Digital Society (2023) 2:41
https://doi.org/10.1007/s44206-023-00066-y

ORIGINAL PAPER

Driving into the Loop: Mapping Automation Bias and Liability Issues for Advanced Driver Assistance Systems

Katie Szilagyi[1] · Jason Millar[2] · AJung Moon[3,4] · Shalaleh Rismani[3]

Received: 22 April 2022 / Accepted: 29 August 2023 / Published online: 7 October 2023
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2023

Abstract
Advanced driver assistance systems (ADAS) are transforming the modern driving experience. Today's vehicles seem better equipped than ever to augment safety by automating routine driving activities. The assumption appears straightforward: automation will necessarily improve road safety because automation replaces the human driver, thereby reducing human driving errors. But is this truly a straightforward assumption? In our contention, this assumption has potentially dangerous limits. This paper explores how well-understood and well-researched psychological and cognitive phenomena pertaining to human interaction with automation should not be properly labelled as misuse. Framing the problem through an automation bias lens, we argue that such so-called instances of misuse can instead be seen as predictable by-products of specific engineering design choices. We engage empirical data to problematize the assumption that automating driving functions directly leads to increased safety. Our conclusion calls for more transparent testing and safety data on the part of manufacturers, for updated notions of misuse in legal contexts, and for updated driver training regimes.

Keywords Advanced driver assistance systems (ADAS) · Automation bias · Automation complacency · Liability · Human-automation interaction (HAI)

1 Introduction

Advanced driver assistance systems (ADAS) are transforming the modern driving experience. Whether through adaptive cruise control or lane keeping assist, high-tech automation systems augment today's drivers with promises to prevent blunders

✉ Katie Szilagyi
Katie.Szilagyi@umanitoba.ca

[1] Faculty of Law, University of Manitoba, Winnipeg, MB, Canada

[2] Faculty of Engineering, University of Ottawa, Ottawa, ON, Canada

[3] Faculty of Engineering, McGill University, Montréal, QC, Canada

[4] Dept. of Electrical and Computer Engineering, McGill University, Montreal, Canada

and achieve tricky tasks. Sirens sound to ward off driver errors, like drifting into an oncoming lane or rolling forward before a stoplight has turned green. With the help of ADAS, today's cars seem better equipped than ever before to improve safety by automating routine driving activities. The assumption appears straightforward: automation will necessarily improve road safety because automation relieves the human driver from certain driving tasks, thereby reducing human driving errors. But is this a safe assumption? In this paper, we argue that this assumption has potentially dangerous limits, especially in the context of next-generation ADAS, which raise important liability issues for drivers and automotive manufacturers. We ask: how do automated driving systems like ADAS force us to rethink the allocation of liability?

In this paper, we analyze the impact of ADAS on the law through an automation bias lens. To us, an automation bias lens means that we approach questions regarding appropriate and effective ADAS operation with special attention to the comprehensive evidence that humans experience automation bias and automation complacency while operating automated systems. Both human factors and psychology literatures, described below, offer compelling evidence that automated systems robustly induce vexing phenomena like automation bias and automation complacency in human operators. These induced effects must be taken into account when we draw the lines between proper and improper uses of an automated system. Drawing these lines is necessary for making sense of the expectations on human drivers. Without a reconceptualization of these concepts, traditional approaches to liability will render ADAS users liable, likely through the tort of negligence, in inappropriate scenarios.

We argue it is reasonable to predict that drivers who are operating ADAS—i.e., increasingly sophisticated automation systems—will demonstrate problematic patterns of behavior consistent with the automation bias and complacency literature. Those effects can include over-trust in automation systems and reduced attention to the driving task, which can understandably lead to accidents. We further argue that in cases where there is evidence, those effects were a contributing factor. Traditional boundaries between proper and improper driving behavior will shift as a result of those effects. To put it bluntly, in such cases, responsibility for resulting accidents shifts away from the driver. Indeed, through the lens of automation bias, these behaviors might more appropriately be framed as predictable by-products of specific engineering design choices, coupled with a lack of appropriate driver training on the effective use of such systems. Our conclusion calls for more transparent testing and safety data on the part of manufacturers, for rethinking notions related to liability in legal contexts, and for updated driver training regimes.

This paper builds upon previous scholarship on automated vehicle systems, asking specific questions about how regulatory regimes should respond to ADAS. Some engineering scholars have been investigating the cognitive challenges that drivers face when ADAS are relied upon as a last line of defense within partially automated driving systems. Our interdisciplinary research contributes to this ongoing effort, drawing from engineering, robot ethics, philosophy, and law to address the legal and policy gaps that exist when responding to the emerging technical complexity in this space. Through our automation bias lens, we engage empirical data to problematize the assumption that automating driving functions directly leads to increased safety.

Some of the empirical research that we draw upon is part of a 5-year (2016–2021) joint project by the Korean Transportation Institute (KOTI) and the Open Roboethics Institute (ORI); we engage with findings from an online survey on drivers' attitudes, training, and familiarity with automated systems (Le Page et al., 2021). By focusing on responsible innovation and getting automation questions right from the outset, we offer policy recommendations for the future of automated mobility.

Taking a whole system approach, making legal sense of new ADAS features requires looking critically at human interaction with ADAS to get a deeper understanding of their impact on legal and regulatory framings. In Section 2, we explain the technical aspects of three common assistive driving technologies and show how they are best situated in the litany of automation concerns documented by the current SAE framework. We elaborate on how existing empirical research helps support our work. In Section 3, we offer a theoretical framing outlining the key concerns that arise when untrained people take control of dangerous automated machinery. This section considers well-studied psychological phenomena, such as robustly observed driver complacency that is induced when humans shift into an operator mode when interacting with partially automated systems. In Section 4, we explore regulatory and legal challenges for the allocation of legal risk created by ADAS, informed by the Boeing 737 MAX crashes as a modern example of the complicated psychology that characterizes driver-ADAS hybrid systems. Finally, we offer policy recommendations for how to better integrate assistive driving technologies with today's driver regulation schemes.

2 Technical Grounding: How ADAS Works

2.1 Automation Overview: SAE Levels and ADAS Features

In the automated driving space, the SAE automated driving levels are used as the de facto standard for categorizing degrees of automation enabled by ADAS tools. Compared to the crude dichotomy between human-driven cars and self-driving cars, the SAE taxonomy offers commentators specificity in describing and developing ADAS technologies across a spectrum of automation (SAE, 2021). Six levels (0 to 5, where 0 indicates no automation) of automation are specified. Ethical analyses of automated driving tend to discount the lower SAE levels, level 1 (driver assistance) and level 2 (partial automation), on the assumption that the human driver remains unproblematically responsible for the most critical driving tasks with those low-level systems (Transport Canada, 2019). Instead, scholars have focused more on higher degrees of automation, usually beginning at level 3 (conditional automation), where drivers are relegated to a monitoring role to ensure the safe operation of automated systems (Transport Canada, 2019). Most ethical and legal analyses are aimed at the top two tiers, level 4 (high automation) and level 5 (full automation) (Goodall, 2021; Jaques, 2019), even though such vehicles are not yet commercially available (Transport Canada, 2019). Meanwhile, cars with ADAS at SAE levels 1–2 are increasingly the norm in today's automotive industry as major manufacturers focus on such technologies, incorporating them into all new vehicles. However, driver awareness of

the availability of such features, as well as associated formal training on how to use them, remains low. One recent study reports that only 50% of survey respondents were aware of their car's ADAS features before test driving the car, casting doubt on the extent of consumer preference for ADAS (Belron Canada, 2019).

While the ethical risks posed by level 4 (high automation) and level 5 (full automation) vehicles are intellectually engaging, directing critical analyses to the hypothetical and farther-future technologies can risk conflating them with—or altogether overlooking—the current, existing issues that lower-level automation systems in today's vehicles raise. Focusing on higher-level automation technologies employs what is sometimes known as the android fallacy—allowing science fiction conceptualizations of imagined futures to displace the reality of existing problems (Richards & Smart, 2016). To better illustrate the capabilities of these lower SAE level technologies, we focus on three examples, each categorized as SAE level 1 or 2: adaptive cruise control (ACC), lane keeping assist (LKA), and parking assist. We then offer some documented challenges with implementing ADAS technologies on the road.

