelief that automated driving is going to revolutionize transportation and transform our daily lives. But the time is right to ask one question. What is the cost of automation?

To answer this question, we reviewed key studies from the literatures on HCI, human robot interaction, and human factors and ergonomics. Our analysis on vehicle automation provides an overview of the advances in safety, security, and comfort that the automotive community has achieved in the last years. But it also illustrates the dangerous unintended consequences of automation technology. While the design of automation in automobiles aims at increasing safety, we have showed how the inadequate design of driver assistance systems can instead *encourage* human operators to become complacent, underaroused and, in turn, less responsive to traffic hazards. By reviewing studies on user automation outside the driving field, we also reported the safety risks that automation misuse, disuse, and abuse may have on drivers' distraction, and their inability to adequately respond to handover requests.

The incoming mass adoption of artificial intelligence in all the areas of vehicle automation has also important implications to how we will interface with these systems. Transparency and explainability of automated driving operation will be critical to construct a cooperative human–system environment. Thus our research needs to focus on the interaction issues that rise from human operators and intelligent vehicle assistants, and do not expect passengers to be passive operators able to discern the operational intricacies of the automated vehicle.

Our review of human–system cooperation models in automated driving provides the readers a refreshed inspiration based on the aspirational models of human–computer symbiosis. We argue the need to move beyond human supervisory roles, and embrace the possibilities of peer-to-peer relationships in models that guarantee system flexibility in coordination and disambiguation of roles and control tasks.

We hope this review provides clear indications of how different disciplines in the field of human–machine interaction will have to come together to build a comprehensive understanding of driver–vehicle cooperation. We are also confident that this document will help researchers and designers overcome some of the design challenges that will arise when developing interfaces for highly and fully automated vehicles.

References

Abbink, D. A., Mulder, M., & Boer, E. R. (2012). Haptic shared control: Smoothly shifting control authority? *Cognition, Technology & Work*, *14*(1), 19–28. doi:10.1007/s10111-011-0192-5

ACEA. (2017). *What is truck platooning?* Retrieved from http://www.acea.be/uploads/publications/Platooning_roadmap.pdf

Argall, B. D., Chernova, S., Veloso, M., & Browning, B. (2009). A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, *57*(5), 469–483. Retrieved from https://www.sciencedirect.com/science/article/pii/S0921889008001772

Bailey, N. R., & Scerbo, M. W. (2007). Automation-induced complacency for monitoring highly reliable systems: The role of task complexity, system experience, and operator trust. *Theoretical Issues in Ergonomics Science*, *8*, 321–348. doi:10.1080/14639220500535301

Banks, V. A., Eriksson, A., O'Donoghue, J., & Stanton, N. A. (2018). Is partially automated driving a bad idea? Observations from an on-road study. *Applied Ergonomics*, *68*(October2017), 138–145. doi:10.1016/j.apergo.2017.11.010

Bartneck, C., Nomura, T., Kanda, T., & Suzuki, T. (2005). A cross-cultural study on attitudes towards robots. *Proceedings of the HCI International*. Retrieved from http://www.academia.edu/download/30706755/bartneckHCII2005.pdf

Bauer, A., Wollherr, D., & Buss, M. (2008). Human–Robot collaboration: A survey. *International Journal of Humanoid Robotics*, *5*(01), 47–66. doi:10.1142/S0219843608001303

Beggiato, M., Pereira, M., Petzoldt, T., & Krems, F. (2015). Learning and development of trust, acceptance and the mental model of ACC. A longitudinal on-road study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *35*, 75–84. Retrieved from https://www.sciencedirect.com/science/article/pii/S1369847815001564

Belderbos, C. (2015). *Authority transition interface: A human machine interface for taking over control from a highly automated truck*. Delft, Netherlands: Delft University of Technology.

Biondi, F., Goethe, R., Cooper, J., & Strayer, D. (2017). *Partial-autonomous frenzy: Driving a level-2 vehicle on the open road*. International Conference on Engineering Psychology and Cognitive Ergonomics (pp. 329–338). Springer, Cham.

Biondi, F., Lohani, M., Hopman, R., Mills, S., Cooper, J., & Strayer, D. (2018). 80 MPH and out-of-the-loop: Effects of real-world semi-automated driving on driver workload and arousal. Proceedings of the Conference of the Human Factors and Ergonomics Society, Los Angeles, CA.

