Driver error or designer error: Using the Perceptual Cycle Model to explore the circumstances surrounding the fatal Tesla crash on 7th May 2016

Victoria A. Banks*, Katherine L. Plant & Neville A. Stanton

Transportation Research Group, University of Southampton, UK

*corresponding author: v.banks@soton.ac.uk

Abstract

"Human error" is often implicated as a causal factor in accident investigation yet very little is done to understand 'why' such errors occur in the first place. This paper uses the principles of Schema Theory and the Perceptual Cycle Model (PCM) to further explore the circumstances surrounding the fatal Tesla crash in May 2016 in which the driver was fatally injured using team-PCM representations. The preliminary National Highway Traffic Safety Administration accident investigation concluded that the driver of the Tesla Model S was at fault. However, the analysis presented in this paper argues that rather than "driver error", the underlying cause of this tragic incident could be in fact more akin to a "designer error" implicating the design of the Autopilot feature itself. This is in line with the National Transportation Safety Boards more recent announcement that suggests systems design may have contributed to the crash. It would therefore appear that the drivers expectation of system functionality may not have matched the real life capabilities of the system. This is likely to be a product of inappropriate mental models relating to system function.

Keywords: Driving automation; human error; mental models; perceptual cycle model; schema theory

1. Introduction

The Society of Automotive Engineers (SAE) taxonomy of automation (SAE J3016, 2016) is a widely accepted industrial standard and defines the allocation of system function between the driver and automated subsystems. It ranges from Level 0 (Driver only) to Level 5 (Full automation). Level 2 (Partial automation) systems first entered the commercial marketplace in 2015 in the form of Mercedes Distronic Plus (Mercedes, 2016), Volvo's Intellisafe Autopilot (Volvo Cars, 2016) and most famously Tesla's Autopilot (Tesla Motors, 2016). These systems use a combined function approach that automate both longitudinal and lateral aspects of control as well as automating aspects of driver decision-making (Stanton et al. 1997; Eriksson & Stanton, 2017). Tesla's Autopilot specifically automates both longitudinal and lateral control, as well as being capable of performing automated lane change manoeuvres if requested by the driver. The impact of automated subsystems on driver behaviour has been extensively researched since the 1970s (e.g. Sheridan, 1970) but it is only in recent years that on-road trials have been conducted (e.g. Banks & Stanton, 2015; 2016, Endsley, 2017; Eriksson et al. 2017). These have provided some worrisome findings fuelling concerns relating to driver trust (Walker et al. 2016), complacency (e.g. Parasuraman et al. 1993; Lee & See, 2004), their ability to resume control (Stanton et al. 1997) and to perform an extended vigilance task associated with 'partial autonomy' and beyond (e.g. Molloy & Parasuraman, 1996; Stanton, 2015). Despite the undeniable benefits of vehicle automation to improve road safety, deliver mobility to all, and reduce the number of accidents occurring on the road (Stanton & Marsden, 1996; Stanton & Salmon, 2009), human-automation interaction is often overlooked throughout the design process in the pursuit of growing functional capabilities (Schaefer et al. 2016) putting the safety of drivers and other road users at risk.

The first fatal accident involving a Tesla Model S, being operated in Autopilot mode, occurred on 7th May 2016. The vehicle collided with a tractor trailer that was crossing an intersection on a highway west of Williston, Florida causing fatal harm to the driver of the Tesla. The driver of the tractor trailer was unharmed. The National Highway Traffic and Safety Administration (NHTSA) commissioned the Office of Defects Investigation to conduct a full investigation of the circumstances surrounding this incident. The findings of this investigation were published in January 2017. Data extracted from the Tesla vehicle in question revealed that the vehicle was being operated in Autopilot mode, that the Autonomous Emergency Brake system had not provided any warnings or attempt to initiate an automated braking manoeuvre and finally, that the driver had made no attempt to override the Autopilot feature by performing evasive action. Overall, the NHTSA (2017) report did not identify any design defects that could have caused the collision to occur. Instead, the preliminary report concluded that "human error" was the primary cause of the incident and speculated that the driver must have been distracted from the driving tasks for an 'extended period'.

Whilst human error can be predicted using Human Error Identification (HEI) techniques such as the Systematic Human Error Reduction and Prediction Approach (SHERPA; Embrey, 1986) and the Human Error Template (HET; Stanton et al. 2009), these techniques seek to predict and classify errors that may occur within complex systems rather than providing explanation. It is essential to understand why actions or assessments made by operators make sense at the time within the context of local rationality (Dekker, 2011). Local rationality refers to the operator goals, current knowledge and focus of attention (Reason, 1990; Dekker, 2011). Rather than “human error” being a terminology used to explain an underlying ‘cause’ of a failure, Dekker (2006) argues that “human error” should in fact be the starting point of an investigation and therefore “demands an explanation” (p.68). This is so that we can further understand how peoples decision-making and responses made sense to them at the time (Plant & Stanton, 2012). In order to do this, we must, to some extent, rely upon models and theories of human behaviour (Dekker, 2006). The Perceptual Cycle Model (PCM; Neisser, 1976) has previously been used to explain human error within a variety of high profile incidents including the Kegworth plane crash (Plant & Stanton, 2012), the Ladbroke Grove rail crash (Stanton & Walker, 2011), the Kerang rail crash (Salmon et al. 2013) and the Stockwell shooting (Jenkins et al. 2011). The PCM heavily emphasises the role of schemata arguing that human thought is closely coupled with their interaction within the world and therefore capable of exploring the principle of local rationality. This is something that current HEI techniques fail to address. As such, the authors use the principles of Schema Theory and PCM to explain the circumstances surround the recent Tesla 2016 crash within the driving automation domain. This novel approach paves way for new appraisals relating to accident causation.

