Check for updates

Believe me! Why tesla's recent alleged malfunction further highlights the need for transparent dialogue

Stephen R. Milford[1,2]*, Bernice S. Elger[1] and David M. Shaw[1,3]

[1]The Institute for Biomedical Ethics, University of Basel, Basel, Switzerland, [2]Faculty of Theology, North-West University, Potchefstroom, South Africa, [3]Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, Netherlands

On November the 13th, 2022 video footage was released purportedly showing a Tesla Model Y malfunctioning by speeding through the streets of a Chinese city killing two people. Video footage such as this has the potential to undermine trust in AVs. While Tesla has responded by stating they will get to the "truth," there are questions as to how this truth is to be decided, and perhaps more importantly how the public can trust either Tesla or negative press. We explore the "facts" of the incident and discuss the challenges of building trust in new AVs systems based on transparency. In this article we argue that transparency is more than simply getting to the "truth." It is fostering a relational dialogue between the facts and stakeholder. Using O'Brien's window metaphor, this article explores the need for AV manufacturers to consider the content of such incidents, the different perceptions of stakeholders, and the medium through which the content is presented. Apart from the need for independent crash investigators, there is a need for AV manufacturers to go beyond simply' getting to the truth' and to engage with the public responsibly.

KEYWORDS

tesla, self-driving cars, transparency, ethics of crash scenarios, crash investigations

Introduction

On the 13th of November 2022 video footage was uploaded to a number of social media sites, including Weibo and YouTube showing CCTV recording of a Tesla Model Y apparently malfunctioning by speeding down the road, killing two people and injuring several more (Happy Days, 2022; Jimu News, 2022). As one can imagine, the footage quickly went viral and necessitated a formal response from Tesla who stated that it was co-operating with authorities in China, but denied its systems were at fault (Lee, 2022). One YouTube channel with only 52 subscribers on the 13th of November garnered 366 thousand views of the video in less than 48 h (Happy Days, 2022). Scrolling through the hundreds of comments one finds mixed responses with many disputing the media reports. However, many users express deep concern for self-driving cars, one user, for example, comments:

Advanced technology and computerized electronics should only be used in computer systems and in F1 for the rest we should go back to the mechanical systems operated by the driver as they used to be! unfortunately that car [Tesla] has become a lethal weapon without control. (EgiGe59 Dima in Happy Days, 2022)

That traditional cars on occasion malfunction, resulting in tragic loss of life, is well known. Investigations take place, companies are held accountable and often large sums of money are paid out to victims (Rhee and Schone, 2010; Thomas, 2010). Yet there is more at stake with autonomous vehicles (AVs). While it is true that questions remain as to AV's ultimate use and benefits (Martínez-Díaz and Soriguera, 2018; Mueller et al., 2020; Litman, 2023)—especially as they relate to vulnerable road users (Owens et al., 2019), as well as to pedestrians and cyclists (Sandt and Owens, 2017) many argue that AVs are a key disruptor technology (Medrano-Berumen and Akbaş, 2021; Nikitas et al., 2021; Othman, 2022). For some AVs offer the enticing prospect of decongested traffic, lowering public transport costs, ecological advantages, not to mention the prospect of radically reduced road deaths (Etienne, 2021).

If AVs are to be deployed on public roads, the public must be willing to accept them. This requires trust (Choi and Ji, 2015) which remains a key hurdle to widespread public adoption of AVs (Kaur and Rampersad, 2018; Raats et al., 2020; Kenesei et al., 2022). In 2020 only 10% of US respondents said they would trust a fully self-driving vehicle, with 28% being unsure (Raats et al., 2020). Incidents such as the one in question demonstrate just how vulnerable non-vehicular road users, including car operators, are and have the potential to erode public trust. To build trust and counter this erosion requires transparency (Miller, 2021). Yet what exactly happened on the 5th of November is not immediately clear.

