Methodologies to assess usability and safety of ADAS and automated vehicle

Annie Pauzie*. O. Orfila. **

*IFSTTAR/TS2/LESCOT, Cité des mobilités, Allée Général Benoist "Parc du Chêne" Case 24 - 69675 -Bron Cedex / France
(email: annie.pauzie@ifsttar.fr)

** IFSTTAR/COSYS/LIVIC 25 allée des Marronniers 78000 VERSAILLES (e-mail: olivier.orfila@ifsttar.fr)

Abstract: In the framework of future innovation and for the sake of road safety, there is a great hope in fully supporting, or even replacing, the human driver by reliable technology. But, due to the novelty of this context, an important care will have to be devoted to investigate drivers’ expectation, needs, behavior and functional abilities to reach this goal. In this context, this paper reviews several human factors issues related to partial and fully automated vehicles, with discussion of strengths and weaknesses of methods investigating driver automation acceptability, trust, situation awareness and workload. Main results of these parameters in relation to automated driving are presented and relevant methodologies to investigate these human variables are discussed in the perspective of real road experiments context.

Keywords: human factors, automation, methodology, road safety, human centred design, acceptability, trust, situation awareness, workload

1. INTRODUCTION

Several studies showed the importance of the human centred approach for the design and the implementation of technology in transport, and more especially in automotive context (Barnard et al., 2010). Indeed, taking into consideration the knowledge on human behavior, drivers’ functional capacities and needs at the early stages of a prototype development helps designers to set up safer systems (Pauzié, 2014). In the perspective of the partial and fully automated vehicles, several issues linked to human factors will have to be carefully studied in order to reach one of the main ambitious objectives of self-driving: increase road safety in a zero accident perspective.

2. ISSUES IN DRIVER CENTRED DESIGN AND AUTOMATED VEHICLE

The fact that human will delegate some or the total control of her/his car to systems is raising new human factors issues in terms of acceptability, trust, situation awareness and mental workload.

Technology acceptability deals specifically with perceived usefulness and perceived ease of use (Davis, 1989) and is influenced by belief, concern and expectation of the population of users. A professional survey compared the attitudes surrounding autonomous vehicles. Polling 17,400 vehicle owners, it showed that opinions are split in the total population (Power and Associates, 2012). Social acceptability is linked to the idea that automation will deprive people of personal control over their vehicles and are possible concerns of automated driving (Howard & Dai, 2014; Shladover, 1998), this specificity being diversely accepted depending on the drivers’ personality, cultural background and generational belonging. An early survey conducted by Bekiaris, Petica, and Brookhuis (1997) found a definite rejection of automated driving. More recently, after experiencing highly automated driving (HAD) in a driving simulator, only 13 of 38 participants stated that they wished to have the system in their car, possibly because the system did not prevent incidents (Schieben et al., 2008). Some drivers liked the increase of comfort and safety while some respondents did not like the idea of handing over control to automation (Flemisch et al. 2011). People can find the lack of control unsettling, believing the technology to be unreliable and the programming to be incapable of proper control, worrying about the risk of computer malfunction (Klayman, 2012). Payre et al. (2015) showed that the driver’s acceptability is related to the type of road context, with preference of delegating the control of the vehicle on highways, in traffic congestion and for automatic parking.

Acceptance is linked to usability characteristics of the system leading to trust. It is vital for successful implementation and is a precondition for these systems to achieve the benefits they claim (Najm, Stearns, Howarth, Koopmann, & Hitz, 2006). In-vehicle system acceptance depends upon drivers’ motivation and driving style, with a high variability among the population. At this stage, despite numerous studies on acceptance of driver assistance systems, neither a common definition nor a standardized measurement procedure is available (Ghazizadeh, Lee, & Boyle, 2012).

Acceptability and acceptance of automation are major challenges allowing success in implementation and a prior condition to get the expected benefits (Najm, Stearns, Howarth, Koopmann, & Hitz, 2006).