1.1.Schema Theory

Schema theory dates back to the early 1900s (e.g. Head, 1920; Piaget, 1926; Bartlett, 1932) and describes how individuals form mental templates of past experiences that can be used to influence their behaviour within the subsequent world. Bartlett (1932) introduced the concept of ‘schema’ and described them as active organisations of past reactions and experiences that could be combined with information in the world to produce behaviour. Similarly, Neisser (1976) describes a schema as an organised mental pattern of thoughts and/or behaviour that can help organise our knowledge and understanding of the world. Neisser suggests that embedded schemata belong in a hierarchical structure, a viewpoint echoed by Plant & Stanton (2012) whom suggest that our knowledge should be considered as networks of information that become activated through our experience of the world. When an individual carries out a task, schemata both affects and directs the way in which they interact with and perceive the information available to them in the world as well as influencing the way in which this information is stored for future reference (Mandler, 1984). This means that schemata can allow individuals to orientate themselves towards incoming stimuli and adapt their responses to it accordingly based upon previous experience (Bartlett, 1932). If the schema is appropriate to the situation, appropriate behavioural responses are produced (Stanton et al. 2009). The initial triggering of a schema is a bottom-up (BU) process. This occurs when situations within the environment initiate the triggering of schemata that are based upon past experiences, expectations or interactions within the world. The process then becomes top-down (TD). Notably, BU and TD processes can occur simultaneously. Norman (1981) argues that if ‘triggers’ within the world are wrongly interpreted, maladaptive behaviour may occur. These culminate in slips of action or lapses in attention. Two types of schemata are proposed; genotype and phenotype (Bartlett, 1932; Neisser, 1976). Genotype schema reflect the residual structure of the mind that can go on to direct behaviour within the world. Genotype schema therefore act as the underlying template for our action responses. These templates have the possibility for continued development, but a key determinant of their development is interaction within the environment (Plant & Stanton, 2013a). In contrast, the phenotype schema reflects ‘in-the-moment’ behaviour and is exhibited through our action within the world (Stanton et al. 2009). According to Norman (1981), there are three basic genotype schema-related errors that can be used to account for the majority of errors: the activation of the wrong schemata, failure to activate appropriate schemata and a faulty triggering of active schemata. All of these error types have been found to occur within the road vehicle environment (Stanton & Salmon, 2009).

1.2.Perceptual Cycle Model

The Perceptual Cycle Model (PCM; Neisser, 1976) is based upon the idea that a reciprocal, cyclical relationship exists between an operator and the environment in which they are situated. The PCM heavily emphasises the role of schemata arguing that human thought is closely coupled with their interaction within the world. This interaction can trigger existing schemata based previous experience and interaction that can lead to anticipation over certain types of information (TD processing). Previous experience directs subsequent behaviour that attempts to interpret information available to them within the environment (BU processing). Notably, environmental experience can modify and update cognitive schemata which in turn can influence future interaction within the environment hence the reciprocal, cyclical, nature of the model (see Figure 1).

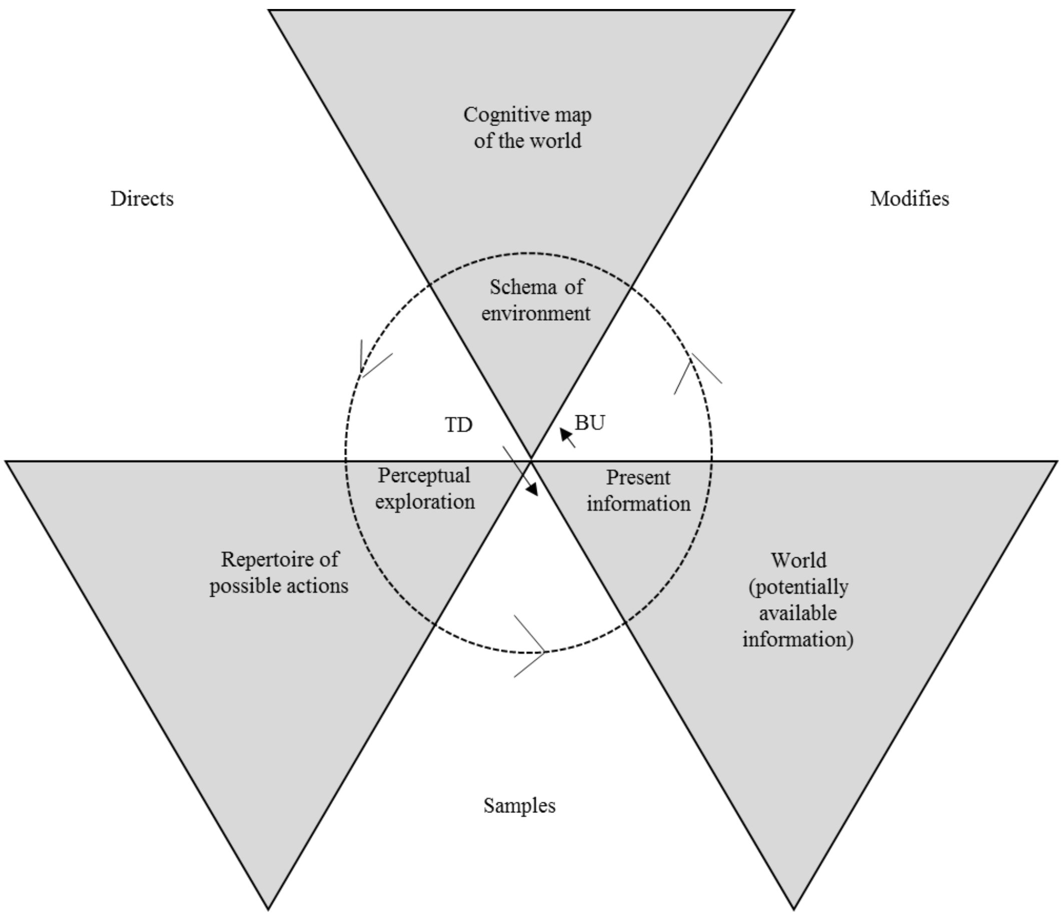

Figure 1. Perceptual Cycle Model (adapted from Neisser, 1976). Note TD: top-down processing; BU: bottom-up processing.

The PCM has been used as a means to explore systemic decision-making processes in the form of retrospective accident analysis (e.g. Stanton & Walker, 2011; Plant & Stanton, 2012). This makes the concept of construct validity particularly important (Annett, 2002). This is because in instances whereby first-hand accounts are not available, ergonomics theories must be used to propose valid explanations of behaviour post event (Salmon et al. 2013). Whilst the discussion surrounding reliability and validity is typically concerned with method selection, Plant & Stanton (2015) argue that reliability and validity are also important for in relation to theory. Whilst methods can be reliable without being valid, they cannot be valid without being reliable (Stanton & Young, 1999). This is also true for Ergonomics theories such as PCM. Plant & Stanton (2015) recognised that the validity of PCM had not been explored and so utilised the PCM to study aeronautical decision-making. They concluded that the PCM does indeed hold both construct validity and test re-test reliability and can be used for accident analysis with confidence.