The facts . . . apparently

CCTV footage of the incident of the 5th of November is both compelling and shocking. The short clip, lasting only about 40 s, shows a Model Y pulling over briefly before re-entering the road and attempting to overtake a scooter. Almost immediately the car rapidly accelerates to what appears to be an almost impossible speed. In a series of apparent CCTV clips captured from different angles along the route, the car races out of control. It strikes another scooter, a bicyclist (not even on the road), and drives straight through a small cargo vehicle (completely destroying it) only to immediately crash into a shop before coming to a complete halt. Reuters reports that authorities in Chaozhou City are investigating, and that Tesla itself is co-operating. Apparently the question as to the state of the 55-year-old driver during the incident has been raised with the driver's son, Mr Zhan, telling reporters that police tested his father for alcohol and drugs but that both tests were negative.

It is alleged that the driver attempted to apply the brakes but that they malfunctioned, and the car took control. Tesla, however, is claiming that data from the vehicle showed that there was no action on the part of the driver to step on the brakes "throughout the vehicle's journey" (Lee, 2022). It is intriguing that Tesla is able to make such a claim just days after the incident and with such confidence in its data. Nevertheless, it may be that the occupant was in such shock at the incident that they did not even think to try to stop the car. Or perhaps the car did malfunction and did not register the brake being depressed.

While it is not unheard of for CCTV footage to be doctored, there are some interesting aspects to the incident. For one thing, it is noted that the source of the footage is Jimu News, a Hubei government-backed official media outlet in China. The footage,

while short, is time stamped and one can determine the distance travelled by the AV over the course of the incident. Using this information one can calculate that the car travelled 2.6 km in 30 s (Dow, 2022). This puts its average speed at 312 km/h. The official top speed of a Model Y is only 250 Km/h (Tesla, 2022).

On the other hand, while Tesla is claiming that the brakes were never depressed throughout the incident–and indeed the CCTV footage appear to show they were not illuminated in most rear shots of the car–at about 23 s after the start of the incident the rear brake lights do appear to turn on very shortly (Dow, 2022). This has led to Jameson Dow of Electrek (an on-line website focusing on transport news) to conclude that: "While it is entirely possible that there is some unexamined cause here, it is almost certainly the same cause as it always is in these situations: Someone pressed the wrong pedal, and then kept pressing it when they panicked" (Dow, 2022). There is precedent for Dow's contention (ITV News, 2022).

Who is to be believed? Tesla, who apparently had performed a diagnostic of the car data so quickly after the incident, the 55-year-old driver who claims that the brakes failed and the car accelerated on its own accord, or the CCTV footage that appears to show a car far exceeding its own maximum speed?

The need for independent third-party investigations

Tesla has vowed to get to the "truth" (Oliveira, 2022) of the incident. This is welcomed as transparency and truth can help to foster trust. System transparency–to use Choi and Ji's term (2015)—is an important aspect of building public trust. It can be defined as "the degree to which users can predict and understand the operating of autonomous vehicles" (Choi and Ji, 2015, p. 694). When incidents such as the one on the 5th of November occur, the public is understandably interested in what happened. If the vehicle acted in an unpredictable manner (e.g., suddenly accelerating without apparent reason) the public may find it hard to predict the operation of an AV and consequently may struggle to trust AVs in general. However, it should not be left up to Tesla alone to determine the truth.

One way to foster public trust is to investigate incidents such as these. Yet, common car accidents are normally subjected to independent investigation, not investigation undertaken by the manufacturer. This may be a competent police accident investigator, or The National Highway Traffic Safety Administration (NHTSA). This should be no different for accidents involving vehicles such as Teslas. Take, for example, the fatal accident involving a tesla Model S in April 2021. The National Transportation Safety Board was tasked with investigating and has released regular updates (NTSB, 2021).

It is not entirely clear how, or why, Tesla was able to obtain the vehicle's data about the incident of the 5th of November 2022 so quickly. However, it seems doubtful that the public would trust an investigation conducted by an invested party that has a clear conflict of interest. If the public did not trust Toyota to investigate its own accelerator malfunctions–ultimately calling a congressional probe (CNN, 2010)—why would they trust Tesla to investigate their own cars?

It is important to note that there has been a notable failure to provide a robust forensic investigation framework for AVs (Hoque and Hasan, 2021). The US has started to move toward this position. The NHTSA, for example, has issued a standing general order that requires all manufacturers and operators of automated driving systems, as well as SAE (Society of Automotive Engineers) Level 2 and above advanced driver assistance systems, to report crashes to the agency (NHTSA, 2022b). While this is welcomed, there are questions as to whether this is adequate or not. Simply reporting a crash does not imply that the NHTSA will actually investigate the cause, nor make preventative recommendations. At present, the NHTSA's Automated Vehicle Transparency and Engagement for Safe Testing program–which enables AV test results to be made publicly available–is voluntary (NHTSA, 2022a).