Biondi, F., Lohani, M., Mills, S., Cooper, J., & Strayer, D. (under review). *On the effects of on-road partial automation on driver workload and arousal*. Human Factors. Los Angeles, CA: SAGE Publications.

Biondi, F., Turrill, J., Coleman, J. R., Cooper, J. M., & Strayer, D. L. (2015). *Cognitive distraction impairs drivers' anticipatory glances: An on-road study*. Proceedings of the Eighth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design Distractive (pp. 23–29). Salt Lake City, Utah.

Biondi, F. N., Getty, D., McCarty, M. M., Goethe, R. M., Cooper, J. M., & Strayer, D. L. (2018). The challenge of advanced driver assistance systems assessment: A scale for the assessment of the human–machine interface of advanced driver assistance technology. *Transportation Research Record: Journal of the Transportation Research Board*. 36119811877356. doi:10.1177/0361198118773569

Bonnefon, J., Shariff, A., & Rahwan, I. (2016). The social dilemma of autonomous vehicles. *Science*, *352*(6293), 1573–1576. Retrieved from http://science.sciencemag.org/content/352/6293/1573.short

Breazeal, C. (2003). Toward sociable robots. *Robotics and Autonomous*. Retrieved from https://www.sciencedirect.com/science/article/pii/S0921889002003731

Carlson, Z., Lemmon, L., Higgins, M., & Frank, D. (2017). This robot stinks! Differences between perceived mistreatment of robot and computer partners. *Arxiv.org*. Retrieved from https://arxiv.org/abs/1711.00561

de Winter, J. C., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*, 196–217. doi:10.1016/j.trf.2014.06.016

Defense Science Board. (2012). *Task force report on: The role of autonomy in DoD systems*. Department of Defense. Retrieved from https://fas.org/irp/agency/dod/dsb/autonomy.pdf

Dias, M., Kannan, B., & Browning, B. (2008). *Sliding autonomy for peer-to-peer human-robot teams*. Proceedings of the international conference on intelligent autonomous systems (pp. 332–341). Nice, France.

Dosovitskiy, A., Ros, G., Codevilla, F., & López, A. (2017). CARLA: An open urban driving simulator. *Arxiv.org*. Retrieved from https://arxiv.org/abs/1711.03938

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, *37*(1), 32–64. doi:10.1518/001872095779049543

Endsley, M. R. (2017). Autonomous driving systems: A preliminary naturalistic study of the Tesla model S. *Journal of Cognitive Engineering and Decision Making*, *11*(3), 225–238. doi:10.1177/1555343417695197

Eriksson, A., & Stanton, N. A. (2017). Takeover time in highly automated vehicles: Noncritical transitions to and from manual control. *Human Factors: the Journal of the Human Factors and Ergonomics Society*, *59*(4), 689–705. doi:10.1177/0018720816685832

European Commisssion. (2016). *Roadmap on highly automated vehicles*. Retrieved from https://circabc.europa.eu/sd/a/a68ddba0-996e-4795-b207-8da58b4ca83e/DiscussionPaper-RoadmaponHighlyAutomatedVehicles08-01-2016.pdf

Fisher, D., & Strayer, D. L. (2014). Modeling situation awareness and crash risk. *Annals of Advances in Automotive Medicine*, *58*, 33–39. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4001668/

Flemisch, F. O., Adams, C. A., Conway, S. R., Goodrich, K. H., Palmer, M. T., & Schutte, P. C. (2003). *The H-Metaphor as a guideline for vehicle automation and interaction*. Hanover, MD: NASA Center for AeroSpace Information.

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems*. Retrieved from https://www.sciencedirect.com/science/article/pii/S092188900200372X

Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: A systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, *19*(1), 121–127. doi:10.1136/amiajnl-2011-000089

Gold, C., Korber, M., Lechner, D., & Bengler, K. (2016). Taking over control from highly automated vehicles in complex traffic situations: The role of traffic density. *Human Factors: the Journal of the Human Factors and Ergonomics Society*, (2014). 18720816634226. doi:10.1177/0018720816634226.