2. Schematic analysis of the Fatal Tesla Crash

The role of the driver within automated driving systems has continued to be a contentious research area (e.g. Banks et al. 2018). Whilst the Society of Automotive Engineers (SAE) have gone some way in standardising the definitions relating to different levels of autonomy, the remaining responsibilities of the driver have been left open to interpretation by regulators, manufacturers and drivers alike (Banks et al. 2018). As the level of automation increases within the driving task, the driver and automated subsystems must coordinate their behaviour in order to ensure safe and normal driving practices (Banks et al. 2014). Taking a systems view, the driver and automated subsystems become analogous to ‘agents’. ‘Agents’ in this sense refer to both human and non-human entities that can receive, hold and share information with others in order to achieve a common goal (Stanton et al. 2006; Salmon et al. 2008). Stanton et al. (2006) recognise that different agents may view their environment in a different way to other agents. The concept of Distributed Situation Awareness (DSA; Stanton et al. 2006; Stanton, 2016) explains how individual situation awareness may be compatible with the awareness of other agents involved within the system. Thus, there is potential for an agent to compensate for degraded individual situation awareness held by another agent. For example, a driver (Agent 1) may not be aware of a vehicle located within their blind spot but is assisted by a Blind Spot Information System (Agent 2) that provides a warning signal located within the wing mirror. Agent-based modelling techniques (e.g. social network analysis; Baber et al. 2013; Banks et al. in press) are predominantly used to explore the interaction that takes place between different system agents in an effort to understand how agents communicate with one another. In contrast, PCM is concerned with providing a detailed explanation of thought process. A team-PCM (taking into consideration multiple system agents) therefore can demonstrate how information held within the world, the schemata held (by agents) and response actions (performed by agents) can be amalgamated to form an overall systems level perceptual cycle (Stanton et al. 2007). This makes it particularly relevant in the exploration of “why” an accident occurs. It is this approach that is used to explain the circumstances surrounding the fatal Tesla crash.

2.1.What was supposed to happen

Tesla’s Autopilot is an enhanced Level 2 automated system (SAE, 2016). The design of the system is such that the human operator remains responsible for safe operation of the vehicle. Whilst longitudinal and lateral control aspects become automated, the human operator is expected to continually monitor the behaviour of the vehicle as well as their environment (e.g. Stanton, 2015). This is so that in the event the Autopilot is unable to cope with its environment, the human operator can resume manual control of the vehicle.

A drivers mental model (e.g. Banks & Stanton, 2015; Stanton & Young, 2000) is likely to be developed through training, experience and the transparency of the system to provide reliable and accurate feedback in an accessible manner (Stanton et al. 2007). Mental models are essentially a set of expectations that enable us to interpret and predict what may happen within the driving environment. The PCM in Figure 2 represents the processes and information available within the world, the ‘intended’ schemata that drivers in Level 2 are expected to hold, and subsequent action responses that could have avoided the accident. It shows that upon activating Autopilot, the Human Machine Interface (HMI) within the Tesla would have changed from grey to blue. The driver, having built up experience in using the vehicle over the course of approximately 9 months, would have known that this meant that the vehicles’ Autopilot feature was fully engaged (i.e. their schemata of ‘Autopilot’ was activated). Both longitudinal and lateral control was now automated meaning that the driver could become figuratively speaking “hands and feet free” (Banks et al. 2014). The driver should have assumed the role of Driver Monitor (for a discussion on driver roles see Banks & Stanton, 2018) and completed the vigilance task associated with this role whilst keeping their hands on the wheel. The driver would have noticed that a tractor trailer was looking to enter the roadway ahead. With the tractor trailer moving out into the roadway ahead, the driver would have recognised this as a potential hazard. They would have monitored both the HMI (to ensure the vehicle had identified the hazard ahead) and also the behaviour of the vehicle to ensure it was taking steps to avoid collision with the vehicle ahead. At this point, we can only assume that the vehicle failed to detect the tractor trailer. Although the NHTSA (2017) report concludes that there were no functional problems that led to the subsequent accident, there is speculation that the radar and camera technology failed to detect the trailer against a brightly lit sky or that the trailer was misclassified as an overhead sign by the software. Regardless, the driver would have detected the conflict between what the HMI was showing, how the vehicle was behaving and the information that was actually available in the real world. At this point, the driver would have resumed manual control of the vehicle and the outcome may have been different.

