Ergonomics

ISSN: 0014-0139 (Print) 1366-5847 (Online) Journal homepage: https://www.tandfonline.com/loi/terg20

Pitfalls of automation: a faulty narrative?

Commentary on Hancock (2019) Some pitfalls in the promises of automated and autonomous vehicles

J. C. F. de Winter

To cite this article: J. C. F. de Winter (2019) Pitfalls of automation: a faulty narrative?, Ergonomics, 62:4, 505-508, DOI: 10.1080/00140139.2019.1563334

To link to this article: https://doi.org/10.1080/00140139.2019.1563334

Published online: 07 Apr 2019.

ERGONOMICS
2019, VOL. 62, NO. 4, 505–508
https://doi.org/10.1080/00140139.2019.1563334

Taylor & Francis
Taylor & Francis Group

COMMENTARY

OPEN ACCESS Check for updates

Pitfalls of automation: a faulty narrative?

Commentary on Hancock (2019) Some pitfalls in the promises of automated and autonomous vehicles

J. C. F. de Winter

Department of Cognitive Robotics, Delft University of Technology, Delft, The Netherlands

In 2018, Elon Musk claimed that the media ignore the benefits of automation and that the perils of automation receive disproportionate amounts of negative attention (Tesla 2018). At present, there are insufficient data to comment on whether Tesla's current Autopilot, or other types of (partially) automated driving systems, are safe or unsafe compared to manual driving, nor is the present commentary about Tesla per se. The quotes from Musk, Table 1, are included to illustrate the possible dangers of an overly sceptical attitude towards nascent technology.

In this brief commentary, it is suggested that an analogue of the problems raised by Musk is existent in the Human Factors and Ergonomics (HF/E) community, as exemplified by Hancock's (2019) contribution. I have also been guilty of using the same narrative, which goes as follows:

> Beware of the pitfalls of automation: Humans are unable (or 'magnificently disqualified'; Hancock 2019) to remain situationally aware and attentive, and to take over control when the automation cannot handle a particular situation.

Along with this narrative, it is usually cautioned that automation does not replace human activity, but rather changes it (Parasuraman, Sheridan, and Wickens 2000) while classical works, such as Bainbridge's (1983) 'Ironies of automation', are invoked as well. Herein, it is argued that, although there is some truth in the above narrative, it is not fully rational.

As correctly pointed out by Musk, 1.2 million fatal road traffic accidents occur every year, mostly due to human error. It is plausible that automation, that is, the removal of the driver from the control loop, is the remedy to this public health problem. Although other solutions, such as improved driver training, stricter police enforcement, improved crashworthiness of cars, and safer roads have been tried for many decades, automation technology is perhaps the only viable candidate for preventing crashes where drivers fail to respond to hazards or lose control of their vehicle. Autonomous emergency braking (AEB), a basic form of automation, already helps to reduce crashes by substantial amounts (Cicchino 2017).

Assuming that technology keeps evolving and automated driving systems will become more capable, the future of automation could be summarised using the Venn diagram shown in Figure 1. This figure illustrates that automation prevents accidents (the hashed segment), but also causes new accidents (the dotted segment). The relative sizes of the white and dotted segments are open to discussion and depend on the state of technology and the level of automation. The main argument here is that Hancock (2019) overemphasises the impact of the pitfalls (the dotted segment), while not lauding or even mentioning the benefits that automation may offer (the hashed segment).

I pointed out above that the potential benefits of automation (the hashed segment) are *large*. There are also several reasons to believe that the number of accidents due to the pitfalls of automation (the dotted segment) may be *small*. Research in automated driving shows that automation can even increase situation awareness. In the military domain, McDowell et al. (2008) found that operators detected targets along the route faster and reported lower workload in automated driving as compared to manual driving. This effect is sensible: if the human does not have to drive manually

CONTACT J. C. F. de Winter j.c.f.dewinter@tudelft.nl Department of Cognitive Robotics, Delft University of Technology
This article has been republished with minor changes. These changes do not impact the academic content of the article.

anymore, more attentional resources are available, which can be devoted to scanning the task environment. Of course, normal drivers are not like military drivers, as they may be less trained and not motivated to stay alert. Still, it is true that automation can complement human driving (providing the driver with 'extra eyes') and help reduce task demands; automation does *not* appear to be "*a formula for extreme stress*", as Hancock called it.

Although there are cases where automation causes low situation awareness because automation enables drivers to engage in non-driving tasks (Llaneras, Salinger, and Green 2013), it should be noted that problems of situation awareness (e.g., distraction, inattention, poor hazard perception) are also present in manual driving. In manual driving, if a driver fails to see a stationary obstacle ahead, this is guaranteed to result in a crash. In automated driving, the same loss of situation awareness results in a crash only if the automation cannot handle the situation at the same time. There are two types of solutions to automation-induced accidents. One solution is to stimulate drivers to remain alert and engaged while supervising the automated driving system (Cabrall et al. in press), for example, by providing a warning when sensors detect that the driver does not periodically touch the steering wheel or when eye-tracking cameras detect that the driver is distracted. However, the real solution is to improve vehicle sensors and software, so that the same accident will not happen again to other drivers. Continual improvement is also the strategy adopted by current autonomous vehicle developers (Favarò, Eurich, and Nader 2018).