2.1.1 Adaptive Cruise Control

Most commercial ACC technologies use radar to ascertain the relative speed difference between the ego vehicle (i.e., the vehicle in which the ACC is installed) and a target vehicle ahead of it on the roadway. The ACC system controls both throttle and brake to alter the speed of the ego vehicle, then the ACC system maintains a particular following distance between the ego vehicle and the target vehicle, or a maximum speed while no target vehicle is present. The following distance is based on a combination of driver input and sensor inputs, sometimes including weather conditions and road conditions. Sophisticated systems will respond spontaneously to situational changes, offering visual and acoustic alerts. For example, a BMW vehicle equipped with ACC will alert the driver when the target vehicle brakes suddenly (BMW, 2021). Drivers are given an opportunity to respond to the alerts by braking, but if they do not, the vehicle will perform the braking function automatically. ACC is usually categorized as SAE level 1, because it implicates braking and acceleration but not steering.

2.1.2 Lane Keeping Assist

LKA is an active ADAS that assists the driver when the vehicle begins to depart the traffic lane. Most systems use a forward-looking camera—usually mounted on top of the rear-view mirror—to detect the lanes on the road, warn the driver, and actively take control of the steering wheel if the car determines it is departing the lane. Some, but not all, systems also assist with keeping the vehicle centered within the lane. Most LKA systems provide support for the drivers in one of two ways: either by actively providing feedback to the steering wheel to steer the car toward the middle of the lane or by giving feedback only if the driver steers the vehicle toward the edges of the detected lanes. LKA relies on clearly visible standard lane markings and good road and visibility conditions to function properly; it also requires that the driver always pay attention to both the road and to ADAS. Absent clear conditions,

the LKA system sensing capabilities are often degraded. Most LKA systems also require a certain minimum speed to operate. The driver is always able to override the system simply by steering; the system will also often be temporarily deactivated while the driver depresses the brake pedal or activates a turn signal (Bekiaris et al., 2011). LKA is categorized as SAE level 1, because it offers steering support.

2.1.3 Parking Assist

Automated parking assist features are being implemented by major car manufacturers. In some implementations, parking is controlled manually by the driver, but the driver receives information and control prompts based on numerous ultrasonic sensors, radar, and video cameras (Wang et al., 2014). Other systems offer a fully automated approach, with the system taking control of the steering and power system to perform the parking maneuver based on the real-time sensor data, and the human can override as needed (Wang et al., 2014). Both approaches are being deployed, and both varieties are implicated when drivers refer to parking assist systems. Recent trends demonstrate a shift to fully automated parking systems. For example, Lexus' LS sedan series (equipped with intelligent parking assist) offers an automated parking system where the driver need only set the reverse gear and the parking assist uses ultrasonic sensors and cameras to park (Wang et al., 2014). Similarly, Volkswagen offers the park assist vision, the first demonstration of a fully automatic parking system, while Audi is developing an automated parking feature that connects to a driver's cellphone (Wang et al., 2014). Simpler parking assist features are categorized as SAE level 1, as the vehicle controls steering while the driver maintains the speed. As manufacturers move to more automated parking systems, these rise to SAE level 2, as the vehicle controls both steering and braking/acceleration simultaneously.

An example of fully automated parking is the Autopark feature in the Tesla Model X, where a combination of sensors and cameras allow the system to control the steering wheel (Bonnet et al., 2013). This allows the car to back autonomously into a parking spot (Tenhundfeld et al., 2020). Due to the complexity of the parking task, drivers tend to require heightened parking skills in situations with increased numbers of cars and fewer parking spaces (Wang et al., 2014).

2.1.4 ADAS Challenges

Studies have documented different ADAS systems' relative success at communicating information to drivers. One of the key considerations is variations between makes, models, and classes of vehicles: drivers familiar with ACC were more easily confused when interacting with conventional cruise control systems (McDonald et al., 2018). Additionally, ACC is unpopular for use in urban and even casual highway driving, since drivers tend to extrapolate from their previous experiences with conventional cruise control, which is typically used primarily for long-distance highway driving. Other users imposed unrealistic expectations upon ADAS, assuming they would be able to handle most aspects of driving. Even though ACC is not universally capable of automatically braking to avoid crash scenarios, one-third of

drivers surveyed believed that it could replace human braking altogether, raising flags for further evaluations of trust and reliance, and ultimately, safety (McDonald et al., 2018).

Furthermore, ADAS function best in ideal circumstances, with limits imposed when operational sensors are dirty, blocked, or otherwise marred by inclement weather. Many LKA systems will not operate in situations when the turning signal is activated and the vehicle is being steered in the direction of the turn signal (McDonald et al., 2018). However, the literature suggests many drivers remain unaware of this limitation, pointing to issues with identifying the environments and conditions under which the system works best (Bekiaris et al., 2011). The Autopark feature in Tesla's Model X is not functional in all circumstances; it is limited to certain types of parking spaces and conditions where the corresponding sensors can obtain sufficient information from its surroundings. Tight parking situations in private car parks do not play well with Autopark (Bonnet et al., 2013). Autopark use may be associated with risking damage to the car or bodily harm, especially in situations of "miscalibrated trust" in the automation (Tenhundfeld et al., 2020).

Most manufacturers avoid describing ADAS as safety features, preferring to couch them in terms of convenience in both marketing materials and owners' manuals (Lee et al., 2022). Scant manufacturer data is available to support conclusions regarding whether ADAS promotes safer outcomes. Appropriately designed and deployed automated parking systems do claim to increase driver safety by improving comfort levels and reducing fatigue, touted as making the overall driving experience more enjoyable (Wang et al., 2014). Other data promotes ADAS features as reducing the possibility of deaths from passenger vehicle collisions (Belron Canada, 2019). Regardless, common messaging around ADAS as safety enhancing is complicated by findings from survey research done by Le Page et al. as part of the KOTI/ORI project, which suggests significant discomfort around ADAS' implementation.

2.2 KOTI/ORI Report Summary

Quantitative public surveys on knowledge regarding ADAS can offer meaningful engagement with the deployment of these technologies and corresponding public attitudes. Le Page et al. (2021) conducted an online gender and age balanced survey of 1018 vehicle owners in June 2020 with consenting participants. Importantly, the authors note that vehicle default settings impact their usage rates: LKAs and parking assist functions are often on by default, different from traditional cruise control systems that require drivers to manually engage the system. As such, the survey authors report that drivers are most likely to use LKA (71.2%), followed by parking assist (63.2%), ACC (45.1%), and regular cruise control (36.6%). These results mirror how the technologies are deployed within new vehicles, underscoring how driver familiarity with system architectures and options are critical for interacting with ADAS—even for turning off undesirable system elements.

Across the range of ADAS technologies, Le Page et al. (2021) explain that a significant proportion of drivers reported feeling more unsafe when SAE level 1 or 2 systems are engaged, as compared with level 0 systems. Parking assist systems

received the most criticism, with 40.7% of respondents reporting feeling as though they compromised driver safety "most or all of the time." Meanwhile, 37% of respondents felt safety was compromised either most or all of the time with LKA, with 30.4% of respondents feeling this way for ACC. These reports are contrasted against traditional cruise control: only 3.5% of respondents report feeling cruise control compromised their safety "every time," with another 13% saying it did so most of the time. Conversely, 41% of respondents state that traditional cruise control "rarely" compromised their safety with 37% saying it "never" did so. Although one's perception of safety is not necessarily an accurate measure of one's actual safety, these reports on drivers' self-reported experiences underscore our concerns in this paper that ADAS systems do not necessarily improve vehicle safety. Other studies attempting to quantify ADAS safety impacts have generated mixed results. Vehicles operating at high levels of automation (level 4) are struck from behind in crashes more often than human-operated vehicles (Goodall, 2021). While this study focused on level 4 automation, which exceeds ADAS capabilities, it remains useful as a comparator for how automation increases certain types of accidents (Goodall, 2021). Conversely, LKA technologies have been found to reduce severity of police-reported crashes by 18% (as compared to vehicles without LKA) (Cicchino, 2019).

Drivers' safety concerns with ADAS signal the need for more transparency of manufacturer's data, research into safety, and driver education and training. When asked about what kinds of training or learning contributed to their knowledge of ADAS, respondents were clear that training efforts were meager and variable. Le Page et al. (2021) report that only 7.7% of respondents confirmed that their driving courses offered some degree of ADAS training, with another 5.8% reporting that the driver licencing authority offered basic information. Meanwhile, most respondents obtained their information either from reading the car manual (37.4%) or trying the feature out while driving (37.9%). Neither of these methods qualify as meaningful training experiences, nor are they likely to counteract the psychological effects described in HAI literature—automation bias and complacency—that undermine effective user behavior.