Beyond simply "the truth"

Notwithstanding the need for independent assessment, merely providing the objective facts of the case may not be enough. Effective transparency is more complex than this. O'Brien (2019), for example, argues for two models of transparency. The first views transparency as merely the transmission of accurate, objective information (truth). In this model, Tesla would satisfy the need for transparency by simply explaining what data was collected during the incident and what the data indicates. Naturally this is easier said than done. AVs make extensive use of machine learning and artificial intelligence (Utesch et al., 2020; Mankodiya et al., 2022). It is quite possible that should a malfunction have actually occurred, the exact cause may be hidden deep within the car's AI "black-box" (Utesch et al., 2020; Umbrello and Yampolskiy, 2022), being inaccessible to the programmers themselves, let alone the public.

Yet even where the information is clearly available, merely disclosing it may be insufficient to foster transparency. To many lay people, the disclosure of algorithmic data may not be understandable, and would require experts to interpret the data, experts who may have a vested interest (such as Tesla's employees). O'Brien, however, presents a second model of transparency that draws attention to the social orientation of transparency. This is to say, there is a relational and dynamic nature to transparency in that it is a communicative action. Transparency itself is value-neutral, rather it is the goal of fostering trust that makes transparency valuable. In this context, transparency is not only about the facts of the case, but a dialogue that provides information to and engages with the public to co-construct interpretations and solutions. One cannot assume that everyone will interpret the information (whatever that may be) in the same way. In O'Brien's words: 'An inherent tension exists between the filtering and framing necessary to make information accessible and the desire for full disclosure of unadulterated information' (2019, p. 758). To achieve trust, one must negotiate this tension.

Consequently, it is important to bear in mind that while trust includes transparency, it goes beyond merely the "truth" (Starke et al., 2022). Humans build trust in very complex and dynamic ways. Much work has been done to investigate user trust in relation to AVs (Adnan et al., 2018; Celmer N. et al., 2018; Hegner et al., 2019; Ha et al., 2020; Hartwich et al., 2021; Waung et al., 2021). Interestingly, there is evidence to suggest that our ability to build trust with

automotive technology (such as AVs) is built upon numerous interactions over time and extends beyond a single–or limited set of–interactions one has personally with such technology. It is rooted in the broader society and culture–what Miller et al. call "situation awareness"—and includes shared mental models (2021).

In this regard, scenarios such as the one that took place on the 5th of November are important not only for those directly involved in the incident, but also the wider public as we consider the risk of erroneous decisions taken by AIs. Xu and Howard (2020) have drawn attention to the importance of AVs in this regard. They argue that in AVs the consequences of faulty AI decisions are unambiguously presented to the public. Considering that prior faulty AI agency significantly impacts trust in human-computer interactions, Xu and Howard demonstrate the importance of building trust among the public for AVs so as to build trust for other computer-human interactions. This topic has received much attention recently (Raats et al., 2020).

O'Brien's window

If we are to build trust through transparency we need to move beyond merely disclosing facts. Here O'Brien's metaphor of a window is a useful tool to draw our attention to three components of the relational dialogue between facts and stakeholders (2019, p. 758). First, O'Brien notes the scene beyond the windowpane, the content itself, i.e., the actual details of what happened as they are presented by someone. We must recognise that what is being shared and not shared are seldom the actual details of the case. Instead it matters more by who and for what purpose the scene beyond the window is described.

Take the incident in question. There are multiple actors engaged, some are powerful, some may be seeking power. Tesla may be perceived as a "Western company" with a particular agenda, while Jimu News (as an official government-backed media outlet) may have an entirely different agenda. Indeed, it is possible that the "victims" in the case (the driver) may also have an agenda, and of course let's not forget that those who shared the story are sharing the "facts" for their own purposes. So while one may ask about the facts such as did the car malfunction, or did the driver mistake the accelerator for the brake, the actors sharing the information do so with a particular goal in mind.

