Transportation Research Part F 43 (2016) 183–198

Contents lists available at ScienceDirect

Transportation Research Part F

journal homepage: www.elsevier.com/locate/trf

Human factors of transitions in automated driving: A general framework and literature survey

Zhenji Lu [a,*], Riender Happee [a], Christopher D.D. Cabrall [a], Miltos Kyriakidis [a,b], Joost C.F. de Winter [a]

[a] Department of BioMechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, The Netherlands
[b] Future Resilient Systems Singapore-ETH Centre, ETH Zurich, Singapore

ARTICLE INFO

Article history:
Received 21 December 2015
Received in revised form 5 September 2016
Accepted 4 October 2016
Available online 31 October 2016

Keywords:
Human factors
Automated driving
Transitions
Transition classification

ABSTRACT

The topic of transitions in automated driving is becoming important now that cars are automated to ever greater extents. This paper proposes a theoretical framework to support and align human factors research on transitions in automated driving. Driving states are defined based on the allocation of primary driving tasks (i.e., lateral control, longitudinal control, and monitoring) between the driver and the automation. A transition in automated driving is defined as the process during which the human-automation system changes from one driving state to another, with transitions of monitoring activity and transitions of control being among the possibilities. Based on 'Is the transition required?', 'Who initiates the transition?', and 'Who is in control after the transition?', we define six types of control transitions between the driver and automation: (1) Optional Driver-Initiated Driver-in-Control, (2) Mandatory Driver-Initiated Driver-in-Control, (3) Optional Driver-Initiated Automation-in-Control, (4) Mandatory Driver-Initiated Automation-in-Control, (5) Automation-Initiated Driver-in-Control, and (6) Automation-Initiated Automation-in-Control. Use cases per transition type are introduced. Finally, we interpret previous experimental studies on transitions using our framework and identify areas for future research. We conclude that our framework of driving states and transitions is an important complement to the levels of automation proposed by transportation agencies, because it describes what the driver and automation are doing, rather than should be doing, at a moment of time.

1. Introduction

Car driving is becoming automated to an ever greater extent. Presently, most car manufacturers have released cars that are equipped with adaptive cruise control (ACC) and/or lane keeping assistance (LKA) systems, which are technologies that assist in the longitudinal and lateral driving tasks, respectively. In field operational tests, these driver assistance systems have been found to raise traffic efficiency and to reduce energy consumption (e.g., Alkim, Bootsma, & Hoogendoorn, 2007). Moreover, such systems may reduce the number of traffic accidents (Kuehn, Hummel, & Bende, 2009), most of which are currently attributed to human error (Brookhuis, De Waard, & Janssen, 2001; Dingus et al., 2006; Storie, 1977; Treat et al., 1979).

---

[*] Corresponding author.
E-mail address: z.lu@tudelft.nl (Z. Lu).

http://dx.doi.org/10.1016/j.trf.2016.10.007
1369-8478/

The existing driver assistance systems function as supportive automation and keep the driver in the loop by requiring the driver to monitor the environment and control part of the driving task. More advanced technologies that allow the driver to be out-of-the-loop for extended periods are now starting to be introduced. Three authorities, namely the German Federal Highway Research Institute (BASt; Gasser & Westhoff, 2012), the Society of Automotive Engineers (SAE, 2014), and the United States National Highway Traffic Safety Administration (NHTSA, 2013) have each formulated definitions that classify automated driving systems from driver assistance to full automation. In fully automated driving, the automation takes care of all monitoring and control activities, and a driver is not strictly needed anymore other than to set a destination. However, several problems, such as limitations of technology, divergent public acceptance, liability issues, and human-machine ethics, are yet to be solved before fully automated driving can become publicly available at a wide scale (e.g., Kyriakidis, Happee, & De Winter, 2015).