Even at low SAE automation levels, ADAS technologies fundamentally transform the nature of driving. ADAS shifts drivers away from directly performing critical driving tasks, instead asking them to monitor one or more automated systems now performing those safety–critical tasks and potentially to intervene immediately when something goes wrong. Thus, drivers and vehicles are transformed into hybrid *driver-ADAS systems*, with the driver and ADAS each performing overlapping aspects of safety–critical driving tasks. Literature exploring whether drivers understand new ADAS technologies has confirmed key issues with their deployment: misperceptions, lack of awareness, uncertainty, and confusion all plague the smooth adoption of ADAS, in addition to raising safety concerns (McDonald et al., 2018). Drivers have struggled to make sense of how some systems work in combination with one another and how various ADAS systems differ. As drivers are increasingly thrust into the role of systems monitor, their tasks behind the wheel expand to ensuring that the system behavior adheres to their expectations for safe driving. Drivers now share responsibility with automated systems for critical driving tasks, yet the automated driving paradigm does not yet seem to have incorporated important

wisdom from critical literature, specifically in related fields such as human factors engineering and ergonomics. With technical capabilities and survey results in mind, we move on to consider the psychological states giving rise to automation bias as a critical lens for conceptualizing this issue.

3 Human-Automation Interaction: Cataloging Concerns

Human factors engineering and ergonomics literatures have both documented the dominant phenomena of human-automation interaction (HAI), which includes the more targeted categories of human–computer interaction (HCI) and human–robot interaction (HRI). In these fields, two closely related terms that are critical to this analysis are used to describe problematic behaviors arising from human interactions with automation: automation bias and complacency. Automation complacency arises when system operators perform multiple tasks simultaneously, increasingly relying on automated decision aids instead of their own instincts (Parasuraman & Manzey, 2010; Wickens et al., 2015). Automation bias arises when system operators rely on imperfect automated decision aids, creating both omission and commission errors due to improper reliance (Parasuraman & Manzey, 2010). Automation complacency and bias present overlapping challenges: they both relate to how humans allocate their attention; they occur in both sophisticated and naïve participants (i.e., well-trained and untrained); and they cannot be easily overcome (Parasuraman & Manzey, 2010). Concurrently, issues of over-trust can result, as system operators become overly reliant on automated decision aids.

To better contextualize the issues faced by human operators in highly automated contexts, we begin this section with the cautionary tale of the twin crashes of the Boeing 737 MAX, which grounded Boeing's fleet of these aircraft in 2019. We explain the contextual circumstances that led to these catastrophic events and the nuanced HAI concerns that exacerbated exigent technical problems.

3.1 The Boeing 737 MAX Example

On March 9, 2019, Boeing 737 MAX aircraft were grounded worldwide after Ethiopian Air Flight 302 from Addis Ababa to Nairobi crashed shortly after takeoff, killing all 157 passengers and crew. It was the second such crash within 6 months; it followed the similar crash of Lion Air Flight 610, a domestic Indonesian flight, which crashed into the Java Sea on October 28, 2018, killing all 189 passengers and crew. In both cases, an automated operating feature—the Maneuvering Characteristics Augmentation System (MCAS)—was directly implicated as a major contributing factor to the cause of the crash (MSCTI, 2020).

To assist with maintaining the Boeing 737 MAX aircraft's safe attitude, two Angle of Attack (AOA) sensors on the outside of the plane provide input data to the MCAS. The day before the Lion Air crash, a mechanic had replaced one of the AOA sensors with a refurbished model. Later that day, an erroneous reading

from the AOA sensor caused the MCAS to push the nose of the plane down, attempting to correct for the false reading. Pilots attempted to override the MCAS to stabilize the plane to no avail. Luckily, a deadhead pilot was aboard and familiar with the issue; he assisted the active pilots, instructing them on how to remove electrical power from the horizontal stabilizer that the MCAS used to maneuver the plane. However, following the flight, the solution was not properly documented and flagged: the same thing happened the next day. This time, the MCAS pressed the plane's nose down more than 20 times. Pilots struggled against the automated system but did not know how to disable the horizontal stabilizer. Eventually, Lion Air Flight 610 crashed into the Java Sea (MSCTI, 2020).

Less than 6 months later, the same issue occurred on Ethiopian Air Flight 302. These two losses in such close succession led to the grounding of the entire Boeing 737 MAX fleet worldwide and a thorough investigation into the nature of the two crashes (MSCTI, 2020). While faulty AOA data and its impact on the MCAS played a critical role in both situations, official statements from the Federal Aviation Administration (FAA) and Boeing placed significant blame on the pilots. Boeing seemed to expect pilots to have a thorough awareness of the MCAS' operation and be able to deal with its malfunction, even though the MCAS was not classified as a safety–critical system, and pilots were not necessarily trained on its operation. Even more problematically, Boeing had specifically declined to define MCAS as a "new" system, which would have prompted costly training and certification programs for pilots. Instead, Boeing obtained FAA approval to remove references to the MCAS from the operations manual, meaning pilots were prevented from knowing about MCAS' existence and potential issues. Concurrently, Boeing's internal test data showed a pilot taking more than 10 s to react to an inappropriate MCAS activation—to catastrophic ends. Instead of making note of this failure, however, Boeing simply hoped that other pilots would be able to respond more quickly and took no action.

The Boeing 737 MAX catastrophe took place in a highly regulated industry, where pilots obtain significant levels of training in order to obtain and maintain their professional licenses. Yet, incorrect assumptions about HAI, system design, and adequate training all contributed to fatal outcomes. Specifically, Boeing wanted pilots to be able to transition from a different aircraft, the 737 NG, without obtaining new experiential training on the 737 MAX. This undermined training ability and diminished development of safety requirements that would have provided the skill development needed to intervene in MCAS malfunction. Furthermore, HAI considerations challenge assumptions that pilots would be able to successfully navigate the multiple alerts they receive in the cockpit. While engineering design best practices stipulate that aircraft systems should be designed to minimize the consequences of human error (Yeh et al., 2016), Boeing did not consider the impact of multiple alerts and indicators on a pilot's ability to prioritize critical issues (NTSB, 2019).

3.2 Automation Bias as Theoretical Lens for Driver-ADAS Interaction

While airline pilots are highly specialized professionals, who receive significant training on HAI before operating updated systems, the everyday driver is regularly thrust into new HAI scenarios simply by getting behind the wheel of a new vehicle. Automation transforms how drivers allocate their attention resources; it creates new cognitive tasks and replaces some of the cognitive and physical aspects of driving (Yamani & Horrey, 2018). Because automated systems reduce a driver's degree of control over a system, drivers become more likely to exhibit behaviors that might be characterized as an improper use of that system, leading to accusations of driver negligence. Such improper uses might sometimes be colloquially referred to as "erroneous uses" or "misuses" (Bahner et al. 2008). Yet, we argue, systems that demand *too much* of a human driver by asking for rapt attention when automation robustly induces over-trust or cognitive overload need a different frame for analysis. Characterizing the scenario one of improper use, erroneous use, or misuse negates the psychological states at play, wrongly places the lion's share of blame on the driver, risking potential accusations of distracted driving.

To wit: the claim that a driver *should have* been paying more attention in a given hybrid driver-ADAS system is fundamentally different from the same claim made in a traditional driving context, in which no automation is present. The addition of automation requires one to consider the potential psychological effects automation injects into the driving context and to consider the resulting impact on liability. A driver-ADAS system that exhibits driving behaviors causally linked to automation bias or complacency does not in and of itself suggest the driver used any part of the system improperly; *such a failure may simply point to a malfunctioning system in toto.* We argue that some uses of driver-ADAS systems might sometimes *look like* a driver has misused that system, but these issues instead flow from the induced psychological states produced by the technology itself. To impose liability on an individual and/or insulate a manufacturer from liability in such instances is, we argue, inappropriate.

Psychological models differentiate between human cognitive tasks that proceed autonomously and those that require people's effortful attention to complete. This distinction can be simplified as being between System 1 cognition—operating automatically and quickly with little to no effort and no voluntary control—and System 2 cognition—operating more slowly and methodically via effortful activities including complex computations, which involves agency, choice, and concentration (Stanovich & West, 2000, cited in Kahneman, 2011). Driving is frequently a System 1 activity, particularly driving a familiar vehicle on an empty road or in familiar circumstances. In the science of human memory, a similar distinction is observed between activating the procedural memory versus the declarative memory (Xie et al., 2019). Tasks dependent on procedural memory proceed without conscious thought; in the context of driving, an example of a procedural memory task is pressing the brake pedal when seeing brake lights in front of you. By contrast, declarative memory requires specific recollections of procedures or facts, like the specific step-by-step engine shutoff procedure in an aircraft (Xie et al., 2019).

