From semi to fully autonomous vehicles: New emerging risks and ethico-legal challenges for human-machine interactions

Thierry Bellet\textsuperscript{a,}, Martin Cunneen\textsuperscript{b,}, Martin Mullins\textsuperscript{b,}, Finbarr Murphy\textsuperscript{b}, Fabian Pütz\textsuperscript{b}, Florian Spickermann\textsuperscript{b}, Claudia Braendle\textsuperscript{c}, Martina Felicitas Baumann\textsuperscript{c}

\begin{itemize}
\item \textsuperscript{a} IFSTTAR (LESCOT: Ergonomics and Cognitive Sciences Laboratory), France
\item \textsuperscript{b} University of Limerick (Kemmy Business School), Ireland
\item \textsuperscript{c} Karlsruhe Institute of Technology (ITAS: Institute for Technology Assessment and Systems Analysis), Germany
\end{itemize}

\textbf{A R T I C L E  I N F O}

Article history:
Received 30 November 2018
Received in revised form 1 March 2019
Accepted 2 April 2019

\textbf{A B S T R A C T}

The provision of an adequate liability regime for ADAS technologies is an essential prerequisite for its roll out over the coming decade. Facing to the challenge of future highly automated vehicles, this paper proposed a Human-Machine Transition (HMT) approach as a common conceptual framework for considering Human Machine Interaction (HMI), liability and ethical issues in a unified way. The issues that arise are interrogated from a legal perspective, more specifically liability regimes and that of applied ethics. The paper highlights the issue of the handover/takeover. Potential consequences for insurance companies are then identified accordingly, with the aim to progress towards the sustainable deployment of automated vehicles on public roads.

\textcopyright{} 2019 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

\section{1. Introduction: Vehicle automation challenges and ethico-legal issues for HMI}

Current thinking on the process of introducing fully autonomous vehicles suggests that the technology will need to pass through a phase in which the driving task will be shared by the human driver and the vehicle. Such a phase is clearly fraught with difficulties as the responsibility for the safety of those within the vehicle and other road users will necessarily become a more fluid concept. As such, the legal community and ethicists are struggling to come to terms with the implications of Automated Vehicles (AV) in terms of liability regimes and questions of responsibility and culpability. Key to understanding the inherent complexities is the notion of handover/takeover transitions: the point at which the driver passes control to the vehicle or the vehicle back to the driver. This manoeuvre and others will be mediated through the human machine interfaces in the vehicle. For liability regimes which currently operate on notions of strict liability wherein the primary locus of responsibility resides with the driver, this represents a major challenge. Moreover, numerous ethical questions, particularly in respect of consent and autonomy, underpin the legal measures obtaining to the transition phase.

This paper is therefore necessarily interdisciplinary in nature and brings together the insights of the EU-funded VI-DAS\textsuperscript{1} project (Vision Inspired Driving Assistance Systems) research team which is comprised of HMI experts, insurance and liability

\textsuperscript{1} The VI-DAS project is an Horizon 2020 research and innovation program (under grant No. 690772) dedicated to the joint monitoring of both (1) the external traffic environment (i.e. exterior of the vehicle) and (2) the car driver's status (i.e. interior of the vehicle). Form this “720°” of monitoring, it is expected to design advanced driving aid systems based on vehicle automation, to be supported by adaptive HMI (i.e. context-dependent). To support these innovations regarding future HMI based on monitoring functions, a working group of experts (including HMI designers, insurance partners and liability specialists) was created, with the aim to define a common “conceptual framework” allowing them to jointly considering liability issues introduced by vehicle automation. This paper is the result of this inter-disciplinary effort.

https://doi.org/10.1016/j.trf.2019.04.004

1369-8478/© 2019 The Authors. Published by Elsevier Ltd.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
HMT phases partially supported by existing laws.

| Table 1 | HMT phases supported by existing laws. |
|---------|--------------------------------------|
| **HMT phase 1: Monitor and INFORM the driver** |
| HMI Key Issues | Inform the driver about: |
| | • The way to adequately perform the driving task in progress |
| | • The driving situation status and the future potential risks |
| | • The driver state/behaviours and the related potential risks |
| Legal Key Issues | What are the current/potential new related Legal Issues? |
| | • HMT phase should generally accord with the current liability framework |
| | • The question arises as to whether the driver acts negligently if the driver relies on information provided by the system. This could be regularly assumed, as the driver is only informed within a reasonable period of time. |
| Implications for Insurance | What are the current/potential new implications for insurance companies? |
| | • No significant implications for insurance, as regress claims against manufacturers should be very limited due to the only supporting function of the system |

**HMT phase 2: Monitor and WARN the driver**
HMI Key Issues | Warn the driver about: |
| | • The driving action to be immediately implemented or stopped |
| | • The driving situation status and the current effective risks |
| | • The driver state/behaviours and the related effective risks |
Legal Key Issues | What are the current/potential new related Legal Issues? |
| | • Question of driver negligence could be more relevant, as s/he potentially has to react in a shorter period of time to verify systems warning/alert (??) |
Implications for Insurance | What are the current/potential new implications for insurance companies? |
| | • No significant implications for insurance, as regress claims against manufacturers should be very limited due to the only supporting function of the system |

**HMT phase 3: assess risk management as performed by the driver**
HMI Key Issues | • If Risk adequately managed: stop to warn & go back to state (MD) |
| | • If Risk not adequately managed: make the decision to activate phase 4 (i.e. Takeover based on automaton decision) |
Legal Key Issues | What are the current/potential related Legal Issues? |
| | • No legal questions, if risky situation is resolved |
| | • If activation of phase 4 is inadequately induced by overriding the driver, this could have implications for manufacturer liability |
Implications for Insurance | What are the current/potential implications for insurance companies? |
| | • Relevance of regress claims against manufacturers could arise. However, high burdens to proof that conditions of manufacturer liability are met could impede the conduct of regress claims |