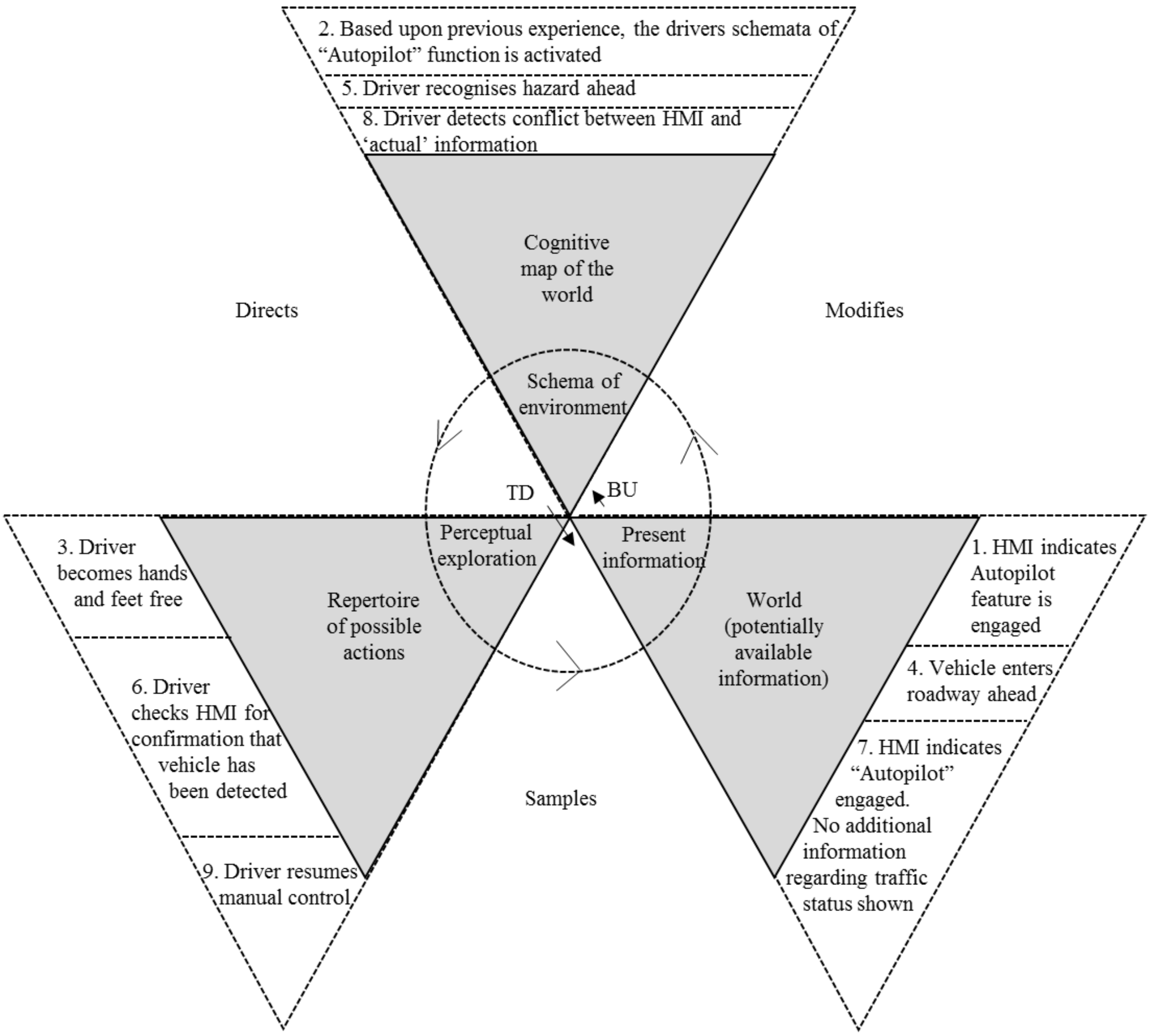

Figure 2. PCM (adapted from Neisser, 1976) depicting what should have happened

2.2. What we think happened

The team-PCM (i.e. Plant & Stanton, 2014) presented in Figure 3 was constructed based upon the NHTSA accident report (2017) and the National Transportation Safety Board (NTSB) initial specialists factual report (June, 2017a). Upon activating Autopilot, the HMI within the Tesla (i.e. the icon in the instrument cluster) would have changed from grey to blue. We know that the drivers schemata relating to Autopilot functionality was strong, especially given evidence widely available within the public domain of the driver in question testing the capabilities of system operation (e.g. YouTube, 2015). In the journey under analysis, the NTSB (2017a) revealed that Autopilot was engaged for approximately 37 minutes. During this time, the drivers' hands were detected on the steering wheel for just 25 seconds. For the remainder of this time, the drivers' hands were not detected. We therefore assume that the driver became quite literally "hands and feet free". During the journey, multiple warnings were issued to the driver via the HMI. In total, 7 visual warnings were triggered (only triggering after 60 seconds of driver inactivity). Six of these warnings escalated to an auditory tone (auditory tone triggered 15 seconds after initial visual warning). At this point, the drivers' hands were detected for 1-3 seconds at a time. Approximately 2 minutes prior to collision, the driver set the cruise speed to 74 miles per hour where it remained until just after the collision took place. The vehicle data showed that for the final 1 minute, 35 seconds of the journey, no lead vehicles were detected from the on-board sensor units. It also reveals that the driver took no evasive action immediately prior to the collision.