For companies like Tesla, this necessitates a response that is more than simply "getting to the truth" of the matter, or referencing past false allegations and cautioning "against people believing "rumors" about the incident" (Dow, 2022). To build trust in itself and its cars, Tesla needs to portray a company concerned with the public's trust in a technology that might be beneficial to the public. Rather than accusing others, denying responsibility, or cautioning the public, it might be better for Tesla to encourage an open discussion and an independent investigation. This would help the public to perceive Tesla as acting in the public's best interests rather than merely trying to protect its own investments.

The second component of O'Brien's transparency model is the viewer who peers through the windowpane. Different viewers will perceive the content differently. Each will have different knowledge, values, and assumptions about the content. They will hold different interpretations and display different emotional responses to the information supplied by the scene. For example, there is a spectrum

of opinion about human-machine systems that ranges from opposition (extreme aversion and distrust about anything related to algorithms), to loafing–extreme complacency or bias in favour of automation (Zerilli et al., 2022). Individuals on each end of this spectrum have different interpretations of the information supplied by the scene.

For Tesla this entails recognising, and engaging with, different interested parties. The shareholder is interested in company value, a Tesla owner wants to know their car is safe, a prospective Tesla client wants to know that Tesla's cars are being developed ethically, while a pedestrian wants to know that Tesla's cars are not putting them at greater risk. Each stakeholder perceives the information through their lens of interest, and each seeks information that will either affirm or challenge their underlying concerns. If Tesla is to build public trust through transparency, their message needs to address a wide audience.

The third component of O'Brien's model is the windowpane itself. The medium through which the scene is mediated affects its interpretation. This frames the scene, limits what viewers see, all the while warping, distorting, and colouring the information presented. A 30 s car ride reduced to 10 s with stark edits and excitable music, makes more for a scene out of an action film than a tragedy in which two people died. Even if the pictures show a malfunctioning AV, the manner in which they present them on social media hardly engenders trust and transparency toward either party (manufacturer, news agent, or driver).

For companies like Tesla this means presenting the same information in different formats to different audiences. It is interesting that nothing about this incident (or indeed other such incidents) is listed on Tesla's own website which includes a section for news. Rather than avoiding potentially harmful stories, Tesla may be better off incorporating information about incidents like the one that took place on the 5th of November and encouraging open discussion about it on its own website.

Discussion

The stakes are high for AVs, especially as the public slowly moves to adopt the new technology. While it is still rare to find a fully autonomous vehicle on public roads, Tesla is advertising its vehicles as being self-driving capable–although there are questions about its actual capabilities (Ivanova, 2023). Consequently, while one may not actually use the full self-driving mode of a Tesla, that it claims it is capable of full autonomy sets it apart from other road vehicles. The incident in question, especially in light of the public attention it has received, draws to mind a concern the public has with this technology. What happens when an autonomous vehicle malfunctions? What does one do when the car becomes fully in control and speeds down the road? What are the consequences of a malfunction, and perhaps more pertinently, how we can ensure it never happens again?

In these situations, the public is looking for transparency. The public wants to know that they can trust AV manufactures to properly investigate such incidents. This requires good procedures be put in place to engage with unbiased independent investigators that will gather information from multiple sources and from different angles. For example, perhaps in this case there were other eyewitnesses, mobile phones that may have captured different angles, or perhaps private CCTV footage.

Yet how Tesla responds to this allegation is perhaps more important than the content of its response. The public is looking for a particular type of transparency, a transparency that takes into account more than simply the facts. The public is seeking a transparency that engages in a conversation about the facts in the context of who presents them, who perceives them, and how they are presented. Communication in these situations is more than simply the ability to present facts and figures about algorithmic performance metrics (Zerilli et al., 2022). Ultimately it is dependent on the ability to go beyond mere explainability. This is true no matter what the facts may be. If it turns out that the incident is an elaborate hoax, then it is unhelpful for Tesla to simply point this out and move on. Worse still, where another incident to appear, it would be unhelpful for Tesla to point to previous incidents of unsubstantiated claims and dismiss public concern–as it has done so in this case by referencing previous fraudulent claims of brake malfunction (Lee, 2022).