**Table 2**
HMT phases partially supported by existing laws.

| **HMT phase 4: Manage TAKEOVER transitions** |
| HMI Key Issues | • All HMI issues related to Human-Based Decision to activate the AD mode and to support a smooth transition from MD to AD |
| | • All HMI issues related to Automated-Based Decision to activate the AD mode and to support a smooth transition from MD to AD |
| Legal Key Issues | What are the current/potential new related Legal Issues? |
| | • If AD is inadequate and results in an accident, manufacturer liability could be relevant, if conditions of product liability are met |
| | • Burden to prove that conditions of manufacturer liability are met could impede the processing of regress claims |
| Implications for Insurance | What are the current/potential implications for insurance companies? |
| | • Lack of insurers ability, motivation and obligation to conduct regress claims on behalf of the insured driver/owner of the vehicle could impede liability costs to be reallocated to the manufacturer side |

**HMT phase 9: Manage HANDOVER transitions**
HMI Key Issues | • HMI functionalities to deactivate the AD mode and resume the manual control |
| | • Continuously inform the driver about the progressive AD status during the transfer, to support a smooth transition from AD to MD |
Legal Key Issues | What are the current/potential new related Legal Issues? |
| | • For level 3, driver has to maintain adequate awareness to takeover driving in a reasonable period of time: |
| | o Definition of an adequate period of time not yet legally defined by law texts or case law |
| | o Technical specifications of a legally required event data recorder are not yet defined |
| | • In case of systems failure: Is the driver entitled for compensation of own injury/property losses (this is regularly exempted for the human “driver” during manual driving action) |
| | • Who has to be held responsible for violation of traffic laws (e.g. speeding) in case of automated driving mode? |
Implications for Insurance | What are the current/potential implications for insurance companies? |
| | • Question of driver negligence in case that he does not maintain adequate level of awareness to takeover driving action |
| | • Question of burden to prove system failure and the ability/practicability of an event data recorder is not yet clear |

specialists, and applied ethicists, and charged with investigating solutions for this transitional phase as we move towards full automation. The central focus of this paper is about the “Human Machine Interactions” (HMI) which are progressively moving as “Human Machine Transitions” (HMT), according to the recent progress in vehicle automation. From the joint effort of this VI-DAS consortium, a typology of current and future Human-Machine Interactions/Transitions issues was identified. This typology
Table 3  
HMT phases to be supported by new laws.

| HMT phase 5: During AUTOMATED DRIVING |
|--------------------------------------|
| **HMI Key Issues**                   |
| * Keep the driver continuously informed about the current status of the Automated Driving (i.e., “mode awareness”) |
| * Allow the drivers to deactivate the AD at any time and support them in this deactivation process. |
| **Legal Key Issues**                 |
| * What are the current/potential new related Legal Issues? |
|   * In case of system failure: Is the driver entitled for compensation of own injury/property losses? (this is regularly exempted for the human “driver” during manual driving action) |
|   * Who is to be held responsible for violation of Traffic laws (e.g. speeding) in case of automated driving mode? |
|   * Can the driver be continuously excluded from driving task (following the Vienna Convention this is arguably not possible within the current status quo of legal environment) |
| **Implications for Insurance**       |
| * What are the current/potential new implications for insurance companies? |
|   * Question of burden to proof system failure and the ability practicality of an event data recorder is not yet clear |
|   * Lack of insurers’ ability, motivation, and obligation to conduct regress claims on behalf of the insured driver/owner of the vehicle could impede liability costs to be reallocated to the manufacturer side |

| HMT phase 6: Monitor and INFORM |
|----------------------------------|
| **HMI Key Issues**               |
| * Monitor the AD system limits and INFORM (prepare) the driver about future Take-Over Requests (liable to be planned) |
| * The driving situation status and the future potential risks |
| * The driver state/behaviours and the related potential risks |
| **Legal Key Issues**              |
| * What are the current/potential new related Legal Issues? |
|   * Driver has to maintain adequate awareness to resume driving within a reasonable period of time: |
|     * Definition of an adequate period of time not yet legally defined by law texts or case law |
|     * Technical specifications of a legally required event data recorder are not yet defined |
| **Implications for Insurance**    |
| * What are the current/potential new implications for insurance companies? |
|   * Questions arise as to thresholds of when the automated functions are suitable for extant condition |
|   * Will communication strategy include reference to impending liability regime? |

| HMT phase 7: Monitor and WARN the driver |
|----------------------------------------|
| **HMI Key Issues**                     |
| * Warn the driver about: |
|   * Monitor the AD system limits and failures, and WARN the Driver of urgent Take-Over Requests |
|   * Monitor driver status to safely resume the manual control of the vehicle and WARN them as required (e.g. distracted or asleep driver) |
| **Legal Key Issues**                   |
| * What are the current/potential new related Legal Issues? |
|   * Driver has to maintain adequate awareness to takeover driving in a reasonable period of time: |
|     * Definition of an adequate period of time not yet legally defined by law texts or case law |
|     * Technical specifications of a legally required event data recorder are not yet defined |
| **Implications for Insurance**         |
| * What are the current/potential new implications for insurance companies? |
|   * Question of burden to prove system failure and the ability/practicability of an event data recorder not yet clear |
|   * Lack of insurers’ ability, motivation, and obligation to conduct regress claims on behalf of the insured driver/owner of the vehicle could impede liability costs to be reallocated to the manufacturer side |