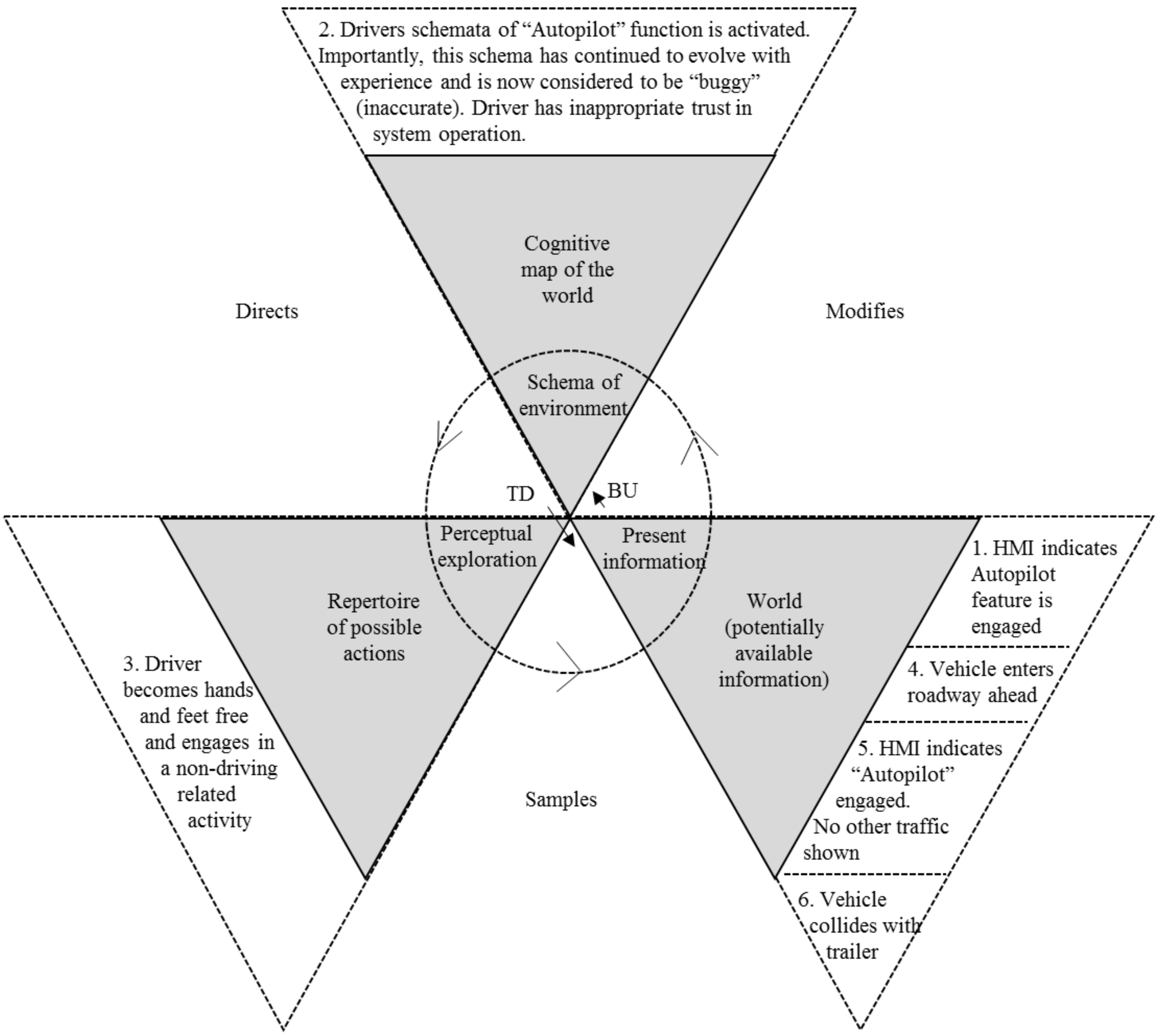

Figure 3. PCM (adapted from Neisser, 1976) depicting what we think happened

In terms of team-PCM, Autopilot provided the driver with all of the necessary information required to inform them of system state. With potential design defects being ruled out and the suggestion that the driver was distracted from the driving task for an 'extended period', NHTSA concluded that the fatal collision occurred as a result of 'driver error'. The driver failed to adhere to their responsibilities under the Level 2 definition in performing a monitoring task (SAE, 2016). Thus, we conclude from the team-PCM that both the drivers and Autopilots situation awareness was compromised at the time of the collision. If the driver had been monitoring the vehicle and environment sufficiently, they would have been in a situation to compensate for the otherwise degraded individual situation awareness held by Autopilot as the driver would have recognised the conflict between the vehicle and tractor trailer ahead – as is the essence of DSA (Stanton, 2016).

Since the initial NHTA and NTSB reports were published, the NTSB have revised their decision and criticised Tesla for allowing the Autopilot feature to be activated on roads it has not been designed for and the way in which it determines whether drivers are engaged (NTSB, 2017b). The NTSB recognises that decades of research has shown that humans are notoriously inefficient at monitoring automated systems (e.g. Casner & Schooler, 2015; Fisher et al. 2016; Molloy & Parasuraman, 1996) and provide a number of recommendations in an effort to improve system transparency and safety. However, the report still fails to explain *why* the driver felt able to relinquish control in the way that they did. Thus, concluding that an accident is partly attributable to human error fails to provide a causal explanation (Plant & Stanton, 2012).

There are a number of reasons why the driver may have felt able to relinquish their full control over vehicle operation to the vehicle. For example, in order to ensure that a system is commercially viable, it must be accurate, reliable, predictable and dependable (Donmez et al. 2006; Eriksson & Stanton, 2017). However, a system that is reliable for 99.9% of the time actually has a negative effect on the formation of mental models relating to its capabilities (Groeger, 1997; Stanton & Young, 2000; Heikoop et al. 2016). In this instance, the drivers cognitive map of Autopilot would have been developed and therefore skewed because of their experiences in using an otherwise seemingly accurate, reliable, predictable and dependable system. We already know from the literature that if individuals are exposed to systems that exhibit such positive behaviours for prolonged periods, they begin to become complacent in its operation (e.g. Parasuraman et al. 1993; Lee & See, 2004; Hollnagel & Woods, 2005). The consequence of inappropriate mental models is the emergence of mode confusion (e.g. Banks & Stanton, 2015; Endsley, 2017; Sarter & Woods, 1995; Stanton & Marsden, 1996) and operational errors (e.g. Young et al. 2007; Stanton et al. 2011) in instances whereby the automated system behaves in a manner inconsistent with predetermined expectations of system operation. In terms of schema theory, it would appear that the driver fell foul to one of the basic genotype schema-related errors outlined by Norman (1981). In this situation, the driver failed to activate appropriate schemata following the activation of Autopilot thus culminating in maladaptive phenotype schema (i.e. prolonged disengagement from the driving task).

The perception of responsibility therefore becomes an important consideration in the design of future automated systems (i.e. the role of the driver within automated driving systems). Lower levels of automation, such as SAE Level 1 (‘Assisted automation’), provide a clear partition in task allocation between the driver and vehicle subsystems. This however becomes increasingly blurred as system complexity increases. We know that driver monitoring behaviour is heavily influenced by the capabilities of the system itself (i.e. the level of automation in which it operates; Kircher et al. 2014) and also the perception of system capabilities (i.e. driver expectations). As this division becomes increasingly blurred, drivers may feel that the onus of responsibility is now shared with their automated counterparts

(Banks et al. 2014), or in extreme cases, feel that the onus of responsibility lies firmly with the automated counterpart (Hoc et al. 2009). For example, if drivers believe the system is more capable than it actually is, they may become completely disengaged from their monitoring responsibilities and/or fail to respond appropriately (e.g. Hancock, 2013; Banks & Stanton, 2018). This appears to have been the case for the driver involved in the fatal Tesla incident signalling a catastrophic "designer error" (Chapanis, 1999) whereby the boundaries between human and automated feature responsibilities and capabilities lacked the transparency required to ensure clear roles and task divisions. The concept of "designer error" was first coined by Chapanis (1999) when describing his research into pilots retracting the landing gear rather than the flaps when landing the 'Flying Fortress' in the Second World War.