Goodrich, M., & Schultz, A. (2018). Human-robot interaction: A survey. *Foundation and Trends in Human - Computer Interaction*, *1*(3), 203–275. Retrieved from https://dl.acm.org/citation.cfm?id=1348100

Haigney, D., Taylor, R., & Westerman, S. (2000). Concurrent mobile (cellular) phone use and driving performance: Task demand characteristics and compensatory processes. *Transportation Research Part F: Traffic Psychology and Behaviour*, *3*(3), 113–121. doi:10.1016/S1369-8478(00)00020-6

Harbluk, J. L., Burns, P. C., Hernandez, S., Tam, J., & Glazduri, V. (2013). *Detection response tasks: Using remote, headmounted and tactile signals to assess cognitive demand while driving*. Proceedings of the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design (pp. 78–84). Bolton Landing, NY.

Hemminghaus, J., & Kopp, S. (2017). *Towards adaptive social behavior generation for assistive robots using reinforcement learning*. Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction (pp. 332–340). Vienna, Austria. Retrieved from https://dl.acm.org/citation.cfm?id=3020217

Hinds, P., Roberts, T., & Jones, H. (2004). Whose job is it anyway? A study of human-robot interaction in a collaborative task. *Human-Computer Interaction*, *19*(1), 151–181. Retrieved from https://dl.acm.org/citation.cfm?id=1466554

Hoc, J. M. (2001). Towards a cognitive approach to human–machine cooperation in dynamic situations. *International Journal of Human-Computer Studies*, *54*(4), 509–540. doi:10.1006/ijhc.2000.0454

Hoc, J. M., & Blosseville, J. M. (2003). Cooperation between drivers and in-car automatic driving assistance. *In Proceedings of CSAPC*, *3*, 17–22.

Hoc, J. M., Mars, F., Milleville-Pennel, I., Jolly, É., Netto, M., & Blosseville, J. M. (2006). Human-machine cooperation in car driving for lateral safety: Delegation and mutual control. *Le travail humain*, *69*(2), 153–182. doi:10.3917/th.692.0153

Hosking, S. G., Young, K. L., & Regan, M. A. (2009). The effects of text messaging on young drivers. *Human Factors: the Journal of the Human Factors and Ergonomics Society*, *51*(4), 582–592. doi:10.1177/0018720809341575

Inagaki, T., & Sheridan, T. B. (2018). A critique of the SAE conditional driving automation definition, and analyses of options for improvement. *Cognition, Technology and Work, (Sae 2016)*, 1–10. doi:10.1007/s10111-018-0471-5

International Organization for Standardization. (2016). *ISO 17488: 2016. Road vehicles — Transport information and control systems — Detection-Response Task (DRT) for assessing attentional effects of cognitive load in driving*. Retrieved from https://www.iso.org/standard/59887.html

Jordan, N., Franck, M., & Hoc (2007, August). *Lateral control support for car drivers: A human-machine cooperation approach*. Proceedings of the 14th European conference on Cognitive ergonomics: invent! explore! (pp. 249–252). London, UK: ACM.

Kahneman, D. (1973). *Attention and effort*. (Englewood Cliffs: Prentice-Hall., Ed.). *Attention and Effort*. doi:10.2307/1421603

Kiesler, S., & Goetz, J. (2002). Mental models of robotic assistants. *Extended abstracts on Human Factors in Computing*. Retrieved from https://dl.acm.org/citation.cfm?id=506491

Kober, J., Bagnell, J. A., & Peters, J. (2013). Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research*, *32*(11), 1238–1274. doi:10.1177/0278364913495721

Lee, J., & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, *35*(10), 1243–1270. doi:10.1080/00140139208967392

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors: the Journal of the Human Factors and Ergonomics Society*, *46*(1), 50–80. doi:10.1518/hfes.46.1.50_30392

Licklider, J. C. (1960). Man-computer symbiosis. *IRE Transactions on Human Factors in Electronics*, (1), 4–11. doi:10.1109/THFE2.1960.4503259

Litman, T. (2017). *Autonomous vehicle implementation predictions*. Retrieved from https://www.vtpi.org/avip.pdf

Liu, R. (2016). *Automated transit: Planning, operation, and applications*. Hoboken, NJ: John Wiley & Sons.