| HMT phase 8: Assess driver’s ability to resume the manual control |
|---------------------------------------------------------------|
| **HMI Key Issues**                                           |
| * Assess drivers’ ability to resume the manual control: |
|   * For AD Level 3: Support the Driver to promptly “come back” to the loop of control (e.g. advanced HMI to support drivers’ Situation awareness rebuilding, their decision-making, etc.). |
|   * For AD Level 4 and 5: define HMI and technical solutions in case of unavailable/unable driver for resuming the manual control (e.g. stop the car in a safe way and inform security services) |
| **Legal Key Issues**                                          |
| * What are the current/potential new related Legal Issues? |
|   * Typically, the difference between an L3 and L4/L5 Automated car is a crucial issue: |
|     * According to SAE L3 definition: The Human Driver MUST be able to resume the manual control (i.e. HANOVER is “legally” MANDATORY). Practically speaking, it means that phase 8 could be optional for L3 systems. Thus, the Vienna Convention may be still valid in its strict sense and existing legal regimes could be potentially adapted. Nevertheless, the driver has to be clearly informed when using these L3 systems about his/her full responsibility in promptly resuming the manual control in case of Take-Over Request. In addition, according to Bainbridge’s ironies of automation (1983) as previously mentioned, it is practically difficult (at both biological and cognitive levels) and ethically questionable to charge the human drivers with the full responsibility for resuming the manual control in all situations, while simultaneously creating an AD environment which totally excludes them from the loop of control. Consequently, technical solutions in case of drivers’ (un)availability in resuming the manual control should be mandatory (not “optional”, as suggested by SAE definition) and new legislations related to liability issues should be probably specifically redefined accordingly for L3. |
|     * According to SAE L4 and L5 definition: The Human Driver CAN be unable to resume the manual control (i.e. HANOVER is “legally” OPTIONAL and Design L4/L5 system able to keep the control is MANDATORY). In this case, the Vienna Convention must be definitively modified to permit the introduction of L4 and L5 vehicles to open roads. |
| **Implications for Insurance**                                |
| * What are the current/potential new implications for insurance companies? |
|   * Part 8 of the cycle has the potential to over-rule the driver. In a strict liability regime this has the potential to cause legal difficulties. |

* The literal interpretation of German Road Traffic Act combines L3 and L4 and approaches both levels equally. Hence, the only differentiating point here is that the time-span for taking over manual driving could maybe be longer for 4 from a legal point of view, but, in fact, the driver would have to takeover to not act negligently.
posits nine different categories of transitions grouped across two main fields, those of manual and automated control, aiming to support the future (r)evolution of human and machine relations introduced by vehicle automation.

1.1. “New Deal” introduced by vehicle automation

Due to remarkable developments in Advanced Driving Aid Systems (ADAS), the entire range of human driver activity in modern vehicles is undergoing change. Like the technological progress which revolutionized aeronautics at the end of the 1970s (Billings, 1991; Sarter & Woods, 1992), automated driving aid systems are now equipped to take the control of the vehicle under certain conditions (Young, Stanton, & Harris, 2007). In extreme scenarios, this could extend to the system assuming control without the consent of the driver. Thus, from a driving task which was once the total responsibility of the human driver in terms of road environment perception, decision-making and sensorimotor control, we are now heading towards a co-managed driving task under the joint authority of a complex entity: the “Human-Machine System” (Bellet, Hoc, Boverie, & Boy, 2011).

Regarding more specifically the Human-Machine Interactions (HMI), partial automation of modern ADAS introduced a Shared Driving Phenomenon (SDP), with a progressive increasing of the machine control during the last decade. However, if the SDP provides a fundamental change in the operation of vehicles, the liability regimes applied to assisted driving stay quite similar to manual driving. By contrast, the recent emergence of highly automated and fully autonomous vehicles (from Levels 3 to 5 of the SAE classification, 2014) introduced a totally new situation, not only regarding the HMI, but also in terms of their consequences for liability regimes and related ethical and legal issues (Kyriakidis et al., 2017). This disruptive technological evolution therefore underscores the urgent need to define a conceptual framework which facilitates a combined analysis of Human-Machine Interactions and related ethico-legal ramifications with a view to anticipating the development and near-future deployment of highly automated vehicles on public roads and managing any attendant problems.

In this general context, this paper examines four main topics:

- What is the new “risk balance” introduced by modern ADAS and vehicle automation?
- What are HMI management strategies and how does the current legal infrastructure support them for existing ADAS?
- What are the new HMI legal issues to be resolved in order to adequately support the design of future highly or fully automated vehicles?
- What are the ethical issues emerging from the advanced HMI solutions to be used in automated cars, particularly in respect of the current legal/liability situation?

---

2 In the SAE classification, Level 3 (L3) corresponds to “conditional automation”, when the driving task is performed by an automated system “with the expectation that the human driver will respond appropriately to a request to intervene”, and the L4 or L5, respectively corresponding to “High” and “Full Automation”, when the automated driving system is able to manage “all aspect of the dynamic driving task, even if the human driver does not respond appropriately to a request to intervene” (abilities however limited to “specific conditions” for L4, against for “all driving conditions” for L5).
The aim of this paper is to propose a general conceptual framework which focuses on the “Human-Machine Transitions” (HMT). It is held that such a framework will be suitable for application to all levels of vehicle automation (from manual to fully automated driving) in a unified way, regardless of HMI specificities and/or legal issues. Implications for insurance companies will be also considered on the basis of this conceptual framework.

1.2. The new “Risk Balance” due to modern ADAS and vehicle automation

The introduction of advanced driving aid systems in modern vehicles fundamentally changes the risk balance of the driving task. On one hand, ADAS actively contribute to increase road safety, but on the other, new types of risk also emerge, as illustrated in Fig. 1.

1.2.1. From the “Risk Reduction” side

The purpose of the ADAS technologies is to mitigate risks associated with driving. Such risks are contained in a range of important envelopes of operation, from parking assist to lane assist. The ADAS technologies involved use sensor and intelligence technologies to reduce risk either by providing additional information to the human driver or by assuming temporary control of the vehicle. The latter case is a prime example of the ADAS technological capacity to supplant the human driver in making more efficient decisions in response to driving scenarios which are judged to present more risk to human control than to machine control. Accordingly, the justification for the machine to take control in a pre-collision scenario is manifest in the potential reduction of harms or damage and it is evident that ADAS technologies offer a range of safety benefits which can also be framed as identifiable risk-reduction measures. Such apparent safety and risk-reduction benefits endorse the implementation of such technologies as standard equipment in all new models. Nevertheless, one challenging aspect of this view relates to Maurer et al. (2018) who reported that a number of ADAS elements, particularly the “guardian angel” pre-collision technology, was misunderstood by many users in their test group.