Kyriakidis et al. (2017) caution that the main human factors challenge related to automated vehicles is managing the expectation that the driver can passively monitor systems and then rapidly resume manual control. By overlooking human-automation relationships, we miss an opportunity to learn about the underlying causation of accidents and near-misses within the automated driving domain. Schema theory and the PCM appears to be a promising avenue to further explore the context of human error in driving automation. The authors argue that rather than "driver error", the primary cause of the fatal Tesla accident could actually centre upon a much larger "designer error" (Chapanis, 1999). This means that the operational design of Autopilot at the time of the fatal crash unintentionally permitted the driver to disengage from the driving task for a prolonged period (NTSB, 2017).

Of course, the application of schema theory is not without its criticisms (Plant & Stanton, 2013a). Walker et al. (2011) acknowledge that 'gaining insight into mental representations…is experimentally and conceptually challenging' (p.879). The PCMs presented in this paper are subjective and based upon the inferences made from the official accident investigation reports published by NHTSA (2017) and the National Transportation Safety Board (2017). As such, they cannot be validated due to the nature of the case study discussed. However, the purpose of this paper was to explore the utility of schema theory and PCM to explore driver error in the context of driving. In other instances, the Critical Decision Method (CDM; Klein et al. 1989) has been successfully applied as a qualitative research method to elicit information relating to the role of PCM in aeronautical decision-making (Plant & Stanton, 2013b). This is an avenue of further investigation recommended for the context of driving by the author of this paper.

Even so, enhanced Level 2 systems appear to be accompanied by an increased risk of unsafe driver behaviours and designers need to do more to limit such hazardous behaviours (Brill et al. 2016, Banks et al. 2018). Proactive measures such as disabling automated features, could be one possible strategy. In fact, such a strategy was adopted by Tesla following the fatal incident. The company introduced an 'Autopilot strikeout' system in which the feature is disabled if the driver fails to respond to system warnings adequately. It remains disabled for the remainder of the drive (or until the vehicle is put into 'park mode'). Despite this, the design of the system itself remains problematic. The automation of longitudinal and lateral control, when implemented in a manner similar to that of Tesla's Autopilot, suggests an increased level of autonomy, beyond 'partial automation'. This disconnect between driver expectation and operational functionality is likely to contribute to the development of an underlying 'trust problem' whereby drivers begin to over-trust automated systems due to inaccurate mental models relating to system operation. Part of this may be attributable to current methods of marketing and deployment. Shladover (2016) argues that the inconsistent use of ‘choice’ words can carry a multitude of interpretations leading to confusion over what the driver *and* vehicle can and cannot do. Thus, in some instances, products may promise more than they can actually deliver (Stilgoe, 2017). However, it is these inaccuracies that have paved the way for errors to be made by the driver. This is because mismanagement of driver expectation may impact upon subsequent driver risk perception (Brill et al. 2016).

Automation is therefore deemed most dangerous when it behaves in a consistent and reliable manner for most of the time (Norman, 2015). This is because drivers are most at risk of becoming complacent (Lee & See, 2004) leading to a mismatch in the SA held by the driver and by the wider driving system (Stanton et al. 2017). Thus, the consequence of “designer error” in this context, is the subsequent occurrence of “driver error” via forms of automation ‘misuse’ and/or ‘disuse’ (Parasuraman & Riley, 1997).

3. Conclusions

Human error is inescapable (Fedota & Parasuraman, 2010) and consistently implicated in accident investigations (Rasmussen, 1990; Stanton & Salmon, 2009). However, many of these investigations fail to (i) explore *why* error occurs in the first place and (ii) recognise that incidents typically arise of as a combination of technical, systemic and human factors (Woods et al. 2010). Concluding that an accident is caused by human error does not provide a causal explanation (Plant & Stanton, 2012; Richardson & Ball, 2009). Schema theory, as represented by the PCM, provides an interesting framework to explore and analyse errors in context. This is because it can provide a human-in-the-system account and can therefore provide a more complete understanding of human-system relationships. Previous research has also shown that PCM provides a valid theoretical framework in which to analyse accidents (Plant & Stanton, 2015).

Whilst prediction and prevention are the primary focus of accident analysis, Plant & Stanton (2012) argue that prediction and prevention can only be truly achieved if a thorough understanding of why an accident occurred is determined. This is where traditional HEI methods fall short. This paper suggests that schema theory, along with the PCM can fill the apparent gap in knowledge. As Aurino (2000) challenged the aviation industry, the authors of this paper also challenge vehicle manufacturers to be more proactive in their safety research. Human error should be viewed as a symptom of safety breakdowns rather than being viewed as a cause. As such, the PCM may provide a contemporary approach to safety as it recognises that environmental conditions, along with their interaction with schemata, can influence operator action. Arguably, the original design of the Autopilot system set the driver up to fail because of the requirement for drivers to continually monitor a highly reliable system (Stanton, 2015). The literature openly reports that humans are not very good at doing this (e.g. Molloy & Parasuraman, 1996). This presents a real challenge for automotive manufacturers as drivers must remain engaged even when the vehicle is being driven by a computer. Banks et al. (2018) suggest that this conundrum is unsolvable and that we should perhaps focus our efforts instead on Level 4 and 5 automation.

References

Annett, J. (2002). A note on the validity and reliability of ergonomics methods. Theoretical Issues in Ergonomics Science, 3(2), 228-232.

Aurino, D. E. M. (2000). Human factors and aviation safety: what the industry has, what the industry needs. Ergonomics, 43(7), 952-959.

Baber, C., Stanton, N. A., Atkinson, J., McMaster, R., & Houghton, R. J. (2013). Using social network analysis and agent-based modelling to explore information flow using common operational pictures for maritime search and rescue operations. Ergonomics, 56(6), 889-905.

Banks, V. A., Eriksson, A., O'Donoghue, J., & Stanton, N. A. (2018). Is partially automated driving a bad idea? Observations from an on road study. Applied Ergonomics, 68, 138-145.

Banks, V. A., & Stanton, N. A. (2015). Discovering driver-vehicle coordination problems in future automated control systems: Evidence from verbal commentaries. Procedia Manufacturing, 3, 2497-2504.

Banks, V. A., & Stanton, N. A. (2016). Keep the driver in control: Automating automobiles of the future. Applied Ergonomics, 53(B), 389-395.