Because driving is largely a System 1 activity, its dominant behaviors are reflexive and unconscious. Studies have shown that drivers execute driving actions an average of 0.3 s before registering conscious awareness of those actions (Ito & Yamazaki, 2016). However, researchers have complicated the intersection between System 1 and 2 driving activities through the addition of "load tasks," such as adding two digits every 6 s or explaining what they had just heard playing on the radio (Ito & Yamazaki, 2016). Adding the System 2 load tasks into the driving context increased the activity of the brain's conscious area and decreased the speed of reflexive, unconscious activities. Categorizing spoken words while braking creates a bottleneck in central attention due to increased cognitive load (Rossi et al. 2012). Other research demonstrates that increased cognitive demand, like cell phone-induced distraction, decreases the completion time for assembly tasks (Biondi, 2021). Additionally, response time for System 2 tasks, like tapping a button in response to stimulus, is impaired by cognitive load (Nilsson et al. 2018).

Kahneman (2011) recounts the story of the "Invisible Gorilla" experiment by Chabris and Simons, in which participants were shown a film strip of a white-shirted team playing basketball against a black-shirted team, with the instruction to count the number of passes between the white-shirted players. Focused on the difficult and absorbing task, the participants tended to miss the 9-s appearance of a woman in a gorilla suit on the basketball court. This, Kahneman offers, showcases the interaction of Systems 1 and 2: while seeing is an unconscious System 1 response, it requires some available mental capacity to be allocated to it. With the participants absorbed in the System 2 activity of counting passes, they would both miss the gorilla's appearance and later be incredulous that they had done so. Not only were participants unaware of the "obvious," they also had poor appreciation of their own limitations.

Translated to ADAS, these observations about human psychology offer powerful baselines for how attention resources are allocated during driving. Since so much of the driving task is stored in procedural memory, drivers perform much of the driving task automatically and unconsciously. ADAS complicates System 1's control of the driving task, since introducing new inputs, signals, and alerts—monitoring activities—engage System 2 processes. The appropriate procedures for responding to a series of alerts often originate in declarative memory, since they require specific recollections about how to respond. In driver-ADAS systems, drivers engage with many new tasks and have less control over the execution of each individual task. This can lead to over-trust and over-reliance on automated systems, occurring at both the attention allocation and information processing levels (Wickens et al., 2015). As ADAS increasingly demands that operators engage with multiple tasks at once, drivers are forced to subconsciously strategize about how to best allocate their attention in response to increased cognitive load.

This increased cognitive task load results from drivers forced to monitor automation while also performing other manual tasks associated with driving. This new monitoring task can be particularly unfamiliar to experienced drivers, who expect driving to proceed differently. Earlier studies examining the cause and effects of improper use in the context of automation complacency measure complacency according to an operator's responsiveness toward the system, independent from

performance consequences. Studies found that higher levels of operator complacency increased the likelihood that an operator's performance would decrease (Bahner et al., 2008; Biassoni et al., 2016). Where automated decision aids (such as ADAS) are correct, complacency can be beneficial to the operator—the decision aid will perform well. Ironically, if the decision aid provides incorrect information—a commission error—complacency is likely to increase as the operator struggles to reconcile the bad information (a result of increased cognitive load), potentially making the operator appear less competent at the operational task. More recent literature looks closely at the relationship between automation complacency and automation bias. For example, Wickens et al. (2015) define complacency more narrowly as "the state of monitoring prior to an automation decision aid failure that expresses itself as a delayed or "guessing" response when the decision aid fails to function at all." They then define automation bias as the "expression of such poor monitoring, when the decision aid provides incorrect advice, by following that advice" (Wickens et al., 2015). Coupled with over-trust of the system and the burden imposed by increased cognitive loads, this can increase complacency (Yamani & Horrey, 2018). Increased complacency then reduces visual attention paid to the primary information sources feeding automation—which must be monitored to detect failures (Yamani & Horrey, 2018). While training in attention strategies was found to decrease complacency, simply adding more practice time had no effect on reducing complacency (Yamani & Horrey, 2018).

Literature distinguishes between two types of automated decision-making aids: alerts and recommendations (Yamani & Horrey, 2018). Relying on these aids can lead to two key types of errors: omission and commission errors (Wickens et al., 2015). Where there is an error of omission, the system fails to alert the operator to some important information. Where there is an error of commission, the system incorrectly alerts the operator of an issue that is not in fact important (Wickens et al., 2015). Automation bias yields particularly problematic results when the alerting system provides false alerts or incorrect recommendations (Sauer et al., 2016). By relying on automated decision aids, users' own decision-making capabilities are negated, amplifying any technological errors. This robust phenomenon is reported across sectors, such as in aviation (Mosier et al., 1998). Similarly, in shooter targeting simulations, military operators' decision-making accuracy declined as the level of automation increased (Rovira et al., 2007). Higher levels of automation reliability correspond with lower decision-making accuracy. In the context of ADAS, automation bias is especially likely when drivers use automatic parking assist systems, due to the reliance on alerting systems required (Yamani & Horrey, 2018). Drivers (or driver-ADAS systems in our preferred terminology) often fail to visually inspect the relevant parking space before selecting to park automatically (Yamani & Horrey, 2018).

Recent literature examining perceived trust and complacency in automated vehicle technology examines partially automated vehicles—SAE level 2 (partial automation)—that transform the driving task by placing the driver in a passive monitoring role (Yamani & Horrey, 2018). Wilson et al. (2020) measured (1) engagement, (2) workload, and (3) perceived safety before and after drivers tested a partially automated vehicle on a public highway, finding high levels of trust, and perceived ease

of use, usefulness, and safety. Although high satisfaction was reported with the technology itself, driver engagement was low, with many drivers feeling as though they were "switch[ing] off" (Wilson et al., 2020). These results are consistent with prior findings that over-trust in automated vehicle technology leads to lower driver engagement and increased distracted and inattentive behavior (Yamani & Horrey, 2018). Safety concerns arise from this driver disengagement, including mode confusion, in which drivers fail to realize whether they were actively operating the vehicle (and, correspondingly, engaging procedural memory), or passively monitoring its automated functions.

Over-trust is more likely to cause drivers to engage in non-driving related tasks, ignore automation system failures, and inhibit driver control over the vehicle (Wilson et al., 2020). Studies by Kundinger et al. (2019) highlight related safety concerns associated with SAE level 2 (partial automation) driving technology. They compared driver fatigue between manual and partially automated driving technology, finding that the SAE level 2 drivers using partially automated features demonstrated significantly higher levels of driver drowsiness than active drivers. Meanwhile, other studies have looked critically at the likelihood of successfully making SAE level 3 automation a commercial reality. The handoff problem—the process of handing control authority over to a driver when an automated system encounters a problem and must disengage—raises significant safety concerns (Inagaki & Sheridan, 2018). Testing on optimal designs for the requests-to-intervene (RTI) processes (i.e., the handoff) has revealed conceptual flaws, since the process of trading authority from computer to person is fraught with challenges, even in the cases of professional, highly trained operators such as pilots (Inagaki & Sheridan, 2018). Automation handoffs focused on an RTI expect that the driver will necessarily notice the alert and accept the trading of authority for vehicle control from the automation to the driver. If the driver is more dependable than the automation in emergency situations, the coupling of an RTI and automatic safety control is beneficial. However, if the driver is less dependable than the automation in emergency situations, handoffs can make a situation worse because they may cause the driver to perform poorly in emergencies. Some authors have attempted to map the complexity of take-over time between semi-automated systems and their human system monitors. In a study on take-over time in semi-autonomous vehicles, Zhang et al. (2019) observed a wide range of take-over times: from the moment of take-over stimulus to the moment of driver intervention, the take-over time ranged from 0.69 to 19.79 s, with a mean time of 2.72 s (Zhang et al., 2019). More urgent alerts led to faster take-over times; higher levels of automation led to longer take-over times.

Taken together, this research suggests that we should assume, as a matter of responsible deployment that drivers will over-trust ADAS and become complacent in the driving task, which introduces risks, perhaps significant risks. ADAS complicates driving by demanding that drivers simultaneously navigate automatic, unconscious System 1 tasks and negotiate specific, procedural System 2 tasks. The interplay between these systems generates a sophisticated handoff problem, where take-over time may be impacted by robustly induced psychological phenomena like automation complacency and automation bias. This creates a novel driving situation, where human monitors are asked to participate in a hybrid driver-ADAS system.