In fact, the study undertaken by Maurer et al. confirms that removing all control from a human driver in emergency situations gives rise to a variety of ethical challenges. As such they conclude that, “In the event of the car sensing a threat and being able to determine a strategy to avoid it, it is ethically obliged to intervene.” (Maurer et al., 2018, p. 341)

1.2.2. From the “New Risks” side

It is clear that the use of ADAS technologies provide opportunities to reduce risk and offer safer responses to driving events. However, such technological envelopes also give rise to new and complex risks. The complexity, in part, concerns the tension between an ethical duty to take action to reduce harm and the societal, ethical, and legal framing of human autonomy and informed consent. Wherever possible then, it is necessary to contrast and assess the potential risk mitigation opportunities with the new risks which often parallel using such technologies. The seemingly obvious choice to justify the adoption of safety benefits requires more careful scrutiny of how arguably safer machine alternatives engage with ethical and/or legal dilemmas. Human-machine interactions are a key element in mitigating any such tensions.

While HMI will deliver many different support functions to the human driver, the provision of important information to the human driver relating to the control and status of the vehicle is at its most critical in the transfer periods from human to machine and machine to human. Accordingly, there are two identifiable HMI functions: communication of desirable information to assist in improved operation; and (2) providing a mechanism that supports periods of control transfer (CT) by informing the driver in a supportive and timely manner to (a) take control, (b) to be prepared to take control, or (c) allow the vehicle’s system to take control. The period of CT represents a period of high risk given that numerous studies (Gold, Happee, & Bengler, 2018; Eriksson and Stanton, 2017) measure the critical timeframe involved as 5–25 s depending on such variable factors as driver distraction, environment, and the overall context of the driving scenario. Given the possible life-threatening nature of this challenging scenario, it is therefore crucial that HMI technologies are properly framed in respect of risk, ethics, and consent. The paper focuses on each of these concepts as integral to the accurate framing of the contextual uses of HMI in semi-autonomous vehicles.

2. HMI management strategies and related legal issues for existing ADAS

The development of current HMI management strategies relating to ADAS is best contextualised in the historical development of on-board driving aid systems. The first generation In-Vehicle Information Systems (IVIS) which appeared in vehicles (typically in navigation systems) during the 1990s, aimed to assist the driver by delivering information relating to the driving task in progress (Bellet et al., 2003). In terms of such information systems, the lower the risk or more distant the danger, the more the ADAS HMI strategy to merely inform the driver: in other words, to support them in perceiving events or negotiating road infrastructures and traffic rules. However, since Maurer et al. (2018) maintain that trust is dependent upon user perception and risk perception, it follows that increasingly complex ADAS technologies will rest on a greater degree of human driver trust to support the safety functions. In fact, complex ADAS will require users to provide a legal confirmation of their consent to yield control of the vehicle to an ADAS system as and when the system detects an emergency scenario.
From HMI delivered pieces of information, the main objective is to support the drivers in enhancing their Situation Awareness (Baumann & Krems, 2007; Bellet, Bailly-Asuni, Mayenobe, & Banet, 2009; Endsley, 1996) and to assist them in performing the driving task. Whenever the criticality and/or the time pressure of the traffic situation increases, the general HMI strategy is to warn the drivers as a means to assist them in “becoming aware” of a hazard or of their risky behaviour, and to invite them to implement a corrective action for managing the situational risk by themselves. At this level, the objective is not only to enhance drivers’ Situation Awareness, but also to increase their Risk Awareness (as defined by Bellet & Banet, 2012) in terms of both risk detection and assessment, and to subsequently support their decision-making process to adapt their driving behaviours accordingly. Many such types of ADAS, such as over-speed alerts, lane-keeping aids, blind-spot monitoring devices, or frontal-collision warning systems, are routinely integrated into the cars of today.

More recent developments in vehicle automation have enabled engineers to further extend the possibilities of HMI solutions. For instance, when the situational criticality is assessed as too high by the system and/or when the required action is so urgent as to be deemed unachievable by the human driver, the ADAS may directly intervene by assuming control of the vehicle through implementing, adapting, correcting, or stopping an action either inadequately or un-performed by the human driver. This is typically the case when “collision warning systems” become “collision mitigation functions” or “Automatic Emergency Braking” systems [AEB]. This eventuality is considered by Maurer et al. (2018) across three hypothetical scenarios; environmental awareness; driver error; and driver fatigue. Such examples are just a few of the many possible scenarios which examine when control of the vehicle should be taken from the human driver for safety reasons.

As demonstrated in Fig. 2 below, when taken together, all the HMI management strategies which support the current ADAS form a “progressive timeline” which is sustained by different types of driving aid systems by jointly considering (1) the cognitive tasks to be performed by the human driver and the adequacy of the driver state (e.g. distracted) or behaviours (e.g. dangerous), and (2) the criticality of the driving situation depending of the external risk and the time budget available to intervene for avoiding the accident. As this continuum of criticality and temporal pressure is traversed, so the legal and ethical issues to be resolved become more complex. Thus, solving the problem of risk communication to the driver is not merely a technical challenge but also requires wider societal input both in terms of legal regime and a more holistic acceptance from the ethical perspective.

With regard to the legalities, all successive generations of ADAS reside in similar frameworks: that is to say, that the driving task generally falls under the full responsibility of the human drivers. This is the case both for information or warning systems and ADAS technologies able to directly intervene on the vehicle controls. One interesting development is that §63a of the German Road Traffic Act now ensures third party access to certain defined recorded data of the vehicle in order to determine whether the accident was due to human driver error or a fault of the automated driving function. While this suggests a move towards a more nuanced liability regime, the law only requires the automated vehicle to record the position and time data when the change of the vehicle control takes place between the vehicle driver and the highly or fully automated system.