Table 1. Quotes from Elon Musk, Tesla's Q1 2018 earnings call (Tesla 2018).

'… Broadly there's over a 1 million, I think 1.2 million automotive deaths per year. And how many do you read about? Basically, none of them. However but, if it's an autonomous situation, it's headline news, and the media fails to mention that - actually they shouldn't really be writing the story, they should be writing the story about how autonomous cars are really safe, but that's not the story that people want to click on. So they write inflammatory headlines that are fundamentally misleading to the readers. It's really outrageous. …
Even if … autonomous cars were 10 times safer, so if instead of a 1 million deaths you had 100,000 deaths. There is still going to be people who will still sue and say, hey, you're responsible for the death here. And it's like, well, the 90% of people who didn't die are not suing. They're still alive, they just don't know it.
So, we've got to deal with that and then obviously regulators respond to public pressure and the press.
it's really incredibly irresponsible of any journalists with integrity to write an article that would lead people to believe that autonomy is less safe. Because people might actually turn it off, and then die.'

The familiar adage 'humans are poor monitors' seems to have been derived from laboratory tasks which have been purposefully designed to feature a very low signal-to-noise ratio. In his highly-cited work, Mackworth (1948) showed that people often miss that a light on a 100-light circle jumps forward two steps instead of one step. This type of task is hard: even blinking or a slight lapse of attention can cause an observer to miss such a signal. Mackworth's findings, however, cannot be used to prove that humans are poor at monitoring automation systems. Loeb and Binford (1963) found that people are quite good at detecting signals for prolonged periods of time (100 minutes), provided that these signals are at least somewhat salient. More specifically, participants were bad (miss rate of 70%) at detecting a steady hiss if this sound was 0.6 dB louder than background noise, but were excellent (miss rate of about 1%) when the stimulus was 2.1 dB louder than background noise. In

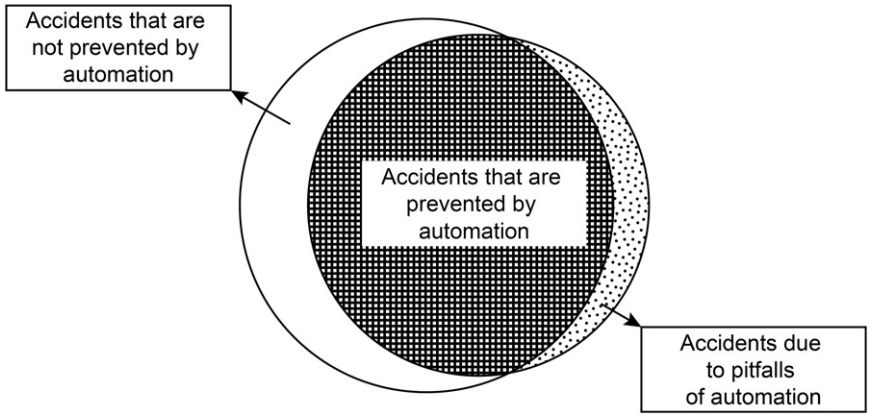

Figure 1. Venn diagram illustrating the possible future impact of automation on road safety. The left circle (the white and hashed segments combined) represents the 1.2 million fatal accidents (mostly caused by human error) when no automation is available among vehicles. When automation is deployed, a large portion of these accidents will be prevented (the hashed segment). However, there may also be new accidents due to pitfalls of automation. HF/E research is concerned with minimising the size of the dotted area.

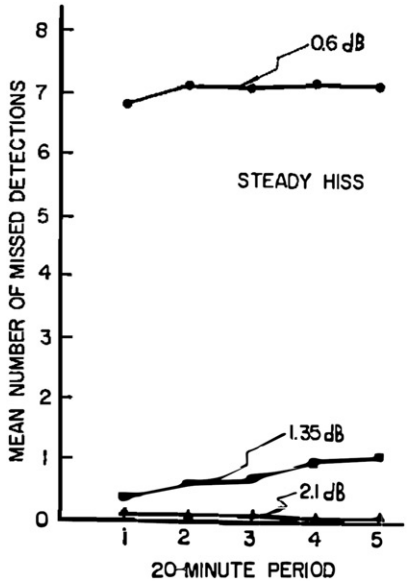

Figure 2. Mean number of misses of 18 participants, per 20-minute period. Ten signals were presented in each 20-minute period. Reprinted from Loeb and Binford (1963) with permission from the Acoustical Society of America.

summary, humans are not fundamentally unable to detect targets for prolonged periods (Figure 2).