Corresponding ADAS risks—induced driver distraction, increased cognitive load, increased fatigue, and a decreased ability to respond to requests to intervene—prompt thorough considerations of what it means to say that a driver (or driver-ADAS system) improperly used a system. They also prompt us to examine the role of driver training and awareness in emerging ADAS-enhanced vehicles.

4 Driver Licencing Regimes and Liability Frameworks

ADAS integrates transformational technologies into the driving experience. As demonstrated above, the emergence of ever more sophisticated ADAS now asks drivers to simultaneously drive the vehicle and monitor complex automated systems. This technical arrangement shifts the driving context from one in which the driver is in direct and constant control of the vehicle's motion to one in which there is a hybrid driver-ADAS system. Flowing from this observation, one might expect a corresponding shift in driver licencing regimes to account for ADAS' emergence. Instead, our regulatory reality demands that drivers be able to effectively monitor and operate ADAS without first obtaining any formal training or licencing to do so. Premised on the notion that the driving task is relatively unchanged by ADAS, this paradigm problematically assumes that drivers are still in full control of critical driving tasks and that ADAS has not transformed critical driving tasks in significant or fundamental ways. With this foundation, we then discuss methods of allocating legal risk in light of ADAS. Informed by the Boeing 737 MAX crash reports, and the automation bias and related psychological literature, we acknowledge the importance of the regulatory sphere. Given the shortcomings of private law remedies, we emphasize the potential role for government regulation through robust driver training and licencing programs.

4.1 Allocating Legal Risk

Understanding liability in new technological environments can require teasing out various legal responsibilities for the different players involved, including appropriate government regulation, manufacturer/product liability, and individual operator liability. Where regulatory frameworks remain underdeveloped, general areas of private law like tort or contract typically expand to encapsulate legal risk in new technological spaces. While manufacturers might bear some liability, it is more likely that responsibility will be passed onto the ultimate consumers: the drivers. Framing the issue with tort law precepts suggests a framework where drivers are accused of negligent behaviors, which, we argue, sits uncomfortably with ADAS design features that induce specific psychological states. Approaching this expansion through our automation bias lens requires thinking critically about how legal terms used in liability might require some reshaping to reflect the HAI realities of hybrid driver-ADAS systems more fairly.

Because ADAS is typically presented as a convenience feature in vehicles, it might sometimes be characterized as a consumer product. Commercial legislative

frameworks provide standards for the merchantability of goods and product warranties, offering remedies against those responsible for consumer product defects. Although definitions vary across jurisdictions, the usual stipulation is that goods for a particular market will be reasonably fit and accepted for the general purposes they serve (*Sale of Goods Act*, 1990; *Uniform Commercial Code*, Article 2 (2002); Directive, 1985/374/EEC). However, such fitness standards rely heavily on the seller's judgment of their own goods, meaning consumers examining products for defects are rarely protected by statutory warranty claims. Moreover, the complexity of ADAS systems mean consumers have little insight into their appropriate operation or what might be within the realm of warranty protection, even if applicable. Taken together, these realities typically lead consumers to the domain of tort law to obtain remedies against manufacturers (O'Leary & Armfield, 2020).

The tort of negligence offers a possible avenue of recovery for technology-related injuries and damages, by imposing a *prima facie* duty of care on manufacturers. Designers, creators, and distributors can be held liable for harms resulting from their products, extending to both the public and their specific users (*Lambert v Lastoplex Chemicals Co*, 1972). In the traditional framework, manufacturer's liability proceeds from the assumption that manufacturers must discharge their duties to produce a safe product, but users are also responsible to use the product safely—i.e., within its range of recommended uses. Manufacturers typically owe a duty to warn to the ultimate consumer; however, the involvement of learned intermediaries can further complicate the relationship where highly technical products are intended to be used under expert supervision or where consumers will not realistically receive direct manufacturer warning before using the product (e.g., *Hollis v Dow Corning Corp*., 1995; *Reyes v Wyeth Laboratories*, 1974). Learned intermediaries are a major component of manufacturer's liability in both Canada and the USA, as well as other common law jurisdictions with a well-developed tort law framework (see, e.g., German Civil Code, 2002). In some situations, reliance on the expertise of a learned intermediary is an appropriate legal response. For example, the *Hollis v Dow Corning Corp* decision provides an example of how learned intermediaries complicate a manufacturer's duty to warn. The case concerned the risks present in breast augmentation surgery, where harm might flow from a silicone breast implant rupturing inside the patient's body. Yet, in this scenario, the plastic surgeon performing the augmentation receives the warning directly from the manufacturer. The patient receiving the implant would not, and could not be expected to, have direct contact with the manufacturer to learn about the risks—instead relying on the plastic surgeon to relay the information. In civil law jurisdictions, statutory mechanisms operate in the negligence space, adding definitions of tortious actions through their civil codes (see, e.g., Article 6:162 of the Dutch Civil Code (1992); see also Quebec's Civil Code (1991) Sections 1468 and 1473).

Technological development and tort liability often progress in tandem, with tort law's affordances shaping emerging technology (Leiman, 2021). While underdeveloped product liability regimes do not usually pose barriers to market introduction, consumer expectations can create a thorny landscape for measuring product defects. For ADAS, this means that confusion is likely: users are often unaware of how ADAS systems ought to perform under certain conditions (Brennan-Marquez

et al., 2019; Ruscio et al., 2015). Even if ADAS systems offer safer alternatives to sole human operation, manufacturers should not be altogether absolved of liability merely for introducing marginal safety improvements. Duties to warn ultimate consumers might be discharged by vehicle retailers, but it is unlikely that most vehicle-purchasing experiences would rise to the level of expertise anticipated by the involvement of learned intermediaries, where case law developed in the context of prescribing physicians and pharmaceutical manufacturers. It is instead likely that most instances of driver-ADAS system malfunction would be characterized as negligence. In such scenarios, any driver failures of attention would be characterized as failures to discharge a duty of care or comply with the appropriate standard of care.

Some critics might assert that our emphasis on driver training and licencing schemes could relieve the pressure on manufacturers to operate in accordance with the nuances of tort law; calling for reforms to driver training might make it seem that the blame for induced psychological HAI effects is a regulatory failure. From this view, the argument might proceed, it is better that manufacturers bear the lion's share of any allocated legal risk to mandate best practices in manufacturing ADAS-enhanced vehicles. However, through our automation bias lens, the roadworthiness of ADAS features seems to include an implied assumption of improved driver safety that ignores the conversion of driver into ADAS systems monitor within a hybrid driver-ADAS system and the corresponding psychological phenomena endemic to HAI. It thus obscures the conceptual impact that automation bias and complacency have on the notions of liability in driving. We argue that the possibility of falling back on traditional notions of driver negligence (those that evolved prior to the introduction of ADAS-like automation) in cases where automation bias or complacency can reasonably be seen as contributing factors will continue to insulate manufacturers, especially so long as training and licencing remain undeveloped.

In other contexts where automation interacts with operating danger, such as aviation, assuming that operators will necessarily be able to compensate for—and overcome—design flaws stemming from automated systems can yield catastrophic results. In the 737 MAX crashes, the Boeing CEO placed some blame on the pilots, noting that runaway stabilizer procedures had not been completely followed (Economy, 2019). After the Lion Air crash, the FAA's Associate Administrator for Aviation Safety called the incident a "one off," attributable to poor pilot performance (MSCTI, 2020). Of course, the Ethiopian Air crash occurred shortly thereafter, underscoring inquiries into whether some larger systemic issue was taking place.

Similar problems have arisen with the explanations following crashes of vehicles with higher levels of automation. Despite the sophistication of the hand-off problem in HAI, crashes that have taken place where vehicles are operating with high levels of autonomy have usually been blamed on the human driver monitoring the system (Wakabayashi, 2018), typically framed as negligence. This was the case when Tesla driver Joshua Brown, using the car's Autopilot feature (level 2 automation), suffered a fatal crash while allegedly viewing a Harry Potter movie (Levin & Woolf, 2016). Statements by the company following Brown's death noted that Autopilot, while impressive and constantly improving, was "not perfect and still requires the driver to remain alert" (Levin & Woolf, 2016). The same was true after a highly

autonomous Uber test vehicle (level 4 automation), being monitored by safety driver Rafaela Vasquez, struck and killed pedestrian Elaine Herzberg (Wakabayashi, 2018). Local police obtained warrants for streaming services Hulu, Netflix, and Google to determine if Vasquez had been distracted at the time, concluding that she had been watching singing competition show "The Voice" when the collision took place (Smiley, 2022).