---

3 “AEB systems generally (though not exclusively) first try to avoid the impact by warning the driver that action is needed. If no action is taken and a collision is still expected, the system will then apply the brakes. Some systems apply full braking force, others an elevated level. Either way, the intention is to reduce the speed with which the collision takes place" (https://www.euroncap.com/en/vehicle-safety/the-rewards-explained/autonomous-emergency-braking/).
In any event, the entire question of Human-Machine Interaction may depend on whether vehicle automated functions are engaged at the driver’s initiative (as is the case for speed regulators, for example) or those that are spontaneously activated as circumstances dictates (as with collision avoidance devices). Provided that vehicle automation functions align with the human drivers’ activity (like Anti-lock Braking Systems [ABS\(^4\)], for instance), the resolution of Human-Machine Interactions issues are relatively straightforward from the liability point of view. Thus, when the driver decides to perform an action (like braking, for instance), the automaton consequently acts on the vehicle in order to optimize the action (and the decision) of the human. In this context, human-machine coupling is arguably intuitive; human drivers decide upon an action and the automaton assists in its implementation. By contrast, the more decisionality inherent in the ADAS, such as the capacity to deploy emergency braking to avoid an obstacle not perceived by the driver, the more increased the complexity of the HMI design. Amendments to the Vienna Convention\(^5\) which came into force as of March 2016 make provision for assisted driving technologies. Nevertheless, according to the articles of the convention, human drivers still bear the duty to monitor all types of ADAS (even those capable of assuming control of the car) and are still assessed as fully responsible for a crash in cases of accidents, with the exception of evident technical failure of the vehicle or of the ADAS algorithms.

3. The new challenges introduced by high automation and autonomous vehicles

In the context of highly automated driving (more specifically, for Levels 3 and 4 of the SAE classification), a new set of challenging paradoxes arise concerning both Human-Machine Interaction and legal aspects which are intrinsically linked to the “liability question”. While on the one hand, designers cannot develop or implement any HMI solution in real vehicles without considering their legal implications, conversely, providing a legal “answer” to the designers is only required for problems which are effectively “asked” by, or indeed created by HMI solutions. Consequently, to better support the development of future highly automated vehicles, a common conceptual framework to consider HMI and Legal is an a priori requirement.

The Human-Machine Transition (HMT) cycle, as presented in Fig. 3, may support the joint conceptual analysis articulated around the “liability issue” (i.e. responsibility in the vehicle control):

In the above HMT cycle, two main topics are of prior importance: namely, the “way to interact”; and the types of transition”. The “way to interact” refers to the manner of interaction between the human drivers and the system interfaces.

---

\(^4\) Anti-lock Braking System [ABS] is a safety system preventing wheel lockup and maintaining both steerability and vehicle stability, thus offering improved vehicle control and decreasing stopping distances.

\(^5\) See [https://www.unece.org/fileadmin/DAM/trans/doc/2014/wp1/ECE-TRANS-WP1-145e.pdf](https://www.unece.org/fileadmin/DAM/trans/doc/2014/wp1/ECE-TRANS-WP1-145e.pdf).
According to the drivers’ needs and the criticality of the driving situation (Fig. 2), the cooperation gradient to interact with the driver may be to: (1) deliver Information, (2) generate Warning, and (3) Takeover/Handover control of the vehicle control. The “type of transition” refers to the “Driving Mode”, i.e. Manual Driving (MD) versus Automated Driving (AD), and their related HMI and legal issues will depend on the initial state in the loop of control.

In cases of Manual Driving (SAE Level 0 to 2) as the initial state (i.e. corresponding to the red section of the Fig. 3), the driving task is initially under the full responsibility of the human drivers. To support them, a set of monitoring functions may be implemented to assess the driving situation and the potential risks (related to the drivers’ state or behaviour and/or to external traffic conditions), and then to accordingly activate the most suitable HMI modality in order to assist the driver (from Information or Warning delivery to vehicle Takeover by the automaton). At this stage of the HMT cycle, Automaton-Based Takeover transitions due to driver errors or inadequate state and behaviour is one of the two ways to manage the transitions from Manual to Automated Driving (Lu & de Winter, 2015; Lu, Happee, Cabral, Kyriakidis & de Winter, 2016). The second way corresponds to voluntary and deliberate transitions implemented by the drivers themselves; in other words, through the “delegation” of the vehicle control towards the automated driving systems (Hoc, Young, & Blosseville, 2009).

During Automated Driving (AD; Levels 3 to 5), and corresponding to the blue section of Fig. 3, the driving task is fully under the responsibility of the automation. However, monitoring functions need to assess AD systems limits or failure, and/or the future driving tasks that require a transfer of the vehicle control to the human driver. If this occurs, there will be a requirement to adequately support the transition from the AD to the manual driving. At this stage of the HMT cycle, a first crucial issue is to generate a Take-Over Request (TOR) for the driver (as an Information or as a Warning, depending on the emergency of the traffic situation and the available time budget), and then to manage the Handover transitions (i.e. from the Automated to the Manual Driving). Compared to pre-existing ADAS then, this presents a totally new challenge which is related to the assessment of human driver abilities to manually perform the current driving task. If the driver is assessed by the AD system as able to resume the manual control in a safe way, the transition can be implemented. In the other cases (for instance a distracted, unconscious or even deceased driver), autonomous systems should maintain full control of the vehicle. Moreover, another very important HMI issue during fully automated driving relates to the information which is continuously delivered to the human drivers to support “mode awareness” (i.e. their mental model about the current status and liability of the automated driving system). At every moment, the driver must be clearly informed by the AD system about the element of the driving task which is currently being performed by the autonomous vehicle and, if required, about incoming situations which will necessitate the resumption of manual control. Ultimately, a new generation of HMI is needed to keep the human appropriately informed and involved in the driving task while continuously supported by vehicle automation (Merat & Lee, 2012).