If the potential benefits of AVs are to be realised, they must be deployed on public roads. This will only be possible if the public trusts them. Incidents such as the one that took place on the 5th of November has the potential to erode public trust in AVs. To foster trust, all parties should strive for transparency. Transparency includes "getting to the truth" (preferably through independent third party investigation), but also recognises the relational and dynamic nature of communicating those facts. As companies like Tesla strive toward building public trust through transparency they need to take into consideration the way they are perceived, that their audiences are made up of diverse stakeholders and that the medium through which they communicate influences the way the facts are perceived. Building trust through transparency entails going beyond presenting the facts. It entails a dynamic and relational engagement with the public.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

SM: Conceptualisation, Writing, Original Draft; BE: Conceptualisation, Writing, Review and Editing; DS: Conceptualisation, Writing, Review and Editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Adnan, N., Md Nordin, S., bin Bahruddin, M. A., and Ali, M. (2018). How trust can drive forward the user acceptance to the technology? In-Vehicle technology for autonomous vehicle. *Transp. Res. Part A* 118, 819–836. doi:10.1016/j.tra.2018.10.019

Celmer, N., Branaghan, R., and Chiou, E. (2018). Trust in branded autonomous vehicles and performance expectations: A theoretical framework. *Proc. Hum. Factors Ergonomics Soc.* 3, 1761–1765. doi:10.1177/1541931218621398

Choi, J. K., and Ji, Y. G. (2015). Investigating the importance of trust on adopting an autonomous vehicle. *Int. J. Human-Computer Interact.* 31 (10), 692–702. doi:10.1080/10447318.2015.1070549

CNN (2010). *Toyota president testifies before congress*. Atlanta, Georgia: CNN. Available at: http://www.cnn.com/2010/POLITICS/02/24/toyota.hearing.updates/index.html.

Dow, J. (2022). *Tesla China responds to dramatic crash that kills two*. Fremont, CA, USA: Electrek. Available at: https://electrek.co/2022/11/13/tesla-china-responds-to-dramatic-crash-that-kills-two-video/.

Etienne, H. (2021). The dark side of the "moral machine" and the fallacy of computational ethical decision-making for autonomous vehicles. *Law, Innovation Technol.* 13 (1), 85–107. doi:10.1080/17579961.2021.1898310

Ha, T., Kim, S., Seo, D., and Lee, S. (2020). Effects of explanation types and perceived risk on trust in autonomous vehicles. *Transp. Res. Part F Psychol. Behav.* 73, 271–280. doi:10.1016/j.trf.2020.06.021

Happy Days (2022). *Tesla brake failure, 2 dead in chaozhou China*. YouTube. Available at: https://www.youtube.com/watch?v=7csgV2CuKNg.

Hartwich, F., Hollander, C., Johannmeyer, D., and Krems, J. F. (2021). Improving passenger experience and trust in automated vehicles through user-adaptive HMIs: "The more the better" does not apply to everyone. *Front. Hum. Dyn.* 3. doi:10.3389/fhumd.2021.669030

Hegner, S. M., Beldad, A. D., and Brunswick, G. J. (2019). In automatic we trust: Investigating the impact of trust, control, personality characteristics, and extrinsic and intrinsic motivations on the acceptance of autonomous vehicles. *Int. J. Human-Computer Interact.* 35 (19), 1769–1780. doi:10.1080/10447318.2019.1572353

Hoque, M. A., and Hasan, R. (2021). "AVGuard: A forensic investigation framework for autonomous vehicles," in ICC 2021—IEEE International Conference on Communications (Montreal, QC, Canada: IEEE), 1–6. doi:10.1109/ICC42927.2021.9500652

ITV News (2022). *Woman pressed accelerator instead of brake before hitting children at school*. London, England: ITV News. Available at: https://www.itv.com/news/london/2022-07-26/woman-pressed-accelerator-instead-of-brake-before-hitting-children-at-school.

Ivanova, I. (2023). *Tesla engineer says company faked "full autopilot" video: Report*. New York, NY, USA: CBS News. Available at: https://www.cbsnews.com/news/tesla-autopilot-staged-engineer-says-company-faked-full-autopilot/.

Jimu News (2022). *Traffic police respond to tesla crash incident*. Hebei, China: Weibo. Available at: https://m.weibo.cn/status/4835310758792108.