A careful look at the impugned circumstances asks us to consider the extent to which each driver fell into highly predictable mental states. The psychological phenomena exhibited by the drivers in these circumstances are consistent with the psychological literature discussed above. This raises important questions about the extent to which it is fair, or factually accurate, to characterize these cases as examples of negligent behavior. While both Brown and Vasquez were surely distracted by watching movies while driving, the well-documented psychological effects of automation bias and/or complacency were contributing factors. Performing driving/monitoring tasks is necessarily complicated by the presence of advanced automation systems. Given the clear potential for driver distraction proposed in the literature, such activities are difficult to avoid; simply calling the drivers negligent does not consider the psychological phenomena at work in the background, precipitating their behaviors. Yet, autonomous vehicle manufacturers seem to demand that human monitors defy their own psychology to provide rigorous, continuous monitoring for autonomous vehicles. Drivers are told to pay careful attention by the very same parties that are dulling their attention via the robustly predictable effects of automation. In the Uber/Vasquez case, further inquiries revealed that Uber's self-driving system had never identified Ms Herzberg as a human (repeat analyses by Uber oscillated between car, bike, and other), and Uber's system had delayed the possibility of automated emergency braking to verify the emergency, avoid false alarms, and allow human takeover only if necessary (Smiley, 2022). After five years of uncertainty, Ms Vasquez pled guilty to one count of "endangerment" in July 2023 in a plea deal; she received three years of probation, but avoided a full trial and possible jail time (Smiley, 2023).

Almost no cases have been litigated in the driver-ADAS hybrid system space so far, leaving open questions as to how courts will allocate legal risk for the operation of ADAS. However, the case of *Justine Hsu v Tesla Inc*., decided in April 2023, offered the first legal test of Tesla's Autopilot system (level 2 automation). Ms Hsu was using Autopilot in stop-and-go traffic, traveling between 25 and 30 mph, when the vehicle swerved into the central median, hitting the curb and deploying the airbag. Being quite petite, Ms Hsu tended to sit quite close to the steering wheel, elbows bent, so she could see the road and reach the foot pedals. The airbag deployed with considerable force, causing numerous breaks in her jaw, damage to her teeth, bleeding from her hands and mouth, among other injuries (Roy et al, 2023a, b). Ms Hsu alleged that the Autopilot had malfunctioned and that the airbags should never have deployed. She made numerous claims against Tesla, including strict product liability, negligence, breach of implied/express warranty, and misrepresentation. A California jury denied Ms Hsu's claim, noting that Autopilot is only an aid, not a self-driving car, and that Ms Hsu was a distracted driver (*Hsu v Tesla*, 2023). Jurors later commented that Tesla was "adamant" about the assistive nature of the technology,

with explicit audio and visual warnings for drivers about their continued responsibility (Roy et al., 2023a). Yet, the record showed that Ms Hsu received Autopilot's notification to put her hands on the wheel less than a second before the crash occurred, certainly not enough time to transition from system monitor to active driver (Roy et al., 2023b). While this is only one case, the jury's reasoning privileges Tesla's technology and the warnings it provides over the well-documented psychological phenomena such technology induces.

Absent other germane case law, we can look for an analogy in a Canadian court's consideration of a dangerous driving violation, in which the driver was driving a manual transmission despite being unfamiliar with its operation. The defendant was charged with dangerous driving under Canada's Criminal Code; in addition to his lack of familiarity with the manual transmission, the evidence included that he was traveling 208 kmph (more than triple the speed limit) and was not a fully licensed driver (R v Routh, 2016). Extrapolating from this view, unfamiliarity with a vehicle's operations—including its technological systems—may eventually be cited in support of a liability claim. While the standard employed in a criminal case is different from that of tort law, the argument is transferable: it is foreseeable that unfamiliarity with ADAS functionality could be argued as a failure to discharge one's duty of care in a negligence claim. With such uncertainty remaining around the allocation of legal risk, robust regulatory frameworks could potentially assist with flagging the new complexity introduced through hybrid driver-ADAS systems.

4.2 Driver Licencing Approaches

Driver training and licencing regimes implicitly acknowledge the importance of proper driving knowledge, verification of that knowledge, especially safety–critical knowledge, and the potential dangers of an inexperienced or undereducated driver behind the wheel. Worldwide, governments take active roles in the licencing of drivers as an important part of public safety. Despite these regulatory powers, the changing face of vehicular technology has not prompted reconsideration of driving training or licencing, even though the sophistication of modern ADAS transforms the driving task. Research has demonstrated that training experience can influence both automation bias and the trust users place in a technology (Sauer et al., 2016).

While initial consultations have identified the problem, ADAS' transformational impact on the driving task is not yet reflected in driving licencing regimes or transportation liability regulations. This paper discusses the Canadian regulatory environment for driver licencing, much of which is transferable to driver licencing in other jurisdictions. In Canada, driver licencing is a provincial responsibility, although the federal regulator, Transport Canada, does offer some over-arching guidelines for the administration of the overall system. Under the *Guidelines for Testing of ADS Vehicles* in Canada, the government provides some recommendations on SAE International definitions for jurisdictions and manufacturers to follow (CCMTA, 2018). The Canadian Council of Motor Transport Administrators (CCMTA) acknowledges the paucity of ADAS information on provincial driver manuals and has stated the need to develop a training method for drivers and licencing examiners to keep pace with the technological evolution (CCMTA, 2018). Although this provides future

guidelines for SAE level 3–5 vehicles, the processes and methods pertaining to the current driver licencing regimes remain outdated. Approximately 15% or 3.7 million vehicles in Canada use ADAS technology bordering definitions of SAE level 2 and 3 guidelines (Belron Canada, 2019). Drivers need to better understand how little they know about their ADAS technology and should be tested to ensure their ability to effectively operate those technologies. Broad interpretations of SAE levels and ambiguous driver licencing rules require immediate attention and amendments.

Safety critical technologies such as backup cameras, lane departure warnings, and blind spot warnings are only sometimes permissible during driver testing. While the CCMTA recommends that safety technologies be engaged during the testing process, those few provinces that have explicitly incorporated the CCMTA's recommendations caution that drivers must still demonstrate traditional safety behaviors like blind spot checking and mirror usage. While driving candidates might be allowed to use some assistive technologies, the emphasis on acting as though the technology is absent privileges traditional driving behaviors, ignoring the environment in which most drivers will operate most of the time. It also underscores the problematic assumption that engaging an ADAS leaves the driving task unchanged, or simply increases driver safety, neither of which seems straightforward or accurate. Admittedly, variability in manufacturing across vehicles presents challenges for potential standardization in the formal testing environment. But failing to directly examine use of ADAS in the testing environment undermines the importance of drivers actually knowing how their ADAS technologies work and how to effectively operate them. Drivers are asked to suspend modern driving aids—and their own disbelief—to operate a vehicle as if it does not possess these transformative features. The current testing measures are therefore not conducive to teaching new drivers how best to effectively operate ADAS technologies.

Under the Motor Vehicle Safety Regulations (2009) (CMVSS 111 Rear Visibility Systems), all passenger cars manufactured after May 2018 are required to accommodate rear-view cameras—a necessary component for parking assist. However, only two provinces (Nova Scotia and Saskatchewan) require additional road training on their use. Although the content of driver training varies regionally, educators are not specifically required to teach about ADAS technology. As a result, the public is not provided with government-issued instructions on the limitations of ADAS technology or the safe driving behaviors when using ADAS features. Most owners do not read owners' manuals, relying instead on their own experiences (Lee et al., 2022). A survey by Belron Canada (2019) showed that 51% of Canadian drivers have little or no familiarity with ADAS as many were never offered any information or training by their vehicle dealers. Le Page et al. (2021) confirm the same. Other studies have shown dealerships, marketing, and salespeople actually cause confusion about the operation of ADAS technologies (Lee et al., 2022). In light of this, policy considerations for amending the current driver licencing regimes are required to educate future drivers in the use of this evolving technology, or impose regulations on manufacturers to provide transparent human factors safety testing data about the use and operational limitations of their ADAS. The CCMTA further recommended that ADAS information be included in future jurisdictional driver manuals and applied in licencing knowledge tests (CCMTA, 2018). Those using public roads need to

have confidence that (a) drivers are appropriately trained in the new hybrid driver-ADAS paradigm and that (b) the training ensures drivers can overcome the well-documented psychological effects induced by new vehicle automation systems.