4. The “driver monitoring issues related to HMT and its ethico-legal consequences

During the two last decades, a large set of research efforts were focused on the driver monitoring issues. The general objective of monitoring functions is to assess the driver’s status in real time, like fatigue or drowsiness (e.g. Williamson & Chamberlain, 2005; Wang, Yang, Ren, & Zheng, 2006; Sahayadhas, Sundaraj, & Murugappan, 2012), visual or cognitive distraction (e.g. Harbluk, Noy, Trbovich, & Eizenman, 2007; Dong, Hu, Uchimura, & Murayama, 2011), mental workload (e.g. Brookhuis & de Waard, 2010), driving errors and risky behaviours assessment (e.g. Bellet et al., 2011), or on-board activity monitoring (e.g. Veeraraghavan, Bird, Atev, & Papanikolopoulos, 2007). With the emerging of highly automated vehicles, a new generation of monitoring functions may be developed for managing Human-Machine Transitions, according to both the drivers and the situational status. In this section we would like to further discuss the ethico-legal implications of human monitoring as introduced by recent progress in vehicle automation.

As presented in Fig. 3, the HMT cycle is divided into 9 phases which raise various ethical issues depending on whether they include monitoring of the driver or handover to the human driver/takeover by the machine. The monitoring to INFORM or WARN phases (i.e. 1, 2, 6 and 7) are especially relevant in regard to privacy and autonomy, given the capacities of the monitoring functions in analysing on-board activities and/or in “judging” driver ability to perform the driving task in a safe way. The management of HANDOVER/TAKEOVER transitions (i.e. 4 and 9) are particularly relevant in regard to responsibility and just distribution of costs and risks in case of accident, as they pose challenges for liability regulation. Finally, phase 8 is related to both aspects, and is probably the most challenging issue introduced by highly automated vehicles.

A first ethical issue introduced by in-vehicle monitoring functions is the continuous supervision and evaluation (in terms of risk taking, for instance) of driver behaviours, and the potential use of this type of data by insurance companies to support Pay As You Drive (PAYD) logic. The concept of PAYD denotes the use of variable pricing of insurance premium according to the performance of the driver (Institute of International Finance, 2016, p.2), and may have both negative and positive effects regarding driving safety and fair pricing (Baumann, Brändle, Coenen, & Zimmer-Merkle, 2018, pp. 52–53; Dijkstra et al., 2015, p. 95; Tong, Lloyd, Durrell, McRae-McKee, Husband, Delmonte, & Buttress, 2015, p. 5). Given the power of the insurance industry as a disciplinary force in society, it casts a long shadow over HMI in cars with regard to autonomy of drivers. It is unclear whether the privacy concerns which arise from data collection within cars (FIPA, 2015) can be fully accommodated by existing laws (such as European data protection regulation) or technical solutions, such as controlling the data access or anonymizing collected data (Derikx & Reuver, 2016, p. 73; Lüdemann, 2015, p. 247).
For insurance companies, the monitoring of individual driver risk is not only designed to create a new, user-oriented business model in order to accurately calculate premiums, but may be crucial to the settlement of liability claims. This is especially relevant in cases of accidents where the system failed to warn or handover control to the human driver during automated driving modes. According to EU product liability regulations the manufacturer would then be liable and the insurance paying for accident costs could take them into regress. The insurers would then have to prove the fault of the automated system, thereby clearly making the data documenting the accident circumstances key to their enquiries. While data collection to settle liability issues and thus compensate victims in a timely manner (especially in fault-based insurance systems such as the UK) is an ethical obligation, it must nonetheless be balanced against the rights of data ownership and privacy of drivers.

Liability issues also play a prominent role in the set of ethical challenges related to the transition phases between automated and manual driving modes (i.e. to manage HANDOVER/TAKEOVER) as responsibility shifts from the machine to the driver and back. In fault-based liability law the driver is liable if s/he caused the accident, such as, by failing to regain control from the machine. Even when jurisdictional liability law is not fault-based, the driver may be sued for negligence. In both cases responsibilities must be regulated by law in a clear manner and driver does and don’ts of automated mode properly communicated to the driver. However, regulations concerning the duties of drivers in automated mode remain rather vague and ambiguous. In Germany, for example, the recently amended law stipulates that the driver should be able to stay “alert” in order to resume “timely” control (StVG section 1(1), see: der Justiz, 2017). Yet, it is not at all clear how much takeover time is expected by regulators or what drivers are permitted to do in automated mode: the arguably critical point, especially in those instances where the takeover is unplanned and has to be virtually instantaneous. Such a legal framework creates uncertainty both for the driver and insurers but does appear to charge the driver with the greater share of responsibility. All of which question German lawmakers’ claims that the manufacturer is clearly liable when the car is in automated mode. Moreover, if taken “literally”, the schematic of a HMT cycle with the line drawn between red and blue phases marking human and machine responsibility is by no means clear.

In reality, there is no absolute separation of tasks or phases wherein the human driver or the machine has full responsibility. This is all the more apparent when one considers that human and machine monitor each other and ideally work together to coordinate their actions. Thus, the human driver has the responsibility to stay alert and monitor the machine in order to be able to take over, while at the same time, the machine monitors the human driver with the ultimate aim of supporting and influencing their decisions, including those to stay alert and not be distracted by other activities. To draw a definitive line between an advisory and an actively influencing effect of the machine is, of course, very difficult. It follows that the concept of phases with clearly defined responsibilities for either machine or human is also highly problematic. Moreover, speaking of a shared responsibility for both human driver and machine does not lead to a solution for the practical problem of liability distribution.