Louw, T., Merat, N., & Jamson, H. (2015). *Engaging with highly automated driving: To be or not to be in the loop?* Paper presented at the Eighth International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, Salt Lake City, Utah

Lu, Z., Happee, R., Cabrall, C. D., Kyriakidis, M., & de Winter, J. C. (2016). Human factors of transitions in automated driving: A general framework and literature survey. Transportation research part F. *Traffic Psychology and Behaviour*, *43*, 183–198. doi:10.1016/j.trf.2016.10.007

Mahi, S., Atkins, M., & Crick, C. (2017). *Learning to assess the cognitive capacity of human partners*. Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction (pp. 63–64). Vienna, Austria. Retrieved from https://dl.acm.org/citation.cfm?id=3038430

Mars, F., Deroo, M., & Hoc, J. M. (2014). Analysis of human-machine cooperation when driving with different degrees of haptic shared control. *IEEE Transactions on Haptics*, *7*(3), 324–333. doi:10.1109/TOH.2013.2295095

Miller, D., Johns, M., Mok, B., Gowda, N., Sirkin, D., Lee, K., & Ju, W. (2016). Behavioral measurement of trust in automation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *60*(1), 1849–1853. doi:10.1177/1541931213601422

Mitsunaga, N., Smith, C., Kanda, T., & Ishiguro, H. (2006). Robot behavior adaptation for human-robot interaction based on policy gradient reinforcement learning. *Journal of the Robotics Society of Japan*, *24*(7), 820–829. Retrieved from https://www.jstage.jst.go.jp/article/jrsj1983/24/7/24_7_820/_article/-char/ja/

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*(3), 134–140. doi:10.1016/S1364-6613(03)00028-7

Nass, C., Jonsson, I. M., Harris, H., Reaves, B., Endo, J., Brave, S., & Takayama, L. (2005, April). *Improving automotive safety by pairing driver emotion and car voice emotion*. CHI'05 extended abstracts on Human factors in computing systems (pp. 1973–1976). Portland, OR: ACM.

National Transportation Safety Board. (2017). *Collision between a car operating with automated vehicle control systems and a tractor-semitrailer truck, Williston, FL, May 7, 2016*. Retrieved from https://www.ntsb.gov/news/events/Documents/2017-HWY16FH018-BMG-abstract.pdf

National Transportation Safety Board. (2018). *Preliminary report issued for investigation of fatal, Mountain View, California, Tesla crash*. Retrieved from https://www.ntsb.gov/news/press-releases/Pages/nr20180607.aspx

Naujoks, F., & Nekum, A. (2014). *Timing of in-vehicle advisory warnings based on cooperative perception*. Proceedings of the Human Factors and Ergonomics Society Europe Chapter Annual Meeting, Turin, Italy. Retrieved from http://hfes-europe.org

NHTSA. (2017). *Automated driving systems automated driving systems 2.0: A vision for safety*. Retrieved from https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/13069a-ads2.0_090617_v9a_tag.pdf

Nishimura, R., Wada, T., & Sugiyama, S. (2015). Haptic shared control in steering operation based on cooperative status between a driver and a driver assistance system. arXiv preprint arXiv:1504.05770.

Noy, I. Y., Shinar, D., & Horrey, W. J. (2018). Automated driving: Safety blind spots. *Safety Science*, *102*(October2017), 68–78. doi:10.1016/j.ssci.2017.07.018

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans.*, *30*(3), 286–297. doi:10.1109/3468.844354

Pashler, H. (1994). Overlapping mental operations in serial performance with preview. *The Quarterly Journal of Experimental Psychology Section A*, *47*(1), 161–191. doi:10.1080/14640749408401148

Payre, W., Cestac, J., & Delhomme, P. (2016). Fully automated driving. *Human Factors: the Journal of the Human Factors and Ergonomics Society*, *58*(2), 229–241. doi:10.1177/0018720815612319

Ranney, T. A., Harbluk, J. L., & Noy, Y. I. (2005). Effects of voice technology on test track driving performance: Implications for driver distraction. *Human Factors: the Journal of the Human Factors and Ergonomics Society*, *47*(2), 439–454. doi:10.1518/0018720054679515