Trust is considered as a key variable for reliance on, and misuse/disuse of, automated systems (Kazi, Stanton, Young, & Harrison, 2005) and has a direct impact on the level of acceptance. Trust is a complex interaction involving dispositional trust (culture, age, gender, personal traits), situational trust (setting, difficulty, task, risk), initial, learned trust (pre-existing knowledge), and dynamic, learned trust (system performance, reliability, validity, errors) (Hoff and
Bashir, 2014). The issue of trust in the system is needed to be considered both in partial and full automation, being crucial in the second case where the driver will be then able to be fully involved in another activity requiring a lot of attention, or to really relax or even have a nap. There might be an issue not only related to lack of trust in the reliability and performance of the automatic system, but also an issue of over-trust where the human simply trusts the technology too much (Strand and al., 2014), leading to misuse and disuse (Parasuraman et al., 1997; Körber et al., 2014). This will have to be highly considered in the scenario of driver getting back control to the driving task.

Situation awareness can be defined as “knowing what’s going on so you can figure out what to do” (Adam, 1993). Endsley (1988a) which states that situation awareness is “the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future”. It is important to consider this variable in the framework of automated vehicle, studies already showed that partly automation can lead to impoverish situation awareness (Merat, Jamson, Lai, & Carsten, 2010; Carsten, 2004) with probably higher effect for highly automated situation. Indeed, a wealth of evidence from simulator studies shows that Highly Automated Driving evoke long response times and an elevated rate of (near-) collisions in critical events as compared to manual driving (Strand, Nilsson, Karlsson, and Nilsson, 2014). Merat and Jamson (2009) found that drivers of a highly automated car took 2.5 s longer to press the brakes in response to a red traffic light than people driving manually in addition to slow brake response times with respect to emerging and oncoming vehicles. Damböck et al. (2013) found longer reaction times in a scenario where the lead vehicle braked hard with the longitudinal controller failing at the same moment and in a scenario where a wild animal ran onto the road, undetected by the sensors. However, there are counterexamples, where drivers successfully avoid collision in critical event scenarios. Essentially, if the automation fails unexpectedly with very little time for the human to respond, then almost all drivers crash (Flemisch et al., 2008), but if drivers receive a timely warning then almost all drivers will safely avoid collision (Gold et al., 2013).

The level of situational awareness is a direct consequence of drowsiness, distraction, health status, fatigue, vigilance and involvement in activities not linked to driving task. Its evaluation will be especially important in a context of highly automated systems that requires the human driver to cycle in and out of the automated driving mode during a trip.

Mental workload is a psychological construct, difficult to define and difficult to assess, depending upon the task demands in relation to the amount of resources the operator is willing or able to allocate, and is therefore a relative concept (De Waard, 1996). In the framework of full automation, one of the most critical scenario in relation to mental workload will be for the driver to perform transitions from manual to automated control, or vice versa. Some authors argued that adjusting the automation status is itself a secondary task that should not induce too much extra workload (Thompson and Tönnis, 2007); nevertheless, in this particular scenario, the issue of the drivers ability to take the control back according to their current state, and the consequently induced cognitive cost for them to manage correctly the action, still needs to be clearly understood. Furthermore, workload can be reduced as the driver is relieved from the cognitive activity associated with manual driving and from the physical activity of moving the pedals and steering wheel (Dragutinovic, Brookhuis, Hagenzieker, and Marchau, 2005).

Workload and situation awareness are two of the most important Human Factors constructs predictive of performance and safety (McCaulley & Miller, 1997; Parasuraman, Sheridan, & Wickens, 2008; Sarter & Woods, 1991; Stanton & Young, 2000). Partial automation could raise driver’s workload and situation awareness issue if the driver has to remain vigilant and monitor the automation status (Winter et al. 2014, Carsten, 2010). They are human cognitive state rather than causal agents (Flach, 1995) and therefore require well defined measurement procedures (Hand, 1996).

The driver’s workload will have also to be assessed in conjunction with stress while performing transition from automated to manual control, in order to ensure safe handoff from the automated vehicle function to the driver in different driving conditions in the time allocated by the system. All these variables will have to be deeply investigated due to the novelty of the situation and the consequently poor level of knowledge and understanding gathered in real road context experiments at this stage. Fortunately, several interesting driving simulator studies allowed setting up some basic understanding of human behavior and mental state in partial and full automation situation (Winter et al. 2014, Merat et al., 2014). Due to the maturity of the technology nowadays, we reach the point where methods to investigate these human factor parameters have to be relevant for real road investigation.