Banks, V. A., & Stanton, N. A. (2018). Analysis of driver roles: Modelling the changing role of the driver in automated driving systems using EAST. *Theoretical Issues in Ergonomics Science.*

Banks, V. A., Stanton, N. A., Burnett, G., Hermawati, S. (in press). Distributed Cognition on the road: Using EAST to explore future road transportation systems. Applied Ergonomics.

Banks, V. A., Stanton, N. A., & Harvey, C. (2014). Sub-systems on the road to vehicle automation: Hands and feet free but not 'mind' free driving. *Safety Science, 62*, 505-514.

Bartlett, F. C. (1932). Remembering: a study in experimental and social psychology. Cambridge: Cambridge University Press.

Brill, J. C., Bliss, J. P., Hancock, P. A., Manzey, D., Meyer, J., & Vredenburgh, A. (2016). Matters of ethics, trust, and potential liability for autonomous systems. Human Factors, 60(1), 308-312.

Casner S. M., & Schooler, J. W. (2015). Vigilance impossible: Diligence, distraction, and daydreaming all lead to failures in a practical monitoring task. Consciousness and Cognition, 35, pp 33–41.

Chapanis, A. (1999). The Chapanis chronicles: 50 years of human factors research, education, and design. Aegean Pub.

Dekker, S. W. (2006). The Field Guide to Understanding Human Error. Aldershot, Ashgate.

Dekker, S. W. (2011). What is rational about killing a patient with an overdose? Enlightenment, continental philosophy and the role of the human subject in system failure. Ergonomics, 54(8), 679-683.

Donmez, B., Boyle, L. N., & Lee, J. D. (2007). Safety implications of providing real-time feedback to distracted drivers. Accident Analysis and Prevention, 39(3), 581-90.

Embrey, D. E. (1986). SHERPA: A systematic human error reduction and prediction approach. In Proceedings of the International Topical Meeting on Advances in Human Factors in Nuclear Power Systems, Knoxville, Tennessee.

Endsley, M. R. (2017). Autonomous Driving Systems: A Preliminary Naturalistic Study of the Tesla Model S. Journal of Cognitive Engineering and Decision Making, DOI: 10.1177/1555343417695197

Eriksson, A., & Stanton, N. A. (2017). Take-over time in highly automated vehicles: non-critical transitions to and from manual control. Human Factors. 59(4), 689-705.

Fedota, J. R., & Parasuraman, R. (2010). Neuroergonomics and human error. Theoretical Issues in Ergonomics Science, 11(5), 402-421.

Fisher, D. L., Lohrenz, M., Moore, D., Nadler, E. D., Pollard, J. K. (2016). Humans and intelligent vehicles: The hope, the help, and the harm. IEEE Transactions on Intelligent Vehicles, 1(1), 56-67.

Groeger, J. A. (1997). Memory and Remembering: Everyday memory in Context. Longman, London.

Hancock, P. A. (2013) In search of vigilance. American Psychologist, 68(2), 97-109.

Head, H., 1920. Studies in neurology. Vol. I, London: Oxford University Press.

Heikoop, D. D., De Winter, J. C. F., Van Arem, B., & Stanton, N. A. (2016). Psychological constructs in driving automation: A consensus model and critical comment on construct proliferation. Theoretical Issues in Ergonomics Science, 17, 284–303

Hoc, J. M., Young, M. S., & Blosseville, J. M. (2009). Cooperation between drivers and automation: implications for safety. Theoretical Issues in Ergonomics Science, 10(2), 135–160.

Hollnagel E., & Woods, D. D. (2005). Joint Cognitive Systems Foundations of Cognitive Systems Engineering, CRC Press.

Jenkins, D., Salmon, P.M., Stanton, N.A., Walker, G.H., & Rafferty, L. (2011). What could they have been thinking? How sociotechnical system design influences cognition: a case study of the Stockwell shooting. Ergonomics, 54(2), 103–119.

Kircher, K., Larsson, A., & Hultgren, J. A. (2014). Tactical Driving Behavior With Different Levels of Automation. Ieee Transactions on Intelligent Transportation Systems, vol 15, no 1, pp 158-167.

Klien, G. A., Calderwood, R., & Macgregor, D. (1989). Critical decision method for eliciting knowledge. IEEE Transactions on Systems, Man and Cybernetics, 19(3), 462– 472.

Kyriakidis, M., de Winter, J, C, F., Stanton, N. A., Bellet, T., van Arem, B., Brookhuis, K., Martens, M. H., Bengler, K., Andersson, J., Merat, N., Reed, N., Flament, M., Hagenzieker, Happee, R. (2017). A human factors perspective on automation driving. Theoretical Issues in Ergonomics Science, DOI: 10.1080/1463922X.2017.1293187.

Lee, J. D. & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. I, 46, 50–80.

Mandler, J. M. (1984). Stories Scripts and Scenes: Aspects of Schema Theory. Lawrence Erlbaum Associates, Hillsdale, NJ.

Molloy, R., & Parasuraman, R. (1996). Monitoring an automated system for a single failure: Vigilance and task complexity effects. Human Factors, 38, 311–322.

National Highway Traffic Safety Administration. (2017). ODI Resume. Available at https://static.nhtsa.gov/odi/inv/2016/INCLA-PE16007-7876.PDF

National Transportation Safety Board. (2017a). Driver assistance systems specialists factual report. Available at: https://dms.ntsb.gov/pubdms/. Accessed 9 Aug 2017.

National Transportation Safety Board. (2017b). Collision between a car operating with automated vehicle control systems and a tractor-semitrailer truck near Williston, Florida, May 7, 2016. Available at: https://dms.ntsb.gov/pubdms/. Accessed 9 Aug 2017.

Neisser, U. (1976). Cognition and Reality. W.H. Freeman and Company, San Francisco.

Norman, D. A. (2015). The human side of automation. In Road Vehicle Automation 2 (pp. 73-79). Springer International Publishing.

Norman, D.A., 1981. Categorization of action slips. Psychological Review, 88, 1–15.

Parasuraman, R., Molloy, R., & Singh, I. L. (1993). Performance consequences of automation-induced ‘complacency’. The International Journal of Aviation Psychology, 3, 1-23.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. Human Factors, 39(2), 230-253.