Reichenbach, J., Onnasch, L., & Manzey, D. (2010). Misuse of automation: The impact of system experience on complacency and automation bias in interaction with automated aids. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *54*, 374–378. doi:10.1177/154193121005400422

Rossi, R., Gastaldi, M., Biondi, F., & Mulatti, C. (2012). Evaluating the impact of processing spoken words on driving. *Transportation Research Record: Journal of the Transportation Research Board*, *2321*(1), 66–72. doi:10.3141/2321-09

Rothenbücher, D., Li, J., Sirkin, D., Mok, B., & Ju, W. (2016, August). *Ghost driver: A field study investigating the interaction between pedestrians and driverless vehicles*. Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on (pp. 795–802). IEEE, Columbia University, NY.

SAE. (2018). *J3016_201806. Taxonomy and .definitions for terms related to driving automation systems for on-road motor vehicles*. Warrendale, PA: SAE International. doi:10.4271/2012-01-0107

Sarter, N. B., & Woods, D. D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors: the Journal of the Human Factors and Ergonomics Society*, *37*(1), 5–19. doi:10.1518/001872095779049516

Saunders, J., Syrdal, D., & Koay, K. (2016). "Teach Me–Show Me"—end-user personalization of a smart home and companion robot. *IEEE Transactions on Human-Machine Systems*, *46*(1), 27–40. Retrieved from http://ieeexplore.ieee.org/abstract/document/7322264/

Scholtz, J. (2003). *Theory and evaluation of human robot interactions*. System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on (pp. 10–pp). IEEE, Honolulu, HI.

Sellner, B., Heger, F., & Hiatt, L. (2006). Coordinated multiagent teams and sliding autonomy for large-scale assembly. *Proceedings of the IEEE*, *94*(7), 1425–1444. Retrieved from http://ieeexplore.ieee.org/abstract/document/1677954/

Sheridan, T., & Verplank, W. (1978). *Human and computer control of undersea teleoperators*. Retrieved from http://www.dtic.mil/docs/citations/ADA057655

Shinohara, Y., Currano, R., Ju, W., & Nishizaki, Y. (2017). *Visual attention during simulated autonomous driving in the US and Japan*. Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Oldenburg, Germany. Retrieved from https://dl.acm.org/citation.cfm?id=3122991

Shu, T., Gao, X., Ryoo, M., & Zhu, S. (2017). *Learning social affordance grammar from videos: Transferring human interactions to human-robot interactions*. IEEE International Conference on Robotics and Automation (ICRA) (pp. 1669–1676). Marina Bay, Singapore. Retrieved from http://ieeexplore.ieee.org/abstract/document/7989197/

Siripanich. (2017). Crossing the road in the world of autonomous cars. VR prototyping for communicating autonomous vehicle intent to pedestrians. Medium. Retrieved from https://medium.com/teague-labs/crossing-the-road-in-the-world-of-autonomous-cars-e14827bfa301

Stanton, N. A., Young, M., & McCaulder, B. (1997). Drive-by-wire: The case of driver workload and reclaiming control with adaptive cruise control. *Safety Science*, *27*, 149–159. doi:10.1016/S0925-7535(97)00054-4

Stanton, N. A., & Young, M. S. (2005). Driver behaviour with adaptive cruise control. *Ergonomics*, *48*(10), 1294–1313. doi:10.1080/00140130500252990

Strayer, D., Cooper, J., Turrill, J., Coleman, J. R., & Hopman, R. J. (2015). *Measuring cognitive distraction in the automobile III: A comparison of ten 2015 in-vehicle information systems. (October 2015)*. Retrieved from http://trid.trb.org/view.aspx?id=1252566

Strayer, D. L., Biondi, F., & Cooper, J. M. (2016). *Dynamic workload fluctuations in driver/non-driver conversational dyads* (pp. 362–367). Iowa Research Online.