4. METHODOLOGIES TO INVESTIGATE HUMAN FACTORS ISSUES IN REAL ROAD CONTEXT

Methodologies to investigate human factors will have to be adapted to the level of automation, as each of this level will raise specific issues regarding driver abilities requirements closely linked to the modalities of automated functionalities, and will have to be conducted in realistic context. So far, there are almost no studies testing highly automated driving in natural environment. Fully autonomous vehicles are not yet ready, and vehicles with highly automated driving technologies are still rather rare. Therefore, existing studies on automation have used interviews, online questionnaires (KPMG, 2013; Payre et al., 2015; Rödel, 2014) or different contexts of driving simulation (Strand et al., 2014; Lee et al., 2015). However, human factors studies required, at this stage of technical maturity, more real-life studies to get ecological valid data. For example, in this purpose, the University of Michigan and MIT have created a mock-up set of busy streets in Ann Arbor to provide tests for self-driving vehicles in an urban environment (Knight, 2014), aiming to conduct a large-scale test with 2,000 driverless cars on the road within the next eight years. In the same vein, other projects using test tracks and dedicated experimental roads will be conducted in the coming decade to be able to gather reliable data on human
behavior and automated vehicle. In this framework, adapted methodologies will have to be identified to investigate relevant human factor variables that could potentially impact system safety and usability.

### 4.1 Acceptability, acceptance and trust

To investigate acceptability of partial and fully automated vehicles, several methodologies have been used, more or less combined, such as online questionnaires, professional surveys, focus group interviews, questionnaires, in-depth interviews (Casley et al., 2013).

The online questionnaire has a great advantage in terms of high number of participants that can be involved in the survey with reasonable time investment for researchers; the limit of this method is that participants cannot be observed while answering (Leggett, Kleckner, Boyle, Duffield, & Mitchell, 2003). So, this type of investigation can be biased being based on imagination and not actual experience (Rödel et al. 2014). Indeed, it is difficult to test something that is not fully implemented in a real world context, as end-users tend to think in an abstract manner rather than thinking about the situation in the real world. The methodology of focus groups and in-depth surveys allowed overcoming this limit, with more contextualized data in relation to the personality and the motivation of the respondent, getting closer to a real world situation and to avoid as much as possible imaginative thoughts. But, due to the length of the process, it induces smaller size of interviewers sample that online questionnaire. Methodology of focus groups and in-depth surveys can be combined with some real road tests to get subjective point of view and opinions based upon real experience of the situation.

To study acceptance, most of the investigations have been conducted on driving simulator (Tanaka et al. 2015). Nevertheless, Neale and Dingus (1998) and Farber (1999) stated that simulators do not provide an accurate representation of the phenomenology of real automated driving.

Trust and autonomous driving have been investigated in simulation and experimental studies (Helldin et al., 2013; Verbene et al., 2012). This variable is usually evaluated through questionnaires before and after experience of driving. In addition to these participants ‘ratings, Gold et al. (2015) recorded gaze behavior during a driving simulator session in order to measure a potential change of trust by a change in scanning behavior. In this study, results indicate that horizontal gaze behavior could not be confirmed as a metric for measuring trust in automation.

### 4.2 Situational awareness, workload and stress

Several measures of situational awareness have been developed, the most widely used among them being the Situation Awareness Global Assessment Technique (SAGAT, Endsley, 1988b) and the Situational Awareness Rating Technique (SART). SAGAT provides an objective measure of situational awareness based on queries provided to the operator to assess his or her knowledge of what was happening at the time during freezes in a simulation while SART provides a subjective rating of situational awareness by operators (Endsley et al., 1998).

The main advantage of SAGAT is that it allows an objective index of situation awareness. Its main disadvantage is that it requires to temporarily freeze the road scenery in a simulation context to test whether the driver has observed and understood the host vehicle’s state, the road infrastructure, objects in the environment, and the behaviours of other road users. It has also been frequently asserted that another disadvantage of this method is that the technique relies on memory (Dreyfus, 1981; Nisbett and Wilson, 1977).

The method SART provides an assessment of situation awareness based on user’s subjective opinion with a set of components determined through analysis to be relevant according to the context. Participants rate on a series of bipolar scales the degree to which they perceive (1) a demand on their resources, (2) supply on their resources and (3) understanding of the situation. These scales are then combined to provide an overall SART score for a given context. The main advantages of SART are that it is easy to use and can be administered in a wide range of task types and can be used in real world tasks as well as simulations. Potential limitations of SART have been asserted to include (Endsley, 1995): (1) the inability of individual to rate his own situation awareness (2) the possible influence of the subjective evaluation of the performance level on the subjective rating of the situation awareness (3) possible confounding with workload issues while situation awareness may operate as an independent factor from workload in many situations (Endsley, 1993).