While in manual/automated driving (phases 0 and 5) the responsibility seems to be clearly distributed. This phase is in fact analogous to Monitor and INFORM (phases 1 and 6), whereby the human driver and machine monitor each other and “share” responsibility in a rather abstract way. Depending on the severity of the situation then, a detection of a risk by either human or machine can require an immediate managing of the takeover or handover (phases 4 and 9); all of which results in blurring the lines between machine/human separation of control and of responsibility. The consequences are the ethical problem of allocating liability fairly and the practical problem of the settlement of liability claims with the aforementioned privacy issues. This problem is further exacerbated by the question of whether human drivers are equipped to monitor activities in automated driving mode as well as for unplanned or emergency handover situations in the first place. As Bainbridge (1983) suggests, this is not necessarily the case, as humans actually fare worse in monitoring and takeover tasks within highly automated environments. Indeed this raises the issue of whether it is ethically right to charge the human driver with responsibility for monitoring and spontaneous handover situations while simultaneously creating an environment where the necessary abilities suffer.

A yet more complex legal question emerges in the HMT cycle from the possible situation of the driver refusing to follow HMI instructions. In phase 8 of the HMT cycle the machine may make a decision that the human lacks the ability to take control of the vehicle and return to phase 5. AUTOMATED DRIVING. From this moment on, both the legal and ethical status of the vehicle is conceptually contingent, as the human owner is over-ruled, raising questions with regard to autonomy and liability. In short, the crux of the matter is whether the final decision about handover or takeover initiation lies with the human driver or the machine. Here the ethical dilemma resides in the asymmetry of responsibility between driver and automated car. Two extreme scenarios may demonstrate this asymmetry more clearly:

- **Scenario A**: Automated mode detects difficulty and wants to handover control to the human driver. If the driver does not react fast enough or does not feel capable and refuses to take over control, and an accident occurs, the driver is liable.

---

6 A rather prominent case is for example the accident involving a test vehicle by Uber in autonomous driving mode. The vehicle systems failed to detect and subsequently killed a 49 year old woman crossing the street while pushing her bike (https://www.theguardian.com/technology/2018/mar/19/uber-self-driving-car-kills-woman-arizona-tempe).

7 E.g., in the accident cases with Tesla and Uber it had to be investigated whether the automated cars did not detect the other traffic participants on the road, or failed to interpret their data in a right manner, or give a warning sign in a timely manner; or whether the driver did not notice the warnings or could not manage to takeover fast enough.

8 This is only a thought experiment and has not been implemented in any prototypes so far, to the authors’ knowledge.
• Scenario B: The driver feels uncomfortable with the driving situation or tired and wants to switch to automated mode. If the machine then refuses or is not capable to take over and an accident occurs, the driver is again liable.

The asymmetry of the situation is all too clear: While the machine is able to abdicate responsibility to the driver in any case, the same is not true of the human driver. S/he cannot freely choose to handover the driving task and with it the responsibility. This asymmetry is however not generally legitimate or in full accordance with the notion of a driving “assistant” as opposed to “co-driver”.

However, apart from a general asymmetry which is arguably biased against the human driver, there is the additional problem of a frequent misconception or miscommunication: namely, that the machine can and will takeover responsibility from the driver in certain situations. While it is not clear if the user of the (semi-)autonomous car should be held responsible for any accidents caused by the car in any mode, even those in autonomous mode or in unclear handover situations, if this is what the legal situation amounts to, then this should be made explicitly clear to any user.

A basic ethical requirement is to distribute costs and risks in a just manner between manufacturers of cars with new HMI technology and drivers. As current liability regulation lags behind the technological developments, the ethical obligation to make driving safer with HMI and take the potential new risks seriously is even more pronounced. Moreover, principles of clear product information about the skills needed to use the HMI system should be provided to avoid unrealistic expectations and discourage unsafe use. It is evident that more legal, ethical, and user-centred research is needed in order to find solutions to the fundamental challenge of HMI in the car, which entails unprecedented shifts of responsibility between human driver and machine.

5. HMT-based “key questions” for insurance and liability

Often overlooked, but nevertheless key stakeholders, the insurance industry and liability experts are central to the sustainability to any move towards automated driving. The technological developments we are witnessing will inevitably result in a complete paradigm shift in the manner in which the motor insurance sector is managed. This is a sweeping and strategically important business which generates some €100 billion of premia annually in Europe alone. From a socio-ethical point of view then, it is a critical piece of infrastructure in that it mitigates some of the vast harm that results from contemporary road transport systems and provides a risk-sharing device to hundreds of millions of consumers across the European Union. As such, the manner in which the insurance industry manages the issues around HMT is clearly important. Foremost amongst their many challenges is the problem of shared responsibility between human and machine.

The tables below provide a summary of how insurers should best engage with this challenge and how relevant liability regimes are adapting. The overarching issue from a legal and insurance perspective is the ability and the burden of proof for producer liability. As such, agency is held to be an interesting point in terms of the intersection of liability and ethics. Returing to point 5 of the HMT: in this instance we could see the machine deny a handover to the human driver; a circumstance which would inevitably raise issues as to appropriate liability regime. While this section is as yet unable to provide “definitive answers”, it nonetheless aims to identify a set of “key issues” related to each HMT phase, for the combined consideration of HMI designers, legislators, and insurances, in support of the on-road introduction of highly automated vehicles. The “Green issues” (summarized in Table 1) are related to concerns which largely parallel existing ADAS. However, new HMI functionalities and vehicle automation technologies may imply new types of liability regimes in the future (Table 1).

5.1. Already legally-managed HMT (phases 1 to 3)

Theoretically, HMT phases 1, 2 and 3, HMI solutions should already be managed by the current laws, supporting the legal agreement of existing ADAS. However, new HMI functionalities and vehicle automation technologies may imply new types of liability regimes in the future (Table 1).