ADAS operation should therefore be adopted into driver's education and training courses. In countries where driver training is locally administered, federal guidance should be issued on how to best achieve these goals. Proper guidance would include identifying key system components, distinguishing between alerts and recommendations, and appropriate interventions for both commission and omission errors. Training recommendations should incorporate psychological and ergonomics research to acknowledge levels of over-trust and overcompliance commonly associated with HAI, with specific warnings about common issues, like feeling "switched off" or driver drowsiness. Data from testing centers could be collected to develop insights for new frameworks for hand-off processes, which might better inform approaches for requests to intervene and take-over time. Such data could also inform regulators of the extent to which driving candidates seem able to overcome any automation-induced psychological effects. Care would need to be taken to ensure training centers and programs had access to vehicles equipped with commonly available ADAS.

In this scenario, ADAS-familiarity could also be acknowledged in formal testing environments, although since most driving candidates provide their own vehicle for driver testing, this may require a minimum ADAS requirement or the provision of a specific test car to achieve a driving credential. In the alternative, ADAS testing could be separately administered, with drivers mandated to complete a specific ADAS course within a particular time period of their other licencing exam. Governments would need to make the training centers accessible, with appropriate attention given to equity-seeking groups. In addition, standard organizations must help manufacturers coordinate the nature of user interfaces associated with various ADAS features, since a sense of interface uniformity from vehicle to vehicle will become more essential as more drivers engage in car sharing over car ownership.

5 Conclusion

To best approach ADAS driver training and regulation, policy makers must acknowledge the reality of automation bias and complacency, and jettison assumptions that introducing ADAS to consumer vehicles necessarily improves safety outcomes. New ADAS-inspired standards in driver licencing regimes with clear evidence-based policy might assist in recalibrating driver trust and reliance appropriately to avoid the automation bias and complacency outcomes identified in Section 2 and to explicitly and effectively address the risks and hazards associated with ADAS. Still, autonomic psychological phenomena impacting the driving task create ongoing challenges that mere training might not resolve. Clear direction to manufacturers through unified regulatory approaches that prioritize testing and safety data can help avoid the patchwork liability outcomes discussed in Section 4.

Focusing on SAE levels to impose liability thresholds is currently obscuring a critical distinction: what complicates liability is the distinction between drivers directly controlling critical driving tasks *versus* monitoring the automated performance of

critical driving tasks. Concurrent research about driver awareness, knowledge, and use of ADAS demonstrates significant shortcomings with their implementation. Literature demonstrates that drivers were unsure about how some systems work in combination with others and how they are distinguishable from one another. Uncertainty and confusion may impact a driver's usage, comfort, and reliance on the technology (Ruscio et al., 2015). Few survey respondents reported seeking information about technologies from any sources beyond the dealership, owner's manual, and their own experience via trial and error; only about 1 in 10 reported seeking information on the Internet and hardly any reported visiting government safety websites (Le Page et al., 2021). More research is needed to determine how best to reach drivers with important information about how to effectively operate ADAS.

Finally, more research is needed to determine how to best allocate legal risk in light of the psychological phenomena at play in HAI systems. Those phenomena suggest that the borders between proper and improper use, and what amounts to negligent use, differ between traditional vehicles and those implementing hybrid driver-ADAS systems. Fairness demands that we examine these issues to better characterize current ADAS and in anticipation of the much more complex automated driving technologies coming down the pike.

Acknowledgements The authors are grateful to the CRAiEDL team, Sophie Le Page, Jae Jin Lee, Julia Mogus, Ariadad Fattahyani, and Jenna Symons. This work benefitted enormously from the careful attention and deep technical expertise of our excellent research assistant, Tommy Friedlich, and from presentation at We Robot 2021, where discussant Meg Leta Jones and our robot law community offered helpful feedback. The research leading to these results received funding from the Korea Agency for Infrastructure Technology Advancement (KAIA).

Author Contribution All authors contributed to the study conception and design. Jason Millar and AJung Moon conceived and conducted the online survey used for data analysis, referenced in the bibliography as Le Page et al. (2021). The first draft of the manuscript was written by Katie Szilagyi, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding The research leading to these results received funding from the Korea Agency for Infrastructure Technology Advancement (KAIA).

Data Availability The datasets generated by the survey research analyzed by the current study are available online via the Open Roboethics Institute: https://openroboethics.org/public-opinion-surveys-on-self-driving-cars-a-collaboration-with-the-korea-transport-institute/.

Declarations

Ethics Approval The survey data reported in this paper, originally published by Le Page et al. (2021) received ethics approval from the Health Sciences and Science Research Ethics Board of the University of Ottawa (REB approval no. H-05–20-5780).

Conflict of Interest The authors declare no competing interests.

References

Bahner, J. E., Hüper, A.-D., & Manzey, D. (2008). Misuse of automated decision aids: Complacency, automation bias and the impact of training experience. *International Journal of Human-Computer Studies, 66*(9), 688–699. https://doi.org/10.1016/j.ijhcs.2008.06.001

Bekiaris, E., Wiethoff, M., & Gaitanidou, E. (2011). *Infrastructure and safety in a collaborative world* (1st ed.). Springer.

Belron Canada. (2019). *The hidden link between windshields and road safety: The role of advanced driver assistance systems (ADAS) in Canada's quest for the world's safest roads*. White Paper prepared by Belron Canada Inc.

Biassoni, F., Ruscio, D., & Ciceri, R. (2016). Limitations and automation. The role of information about device-specific features in ADAS acceptability. *Science Safety Direct, 85*, 179–186. https://doi.org/10.1016/j.ssci.2016.01.017

Biondi, F. (2021). Overloaded and at work: Investigating the effect of cognitive workload on assembly task performance. *Human Factors, 63*(5), 813–820. https://doi.org/10.1177/0018720820929928

BMW. (2021). *BMW driver assistance: Intelligent driving*. BMW. Retrieved November 17, 2021, from https://www.bmw.ca/en/topics/experience/connected-drive/bmw-connecteddrive-driver-assistance.html#

Bonnet, C., Hiller, A., Kuenzel, G., Moser, M., & Schiemenz, H. (2013). *Method for autonomous parking of a motor vehicle, driver assistance device for performing the method and motor vehicle with the driver assistance device* (Patent No. US20150088360A1).

Brennan-Marquez, K., Levy, K., & Susser, D. (2019). Strange loops: Apparent versus actual human involvement in automated decision making. *Berkeley Technology Law Journal, 34*(3), 745–772. https://doi.org/10.15779/Z385X25D2W

Canadian Council of Motor Transport Administrators (CCMTA), Automated Vehicles Working Group. (2018). Canadian jurisdictional guidelines for the safe testing and deployment of highly automated vehicles. American Association of Motor Vehicle Administrators. Retrieved January 30, 2023, from: https://www.ccmta.ca/web/default/files/PDF/CCMTA-AVGuidelines-sm.pdf

Cicchino, J. B. (2019). Real-world effects of rear automatic braking and other backing assistance systems. *Journal of Safety Research, 68*, 41–47. https://doi.org/10.1016/j.jsr.2018.12.005

Civil Code of Quebec, § 1468 (1991).

Directive 1985/ 374/ EEC. *The approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products.* European Parliament, Council of the European Union.

Dutch Civil Code: Book 6 The Law of Obligations, § Article 6:162 (1992).

Economy, P. (2019, April 30). *Boeing CEO puts partial blame on pilots of crashed 737 MAX aircraft for not "completely" following procedures*. Inc.Com. https://www.inc.com/peter-economy/boeing-ceo-puts-partial-blame-on-pilots-of-crashed-737-max-aircraft-for-not-completely-following-procedures.html

German Civil Code, BGB, (2002).

Goodall, N. J. (2021). Comparison of automated vehicle struck-from-behind crash rates with national rates using naturalistic data. *Accident Analysis & Prevention, 154*, 106056. https://doi.org/10.1016/j.aap.2021.106056

Hollis v Dow Corning Corp., [1995] 4 SCR 634.

Hsu, Justine v Tesla Inc., Case No. 20STCV18473 – Special Verdict Form (2023).

Inagaki, T., & Sheridan, T. B. (2018). A critique of the SAE conditional driving automation definition, and analyses of options for improvement. *Cognition, Technology & Work, 21*, 569–578. https://doi.org/10.1007/s10111-018-0471-5

Ito, T., & Yamazaki, K. (2016). A study of a driver model using unconscious driving behaviors. *Procedia Computer Science, 96*, 754–761. https://doi.org/10.1016/j.procs.2016.08.260

Jaques, A. E. (2019). *Why the moral machine is a monster*. Proceedings of the 2019 WeRobot Conference, Miami, FL. Retrieved January 22, 2021, from https://robots.law.miami.edu/2019/wp-content/uploads/2019/03/MoralMachineMonster.pdf

Kahneman, D. (2011). *Thinking, fast and slow*. Straus and Giroux.