Furthermore, it has been shown that the SART scores were highly correlated with confidence level and subjective performance, leading to consider that subjective situation awareness ratings be viewed as good indices of these aspects, but perhaps not veridical representations of situation awareness itself (Endsley et al., 1998).

Situation awareness can be also evaluated while recording driver’s behavior and attitude: in an interesting study, Omae, Hashimoto, Sugamoto, and Shimizu (2005) let 30 drivers experience a ride in a real highly automated vehicle on a test track. Even though the participants were told that the automation could display steering failures that required manual intervention, 8 participants fell asleep during the drives. Some participants started reading, operating their mobile phone, crossing their legs, or leaning out of the window. The drivers stated that they engaged in these behaviours because the task was boring and they had nothing to do.

Situation awareness can be inferred from drivers’ eye-movements recording. For example, drivers in highly automated driving contexts are less likely to gaze at the road centre than manual drivers (Barnard and Lai 2010, Carsten et al. 2012), which indicates that they have altered situation awareness compared to manual driving. This variable can also be inferred by measuring performance of driver’s response at critical events.

Several methods have been developed to measure mental workload (Colle H.A, 1998): measurements of physiological parameters, dual-task method (Reid G., Nygren T., 1988) and methods that elicit drivers’ subjective judgments about
the workload they have experienced. In terms of physiological measurements, it has been shown that cognitive load can reduce skin conductance (Cha, 2003), increases eye-blink rate (Cha, 2003; Damböck et al., 2013; Merat et al. 2012) and can be measured through heart rate variability (e.g., Brookhuis, Van Driel, Hof, Van Arem, & Hoedemaeker, 2009; De Waard et al., 1999; Maysers, Piechulla, Weiss, & König, 2003; Takada & Shimoyama, 2001; Takano & Kobayashi, 2004; Törnros et al., 2002; Wille, Röwenstrunk, & Debus, 2007). Using this variable, it has been shown that partial automation tend to reduce heart rate as compared to manual driving, indicating a reduction of workload (Carsten et al., 2012; De Waard et al. 1999). However, not all studies are consistent in this respect. Furthermore, measurement of physiological parameters raises important difficulties in real road context, that is why most of the workload driver’s measurement in this case has been conducted using questionnaires or tests to assess cognitive cost. Subjective or self-assessed measures allow estimates from individual’s reports concerning the workload or effort expenditure that was experienced during the task (Tokunaga R.A., 2000). These measures are often used in practice because they have many advantages in terms of easy running process over objective measures (Patten, 2004). Among these methodologies, one of the more adapted to the context of driving is the DALI (Driving Activity Load Index, Pauzie, 2008), a revised version of the NASA-TLX, the latter one being created by the army and originally designed to assess pilot workload in the aviation domain. The DALI has allowed testing drivers’ mental workload in various contexts linked to the use of in-vehicle technologies (Harvey et al., 2011; Tretten, 2011; Tretten et al., 2009, Kim & Woh, 2011, Kern et al., 2010).

5. CONCLUSIONS
Automated driving is an important hope for road safety. Some results from a real road study showed that a partially automated car outperformed human drivers with better responses to unanticipated events, faster than manual drivers did (Davis et al., 2008). In the same vein, Marcus (2012) stated that, in a few decades, automated driving will be so reliable that humans will not be legally allowed to drive. Nevertheless, at this stage, several questions need to be resolved in terms of human factors and will require deep investigations in realistic conditions to evaluate how the driver will accept, face, use and manage this situation according to the diversified scenarios linked to the several levels of automation. The methodologies to assess drivers’ requirements and capacities will need to be tested and maybe to be adapted taking into account the novelty of this situation in order to be fully efficient to prepare the reality of this huge step in the driving task concept and to ensure expected positive consequences for the road safety.

REFERENCES
http://dx.doi.org/10.4271/2003-01-0119
http://dx.doi.org/10.1109/itsc.2013.6728467
expertise (ORC 81-3). Berkeley: Operations Research Center, University of California.


Kern D., Mahr A., Castronovo S., Schmidt A., Müller C., (2010). Making Use of Drivers’ Glances onto the Screen for Explicit Gaze-Based Interaction. Second International Conference on Automotive User Interfaces and Interactive Vehicular Applications, November 11-12, Pittsburgh, Pennsylvania, USA.