Previous human factors research indicates that automation resolves the imprecision and variability of human task performance, but also yields new types of safety concerns. It has been found that a high level of automation can cause out-of-the-loop problems such as complacency, skill degradation, mental underload (when the automation functions reliably), mental overload (when the operator suddenly needs to solve an automation-induced problem), and loss of situation awareness (Bainbridge, 1983; Bibby, Margulies, Rijnsdorp, & Withers, 1975; Endsley & Kiris, 1995; Hancock et al., 2013; Kaber & Endsley, 1997; Parasuraman & Riley, 1997; Vlakveld, 2015), which are issues that have also been implicated in the domain of automated driving (De Winter, Happee, Martens, & Stanton, 2014; Seppelt & Victor, 2016; Young & Stanton, 2002). Recently, a meta-analysis of 18 experiments on human-automation interaction found statistical support for the so-called lumberjack hypothesis, which postulates that as the degree of automation increases, the side effects of automation (e.g., performance impairment if the automation fails) increase as well (Onnasch, Wickens, Li, & Manzey, 2014). In the domain of automated driving, it has been argued that not the highest levels of automation, but intermediate levels in which the human is expected to monitor the automated driving system, may be particularly hazardous because humans are unable to remain vigilant for prolonged periods of time (Casner, Hutchins, & Norman, 2016; Norman, 2015). These studies make clear that due to the changes in the driver's role in automated vehicles compared to manually driven vehicles, human factors need to be carefully considered by researchers, designers, and policy makers (see also Kyriakidis et al., submitted for publication; Merat & Lee, 2012).

Bainbridge (1983) argued that 'taking over control' is a primary task left for the human operator who supervises an automated system. Indeed, one cannot ignore the fact that automated driving systems will occasionally fail (Goodall, 2014), which implies that a driver has to resume control to avoid crashing. Moreover, automated driving systems of the near future will probably not be able to cover all traffic conditions, which implies that the driver has to take over control to avoid a collision or traffic violation. Empirical studies have confirmed that accidents and near-accidents are likely to occur in situations where drivers suddenly have to resume manual control from an automated driving system (e.g., De Waard, Van der Hulst, Hoedemaeker, & Brookhuis, 1999; Flemisch, Kelsch, Löper, Schieben, & Schindler, 2008; Jamson, Merat, Carsten, & Lai, 2013; Schermers, Malone, & Van Arem, 2004; Zeeb, Buchner, & Schrauf, 2015). The aforementioned out-of-the-loop problems exacerbate the inability of the driver taking back control from automation. Thus, it is important to investigate control transitions in automated driving, especially when considering that human factors studies have repeatedly demonstrated that humans are not good at supervisory tasks (Hancock, 2015; Mackworth, 1950).

One issue that occurs when interpreting the experimental literature on control transitions is that the results are much determined by the specific automation functions, traffic conditions, and task instructions (see De Winter et al., 2014 for a review). To be able to derive more general conclusions on driver behaviour across different automated driving systems and traffic situations, this paper proposes a framework that defines and classifies transitions focusing on changes of driving states. This framework is intended to build a dialogue among researchers who share common interests in understanding how drivers behave during transitions in automated driving. Our concept of driving states differs from the existing BASt, SAE, and NHTSA levels of automation because it formally outlines possible allocations of primary driving tasks and is descriptive rather than normative. That is, our framework describes what the driver and automation are doing at a given moment of time (descriptive approach) rather than what they should be doing according to design criteria/standards of conduct (normative approach).

This paper is organised as follows. Section 2 defines transitions between driving states. We explain that the driving states represent how the primary driving tasks of longitudinal control, lateral control, and monitoring are distributed between the automation and the driver, and that transitions are defined as a change from one driving state to another. Section 3 introduces a classification tree that categorizes different types of control transitions. In Section 4, we review experimental studies that are concerned with transitions in automated driving, and interpret the findings using our transitions framework. Finally, Sections 5 and 6 present research gaps and draw conclusions arising from this review and applications of the new framework.

2. Definition of transitions in automated driving

Most studies on transitions in automated driving have defined a 'transition' as either an activation or a deactivation of a function (Gold, Damböck, Lorenz, & Bengler, 2013; Miller, Sun, & Ju, 2014; Nilsson, Falcone, & Vinter, 2015; Pauwelussen & Feenstra, 2010; Toffetti et al., 2009), or a change from one level of automation to another (Merat, Jamson, Lai, Daly, & Carsten,