Kundinger, T., Wintersberger, P., & Riener, A. (2019). (Over)Trust in automated driving: The sleeping pill of tomorrow? *CHI EA 2019: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems; Paper No. LBW2418*, 1–6. https://doi.org/10.1145/3290607.3312869

Lambert v Lastoplex Chemicals Co, [1972] SCR 569.

Lee, J. J., Patel, B., Millar, J., Friedlich, T., Kearney, A., & Mattesco, C. (2022). A legal textual analysis of vehicle owner's manuals: Gauging manufacturers' relative risk perception of advanced driver assistance systems in the absence of real-world data. *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society (AIES '22).*

Le Page, S., Millar, J., Bronson, K., Rismani, S., & Moon, A. (2021) Driver perceptions of advanced driver assistance and safety. Retrieved September 30, 2021, from https://arxiv.org/abs/1911.10920

Leiman, T. (2021). Law and tech collide: Foreseeability, reasonableness and advanced driver assistance systems. *Policy and Society, 40*(2), 250–271. https://doi.org/10.1080/14494035.2020.1787696

Levin, S., & Woolf, N. (2016, July 1). Tesla driver killed while using autopilot was watching Harry Potter, witness says. *The Guardian.* https://www.theguardian.com/technology/2016/jul/01/tesla-driver-killed-autopilot-self-driving-car-harry-potter

Majority Staff of the Committee on Transportation and Infrastructure (MSCTI). (2020). *Final Committee report: The design, development & certification of the Boeing 737 MAX*. The House Committee on Transportation & Infrastructure. Retrieved February 9, 2022, from https://www.corporatecrimereporter.com/wp-content/uploads/2020/11/Summary-Committee-Report-Boeing-737MAX-9-20.pdf

McDonald, A., Carney, C., & McGehee, D. V. (2018). Vehicle owners' experiences with and reactions to advanced driver assistance systems. AAA Foundation for Traffic Safety. Retrieved February 14, 2021, from https://aaafoundation.org/wp-content/uploads/2018/09/VehicleOwnersExperiencesWithADAS_TechnicalReport.pdf

Mosier, K. L., Skitka, L. J., Heers, S., & Burdick, M. (1998). Automation bias: Decision making and performance in high-tech cockpits. *The International Journal of Aviation Psychology, 8*(1), 47–63. https://doi.org/10.1207/s15327108ijap0801_3

National Transportation Safety Board (NTSB). (2019). Safety recommendation report: Assumptions used in the safety assessment process and the effects of multiple alerts and indications on pilot performance. Retrieved November 15, 2021, from https://www.ntsb.gov/investigations/accidentreports/reports/asr1901.pdf

Nilsson, E. J., Aust, M. L., Engstrom, J., Svanberg, B., & Linden, P. (2018). Effects of cognitive load on response time in an unexpected lead vehicle braking scenario and the detection response task (DRT). *Transportation Research Part F: Traffic Psychology and Behaviour, 59*(A), 463–474. https://doi.org/10.1016/j.trf.2018.09.026

O'Leary, T., & Armfield, T. (2020). Adapting to the digital transformation. *Alberta Law Review*, *58*(2), 249–272. https://albertalawreview.com/index.php/ALR/article/view/2618/2577

Parasuraman, R., & Manzey, D. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors the Journal of the Human Factors and Ergonomics Society, 52*(3), 381–410. https://doi.org/10.1177/0018720810376055

R v Routh, 2016 ABPC 26.

Richards, N., & Smart, W. D. (2016). How should the law think about robots*?*. In R. Calo, M. Froomkin, & I. Kerr (Eds.), *Robot Law* (pp. 3–24). Edward Elgar Publishing.

Reyes v. Wyeth Laboratories, 498 F.2d 1264 (5th Cir. 1974).

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.* https://doi.org/10.3141/2321-09

Rovira, E., Mcgarry, K., & Parasuraman, R. (2007). Effects of imperfect automation on decision making in a simulated command and control task. *Human Factors the Journal of the Human Factors and Ergonomics Society, 49*(1), 76–87. https://doi.org/10.1518/001872007779598082

Ruscio, D., Ciceri, M.-R., & Biassoni, F. (2015). How does a collision warning system shape driver's brake response time? The influence of expectancy and automation complacency on real-life emergency braking. *Accident Analysis and Prevention, 77*, 72–81. https://doi.org/10.1016/j.aap.2015.01.018

Roy, A., Levine, D., & Hyunjoo, J. (2023a, April 22). Tesla wins bellwether trial over Autopilot car crash. *Reuters.* https://www.reuters.com/legal/us-jury-set-decide-test-case-tesla-autopilot-crash-2023-04-21/3

Roy, A., Levine, D., & Hyunjoo, J. (2023b, April 22). Exclusive: Tesla's Autopilot never claimed to be self-pilot, juror says. *Reuters.* https://www.reuters.com/business/autos-transportation/teslas-autopilot-never-claimed-be-self-pilot-juror-2023-04-21/

Sale of Goods Act, R.S.O. 1990, c. S.1.

SAE. (2021, April 30). Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. SAE International. https://www.sae.org/standards/content/j3016_202104/

Sauer, J., Chavaillaz, A., & Wastell, D. (2016). Experience of automation failures in training: Effects on trust, automation bias, complacency and performance. *Ergonomics, 59*(6), 767–780. https://doi.org/10.1080/00140139.2015.1094577

Smiley, L. (2022, March 8). 'I'm the operator': The aftermath of a self-driving tragedy. *Wired.* https://www.wired.com/story/uber-self-driving-car-fatal-crash/

Smiley, L. (2023, July 28). The legal saga of Uber's fatal self-driving car crash is over. *Wired.* https://www.wired.com/story/ubers-fatal-self-driving-car-crash-saga-over-operator-avoids-prison

Tenhundfeld, N. L., de Visser, E. J., Ries, A. J., Finomore, V. S., & Tossell, C. C. (2020). Trust and distrust of automated parking in a Tesla Model X. *Human Factors: THe Journal of the Human Factors & Ergonomics Society, 62*(2), 194–210. https://doi.org/10.1177/0018720819865412

Transport Canada. (2019). Safety assessment for automated driving systems in Canada. *Transport Canada*. Retrieved January 30, 2021, from https://tc.canada.ca/sites/default/files/migrated/tc_safety_assessment_for_ads_s.pdf

Uniform Commercial Code, Article 2. Sales § 2–314(2)(c) (2002).

Wakabayashi, D. (2018, March 19). Self-driving Uber car kills pedestrian in Arizona, where robots roam. *The New York Times.* https://www.nytimes.com/2018/03/19/technology/uber-driverless-fatality.html

Wang, W., Song, Y., & Deng, H. (2014). Automatic parking of vehicles: A review of literatures. *International Journal of Automotive Technology, 15*, 967–978. https://doi.org/10.1007/s12239-014-0102-y

Wickens, C. D., Clegg, B. A., Vieane, A. Z., & Sebok, A. L. (2015). Complacency and automation bias in the use of imperfect automation. *Human Factors, 57*(5), 728–739. https://doi.org/10.1177/0018720815581940

Wilson, K. M., Yang, S., Roady, T., Kuo, J., & Lenné, M. G. (2020). Driver trust & mode confusion in an on-road study of level-2 automated vehicle technology. *Safety Science, 130*, 104845. https://doi.org/10.1016/j.ssci.2020.104845

Xie, T.-T., Wang, T.-Z., Wei, Y.-P., & Ye, E.-C. (2019). Declarative memory affects procedural memory: The role of semantic association and sequence matching. *Psychology of Sports & Exercise, 43*.

Yamani, Y., & Horrey, W. J. (2018). A theoretical model of human-automation interaction grounded in resource allocation policy during automated driving. *International Journal of Human Factors and Ergonomics, 5*(3), 225–239. https://doi.org/10.1504/IJHFE.2018.095912

Yeh, M., Swider, C., Jin Jo, Y., & Donovan, C. (2016). *Human factors considerations in the design and evaluation of flight deck displays and controls.* Federal Aviation Administration.

Zhang, B., de Winter, J., Varotto, S., Happee, R., & Martens, M. (2019). Determinants of takeover time from automated driving: A meta-analysis of 129 studies. *Transportation Research, 64*.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.