> Hancock (2019) correctly points out that 'we should not anticipate that human society will act in a rational manner' and that 'it will be little comfort to the family of specific victim(s) … that a marginal, overall system improvement has been experienced'.

Indeed, as noted by Musk (Tesla 2018) as well, the pitfalls of automation are weighted more heavily than the benefits, and automotive companies will be wary of litigation. It is therefore foreseeable that automation will be introduced at a widespread scale only if considerably safer than manual driving. Herein, I caution that HF/E scientists should not reinforce irrationality by presenting a narrative which suggests that automated driving systems are dangerous and undesirable.

In conclusion, in this brief commentary, I argued that the typical 'pitfalls of automation' narrative may be exaggerated. Some examples from the literature were provided showing that automated driving can increase situation awareness compared to manual driving and that humans are well able to detect salient signals for prolonged periods. HF/E scientists should not downplay the potential benefits of automated driving while 1.2 million people die each year in traffic due to human error. Negatively oriented claims towards automation may stifle innovation and thereby harm public health. Similar discussions on how invalid claims about technological safety can be detrimental to safety are provided by Evans (2014) and Juma (2016).

This commentary does not aim to imply that HF/E is a useless endeavour. HF/E scientists highlight problems of human-automation interaction at an early stage and contribute to minimising the pitfalls of automation (i.e., the dotted segment in Figure 1). Here, HF/E scientists provide important contributions, such as improved human-machine interfaces. It is imaginable that HF/E research will become more relevant in the years to come, as automation becomes more capable and reliable. In other words, as the white segment in Figure 1 will become smaller and smaller, the dotted segment will become *relatively* larger, meaning that the human factor will become the last frontier of road safety (cf. ICAO 1984).

ORCID

J. C. F. de Winter http://orcid.org/0000-0002-1281-8200

References

Bainbridge, L. 1983. "Ironies of Automation." *Automatica* 19: 775–779. doi:10.1016/0005-1098(83)90046-8.

Cabrall, C. D. D., A. Eriksson, D. Dreger, R. Happee, and J. C. F. de Winter. In press. "How to Keep Drivers Engaged While Supervising Driving Automation? A Literature Survey and Categorization of Six Solution Areas." *Theoretical Issues in Ergonomics Science*. doi:10.1080/1463922X.2018.1528484.

Cicchino, J. B. 2017. "Effectiveness of Forward Collision Warning and Autonomous Emergency Braking Systems in Reducing Front-to-rear Crash Rates." *Accident Analysis and Prevention* 99: 142–152. doi:10.1016/j.aap.2016.11.009.

Evans, L. 2014. "Twenty Thousand More Americans Killed Annually Because US Traffic-safety Policy Rejects Safety Science." *American Journal of Public Health* 104 (8): 1349–1351. doi:10.2105/AJPH.2014.301919.

Favarò, F., S. Eurich, and N. Nader. 2018. "Autonomous Vehicles' Disengagements: Trends, Triggers, and Regulatory Limitations." *Accident Analysis and Prevention* 110: 136–148. doi:10.1016/j.aap.2017.11.001.

Hancock, P. A. 2019. "Some Pitfalls in the Promises of Automated and Autonomous Vehicles." *Ergonomics* doi:10.1080/00140139.2018.1498136.

Juma, D. 2016. *Innovation and its Enemies: Why People Resist New Technologies*. Oxford University Press.

International Civil Aviation Organization (ICAO). 1984. *Accident Prevention Manual*. Montreal: International Civil Aviation Organization.

Llaneras, R. E., J. Salinger, and C. A. Green. 2013. "Human Factors Issues Associated with Limited Ability Autonomous Driving Systems: Drivers' Allocation of Visual Attention to the Forward Roadway." *Proceedings of the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design* 92–98.

Loeb, M., and J. R. Binford. 1963. "Some Factors Influencing the Effective Auditory Intensive Difference Limen." *The*

Journal of the Acoustical Society of America 35: 884–891. doi:10.1121/1.1918621.

Mackworth, N. 1948. "The Breakdown of Vigilance during Prolonged Visual Search." *Quarterly Journal of Experimental Psychology* 1 (1): 6–21. doi:10.1080/17470214808416738.

McDowell, K., P. Nunez, S. Hutchins, and J. S. Metcalfe. 2008. "Secure Mobility and the Autonomous Driver." *IEEE Transactions on Robotics* 24 (3): 688–697. doi:10.1109/TRO.2008.924261.

Tesla. 2018. "Tesla (TSLA) Q1 2018 Results – Earnings Call Transcript." Accessed 8 January 2019. https://seekingalpha.com/article/4169027-tesla-tsla-q1-2018-results-earnings-call-transcript

Parasuraman, R., T. B. Sheridan, and C. D. Wickens. 2000. "A Model for Types and Levels of Human Interaction with Automation." *IEEE Transactions on Systems, Man, and Cybernetics. Part A, Systems and Humans: A Publication of the IEEE Systems, Man, and Cybernetics Society* 30 (3): 286–297. doi:10.1109/3468.844354.