5.2. Partially legally-managed HMT (phases 4 and 9)

Strictly-speaking, the HMT phases 4 and 9 already exist in L2 systems (such as the Tesla Auto-pilot, for instance) regarding the requirement of “vehicle control transitions” between the human and the machine, and vice versa. However, for L2 systems, it is mandatory for the driver to “stay in the loop of control” by continuously monitoring vehicle automation (the driving task may be performed by one or several ADAS, but “with the expectation that the human driver perform all remaining aspect of the dynamic driving task”: SAE definition). As such, the Vienna Convention is applicable, and the human driver is ultimately responsible in case of accident. The totally new challenge introduced by fully automated car of L3 and more (see Table 2), is the opportunity for fully disengaged drivers during driving, until a takeover request of the AD (“with the expectation that the human driver will respond appropriately to a request to intervene”, the SAE definition for Level 3 systems, that is per contrast not expected for L4 and 5).
5.3. Totally new challenges (phases 5 to 8)

Finally, phases 5 to 8 of the HMT cycle require new specific HMI, laws and ethical considerations when compared with existing systems. The resolution of these issues is required to permit the deployment of fully automated vehicles (from L3 to L5) on open roads. The main points are summarized in the following Table 3:

Despite being a key stakeholder in terms of the development and indeed the acceptance of new technologies, the insurance industry is often overlooked. The analysis presented in the tables above is derived from the expertise of insurance professionals working in the field of assisted and automated driving. It confirms that much work remains to be done to provide an adequate legal/liability infrastructure around ADAS technologies.

6. Conclusion

The increasing use of the technologies and the move to the shared driving phenomenon and merging of driving responsibility between the human driver and the vehicle present unique conceptual difficulties. These difficulties emphasise the need to rehearse the theorizations of the phenomenon of shared driving since it is only by developing a comprehensive conceptual framework that each of the machine driving technologies can be assessed in terms of an informed metrics of risk, ethics, and governance.

Despite the many technical challenges and HMI ethical and legal issues to be resolved, the core question is not whether it is possible to permit Automated Vehicles on the public roads. Rather, it is a question of how, when, and under which conditions they should be introduced. Of course, the Bainbridge (1983) famous “ironies of automation” remain exactly the same, “but now the time has come to propose solutions” (Kyriakidis et al., 2017, p.6). One such key set of solutions relates to the formulation of the future liability regimes which must be put in place for the successful implementation of the technology.

The development of the HMT cycle outlined in Fig. 3 offers a common cognitive tool to enable HMI designers, legal specialists, and ethicists to interrogate issues around the fluidity of control of the automobile implicit in ADAS technologies. In this context, one of the central goals of the VI-DAS project, that of finding a technological solution for determining the state of readiness of the human driver, is a prerequisite of the implementation of the ADAS technologies. Technological solutions, in the case of VI-DAS, and video analysis of the human driver status, are clearly vital. However, the legal/liability and ethical components of the HMT cycle also need to be instantiated.

The propensity for technology to advance at a speed which outstrips the ability of governance regimes to keep pace is well documented. As insurance law is a key component of such regimes, this paper seeks to elucidate the implications of the various aspects of the HMT cycle in terms of liability. At the same time, legal regimes which are underpinned by ethical considerations that are at odds with the dominant ethical norms are unlikely to be successful. Hence, this paper examines both legal and ethical thinking around the HMT cycle, and as such, represents a significant step forward in the ongoing debates on the design and implementation of future automated vehicles.

Acknowledgement

The VI-DAS Project is funded by the European Union’s Horizon 2020 Research and Innovation Program under Grant No. 690772.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.trf.2019.04.004.

References

Bainbridge, L. (1983). Ironies of automation. Automatrica, 19, 775–779.
Baumann, M., & Krems, J. (2007). Situation awareness and driving: A cognitive model. In P. C. Cacciabue (Ed.), Modelling driver behaviour in automotive environments. Critical issues in driver interactions with intelligent transport systems (pp. 253–265). London: Springer.
Baumann, M. F., Brändle, C., Coenen, C., & Zimmer-Merkle, S. (2018). Taking responsibility: A responsible research and innovation (RRI) perspective on insurance issues of semi-autonomous driving. Transportation Research Part A: Policy and Practice, 1, https://doi.org/10.1016/j.tra.2018.05.004.
Bellet, T., & Banet, A. (2012). Towards a conceptual model of motorists’ Risk Awareness: A comparative study of riding experience effect on hazard detection and situational criticality assessment. Accident Analysis & Prevention, 49, 154–164.
Bellet, T., Bailly-Asuni, B., Mayenobe, P., & Banet, A. (2009). A theoretical methodological framework for studying and modelling drivers’ mental representations. Safety Science, 47, 1205–1221.
Bellet, T., Hoc, J.-M., Boverie, S., & Boy, G. A. (2011). From human-machine interaction to cooperation: Towards the Integrated Co-pilot. In C. Kolski (Ed.), Human-computer interaction in transport (pp. 129–156). Ashgate.
Bellet, T., Tattegrain-Vente, H., Chapon, A., Bruyas, M. P., Pachiaudi, C., Deleurence, P., & Guillhon, V. (2003). Ingénierie Cognitive dans le contexte de l’assistance à la conduite automobile [Cognitive engineering within the context of driving assistance]. In G. Boy (Ed.), Ingénierie Cognitive: HMI et Cognition (pp. 323–414). Paris: Hermes Science-Lavoisier.
Billings, C. E. (1991). Human-centered Aircraft Automation: A Concept and Guidelines (Technical Memorandum). Moffett Field, CA: NASA Ames Research Center, Brookhuis, K. A., & de Waard, D. (2010). Monitoring drivers’ mental workload in driving simulators using physiological measures. Accident Analysis & Prevention, 42(3), 898–903.
Bundesministerium der Justiz und für Verbraucherschutz (2017). Straßenverkehrsgesetz (StVG) (Road Traffic Act): https://www.gesetze-im-internet.de/stvg/index.html, last accessed: 05.07.2018.