Piaget, J. (1926). The child's conception of the world. New York: Harcourt, Brace

Plant, K. L., & Stanton, N. A. (2012). Why did the pilots shut down the wrong engine? Explaining errors in context using Schema Theory and the Perceptual Cycle Model. *Safety Science*, *50*(2), 300-315.

Plant, K. L., & N. A. Stanton. (2013a). The explanatory power of schema theory: Theoretical foundations and future applications in ergonomics. Ergonomics, 51(6), 1-15.

Plant, K. L., & N. A. Stanton. (2013b). What is on your mind? Using the perceptual cycle model and critical decision method to understand the decision-making process in the cockpit. Ergonomics, 56(8), 1232-1250.

Plant, K. L., & Stanton, N. A. (2014). All for one and one for all: Representing teams as a collection of individuals and an individual collective using a network perceptual cycle approach. International Journal of Industrial Ergonomics, 44(5), 777-792.

Plant, K. L., & Stanton, N. A. (2015). The process of processing: Exploring the validity of Neisser's perceptual cycle model with accounts from critical decision-making in the cockpit. Ergonomics, 58(6), 909-923.

Rasmussen, J., Nixon, P., & Warner, F. (1990). Human error and the problem of causality in analysis of accidents. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 449-462.

Reason, J. (1990). Human Error. Cambridge: Cambridge University Press.

Richardson, M., & Ball, L. J. (2009). Internal representations, external representations and ergonomics: Towards a theoretical integration. Theoretical Issues in Ergonomics Science, 10(4), 335-376.

Society of Automotive Engineers. (2016). Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, J3016_201609: SAE International.

Salmon, P. M., Read, G. J., Stanton, N. A., & Lenné, M. G. (2013). The crash at Kerang: Investigating systemic and psychological factors leading to unintentional non-compliance at rail level crossings. Accident Analysis & Prevention, 50, 1278-1288.

Salmon, P. M., Stanton, N. A., Walker, G. H., Baber, C., Jenkins, D. P., McMaster, R., & Young, M. S. (2008). What really is going on? Review of situation awareness models for individuals and teams. *Theoretical Issues in Ergonomics Science*, *9*(4), 297-323.

Sarter, N. B., & Woods, D. D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. Human Factors, vol 37, pp. 5–19.

Schaefer, K. E., Chen, J. Y. C., Szalma, J. L., & Hancock, P. A. (2016). A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. *Human Factors*, 58(3), 377-400.

Shladover, S. E. (2016). The truth about "self-driving" cars. Scientific American, 314(6), 52-57.

Sherdian, T. B. (1970). Big brother as driver: New demands and problems for the man at the wheel. Human Factors, 12, 95-101.

Stanton, N. A. (2015). Responses to autonomous vehicles. Ingenia, 62: 9.

Stanton, N. A. (2016). Distributed Situation Awareness. Theoretical Issues in Ergonomics Science, 17(1), 1–7.

Stanton, N. A., Dunoyer, A., & Leatherland, A. (2011). Detection of new in-path targets by drivers using Stop & Go Adaptive Cruise Control. Applied Ergonomics, 42(4), 592-601.

Stanton, N. A., & Marsden, P. P. (1996). From fly-by-wire to drive-by-wire: Safety implications of automation in vehicles. Safety Science, 24(1), 35-49.

Stanton, N. A., & Salmon, P. M. (2009). Human error taxonomies applied to driving: A generic driver error taxonomy and its implications for intelligent transport systems. Safety Science, 47(2), 227-237.

Stanton, N. A., Salmon, P., Harris, D., Marshall, A., Demagalski, J., Young, M. S., Waldmann, T., & Dekker, S. (2009). Predicting pilot error: testing a new methodology and a multi-methods and analysts approach. Applied Ergonomics, 40(3), 464-471.

Stanton, N. A., Salmon, P. M., Walker, G. H., & Jenkins, D. (2009). Genotype and phenotype schemata and their role in distributed situation awareness in collaborative systems. Theoretical Issues in Ergonomics Science, 10(1), 43-68.

Stanton, N. A., Salmon, P. M., Walker, G. H., Salas, E., & Hancock, P. A. (2017). State-of-science: Situation awareness in individuals, teams and systems. Ergonomics, 60(4), 449-466.

Stanton, N.A., Stewart, R., Harris, D., Houghton, R.J., Baber, C., McMaster, R., Salmon, P. (2006). Distributed situation awareness in dynamic systems: Theoretical development and application of an ergonomics methodology. Ergonomics 49(12–13), 1288–1311.

Stanton, N. A., & Walker, G. H. (2011). Exploring the psychological factors involved in the Ladbroke Grove rail accident. *Accident Analysis and Prevention, 43*, 1117–1127.

Stanton, N. A., & Young, M. S. (1999). What price ergonomics? Nature, 399, 197-198.

Stanton, N. A., & Young, M. S. (2000). A proposed psychological model of driving automation. Theoretical Issues in Ergonomics Science, 1(4), 315-331.

Stanton, N., Young, M., & McCaulder, B. (1997). Drive-by-wire: the case of mental workload and the ability of the driver to reclaim control. Safety Science, 27(2-3), 149-159.

Stanton, N. A., Young, M. S., & Walker, G. H. (2007). The psychology of driving automation: a discussion with Professor Don Norman. International journal of vehicle design, 45(3), 289-306.

Stilgoe, J. (2017). Tesla crash report blames human error – This is a missed opportunity. The Guardian. Retrieved from https://www.theguardian.com/science/political-science/2017/jan/21/tesla-crash-report-blames-human-error-this-is-a-missed-opportunity. Accessed 14 Aug 2017.

Walker, G. H., Stanton, N. A., & Salmon, P. M. (2011). Cognitive compatibility of motorcyclists and car drivers." Accident Analysis & Prevention, 43, 878– 888.

Walker, G. H., Stanton, N. A., & Salmon, P. (2016). Trust in vehicle technology. International Journal of Vehicle Design, 70(2), 157-182.

Woods, D. D. (2010). Behind Human Error. Ashgate Publishing, Ltd.

Young, M. S., Stanton, N. A., & Harris, D. (2007). Driving automation: learning from aviation about design philosophies. International Journal of Vehicle Design, 45(3), 323-338.

YouTube. (2015). https://www.youtube.com/watch?v=5TjbqVartjM. [Accessed 03.07.2017].