Strayer, D. L., Cooper, J. M., McCarty, M. M., Getty, D. J., Wheatley, C. L., Motzkus, C. J., … Biondi, F. N. (2018). *Visual and cognitive demands of using Apple's CarPlay, Google's Android Auto and five different OEM infotainment systems*. Retrieved from https://

aaafoundation.org/wp-content/uploads/2018/06/AAA-Phase-6-CarPlay-Android-Auto-FINAL.pdf

Strayer, D. L., & Drews, F. (2007). Cell-phone? Induced driver distraction. *Current Directions in Psychological Science*, *16*(3), 128–131. doi:10.1111/j.1467-8721.2007.00489.x

Strayer, D. L., & Fisher, D. L. (2016). SPIDER. *Human Factors: the Journal of the Human Factors and Ergonomics Society*, *58*(1), 5–12. doi:10.1177/0018720815619074

Strayer, D. L., Getty, D., Biondi, F. N., & Cooper, J. (in press). *The multitasking motorist and the attention economy*. Washington DC: APA Press.

Strayer, D. L., Watson, J. M., & Drews, F. A. (2011). *Cognitive distraction while multitasking in the automobile*. Psychology of Learning and Motivation Volume 54 (Vol.54, pp. 29–58). Burlington, MA. doi:10.1016/B978-0-12-385527-5.00002-4

Sullivan, J. M., Flannagan, M. J., Pradhan, A. K., & Bao, S. (2016). *Literature review of behavioral adaptations to advanced driver assistance systems*. AAA Foundation for Traffic Research. Washington DC: AAA Foundation for Traffic Research.

Taylor, T., Pradhan, A. K., Divekar, G., Romoser, M., Muttart, J., Gomez, R., … Fisher, D. L. (2013). The view from the road: The contribution of on-road glance-monitoring technologies to understanding driver behavior. *Accident Analysis & Prevention*. doi:10.1016/j.aap.2013.02.008

Templeton, B. (2015). *A critique to NHTSA and SAE "Levels" of self-driving*. Retrieved from https://www.templetons.com/brad/robocars/levels.html

Tombu, M., & Jolicœur, P. (2003). A central capacity sharing model of dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(1), 3–18. doi:10.1037/0096-1523.29.1.3

Vanderhaegen, F. (2012). Cooperation and learning to increase the autonomy of ADAS. *Cognition, Technology & Work*, *14*(1), 61–69. doi:10.1007/s10111-011-0196-1

Vollrath, M., Schleicher, S., & Gelau, C. (2011). The influence of cruise control and adaptive cruise control on driving behaviour – A driving simulator study. *Accident Analysis and Prevention*, *43*(3), 1134–1139. doi:10.1016/j.aap.2010.12.023

Walch, M., Lange, K., Baumann, M., & Weber, M. (2015). *Autonomous driving: Investigating the feasibility of car-driver handover assistance*. Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 11–18). New York, NY: ACM.

Waytz, A., Heafner, J., & Epley, N. (2014). The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology*, *52*, 113–117. Retrieved from https://www.sciencedirect.com/science/article/pii/S0022103114000067

Wickens, C. (2008). Multiple resources and mental workload. *Human Factors: the Journal of the Human Factors and Ergonomics Society*, *50* (3), 449–455. doi:10.1518/001872008X288394

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit- formation. *Journal of Comparative Neurology of Psychology*, *18*, 459–482. doi:10.1002/cne.920180503

Złotowski, J., Yogeeswaran, K., Human, C., & Bartneck, C. (2017). Can we control it? Autonomous robots threaten human identity, uniqueness, safety, and resources. *International Journal of Human Computer Studies*, *100*(4), 48–54. Retrieved from https://www.sciencedirect.com/science/article/pii/S1071581916301768

About the Authors

Francesco Biondi is an Assistant Professor at the University of Windsor. He received his Ph.D. from the University of Padova, Italy, in 2015. He worked at Jaguar Land Rover as HMI Engineer, and as Research Scientist at the University of Utah.

Ignacio Alvarez is senior research scientist at the Autonomous Driving Research Lab in Intel Labs. Previously Ignacio worked for BMW in HMI & Telematics. He received his PhD in Computer Science from University of the Basque Country and Clemson University (USA) and B.S. in Communication Sciences from Burgos University (Spain).

Kyeong-Ah Jeong is senior human-factors engineer | user-experience researcher, Core & Visual Computing Group, Intel Corporation. Throughout her academic and industry careers, she focused on human-factors/usability engineering and human–computer interaction in software, automated driving, and smart home areas. She received her PhD in Industrial Engineering, Purdue University, West Lafayette, IN, USA.