Kaur, K., and Rampersad, G. (2018). Trust in driverless cars: Investigating key factors influencing the adoption of driverless cars. *J. Eng. Technol. Manag.* 48, 87–96. doi:10.1016/j.jengtecman.2018.04.006

Kenesei, Z., Ásványi, K., Kökény, L., Jászberényi, M., Miskolczi, M., Gyulavári, T., et al. (2022). Trust and perceived risk: How different manifestations affect the adoption of autonomous vehicles. *Transp. Res. Part A* 164, 379–393. doi:10.1016/j.tra.2022.08.022

Lee, L. (2022). *Tesla says it will assist police probe into fatal crash in China*. London, UK: Reuters. Available at: https://www.reuters.com/business/autos-transportation/tesla-says-it-will-assist-police-probe-into-fatal-crash-china-2022-11-13/.

Litman, T. (2023). *Autonomous vehicle implementation: Implications for transport planning*. Victoria, Canada: Victoria Transport Policy Institute.

Mankodiya, H., Jadav, D., Gupta, R., Tanwar, S., Hong, W.-C., and Sharma, R. (2022). OD-XAI: Explainable AI-based semantic object detection for autonomous vehicles. *Appl. Sci.* 12 (11), 5310. doi:10.3390/app12115310

Martínez-Díaz, M., and Soriguera, F. (2018). Autonomous vehicles: Theoretical and practical challenges. *Transp. Res. Procedia* 33, 275–282. doi:10.1016/j.trpro.2018.10.103

Medrano-Berumen, C., and Akbaş, M. İ. (2021). Validation of decision-making in artificial intelligence-based autonomous vehicles. *J. Inf. Telecommun.* 5 (1), 83–103. doi:10.1080/24751839.2020.1824154

Miller, C. A. (2021). "Trust, transparency, explanation, and planning: Why we need a lifecycle perspective on human-automation interaction," in *Trust in human-robot interaction*. Editors C. S. Nam, and J. B. Lyons (Cambridge, MA, USA: Elsevier Academic Press), 233–257. 2021-21101-011. doi:10.1016/B978-0-12-819472-0.00011-3

Mueller, A. S., Cicchino, J. B., and Zuby, D. S. (2020). What humanlike errors do autonomous vehicles need to avoid to maximize safety? *J. Saf. Res.* 75, 310–318. doi:10.1016/j.jsr.2020.10.005

NHTSA (2022a). *Automated vehicle transparency and engagement for safe testing initiative*. Washington, DC, USA: The National Highway Traffic Safety Administration. Available at: https://www.nhtsa.gov/automated-vehicle-test-tracking-tool.

NHTSA (2022b). *Automated vehicles for safety [text]*. Washington, DC, USA: The National Highway Traffic Safety Administration. Available at: https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety.

Nikitas, A., Vitel, A.-E., and Cotet, C. (2021). Autonomous vehicles and employment: An urban futures revolution or catastrophe? *Cities* 114, 103203. doi:10.1016/j.cities.2021.103203

NTSB (2021). *Electric vehicle run-off-road crash and postcrash fire*. Washington, DC, USA: National Transportation Safety Board. Available at: https://www.ntsb.gov/investigations/Pages/HWY21FH007.aspx.

O'Brien, B. C. (2019). Do you see what I see? Reflections on the relationship between transparency and trust. *Acad. Med.* 94 (6), 757–759. doi:10.1097/ACM.0000000000002710

Oliveira, A. (2022). *Out of control Tesla speeds through Chinese streets, killing two people and injuring three others | Daily Mail Online*. London, UK: Daily Mail. Available at: https://www.dailymail.co.uk/news/article-11423865/Out-control-Tesla-speeds-Chinese-streets-killing-two-people-injuring-three-others.html.

Othman, K. (2022). Exploring the implications of autonomous vehicles: A comprehensive review. *Innov. Infrastruct. Solutions* 7 (2), 165. doi:10.1007/s41062-022-00763-6

Owens, J. M., Sandt, L., Morgan, J. F., Sundararajan, S., Clamann, M., Manocha, D., et al. (2019). "Challenges and opportunities for the intersection of vulnerable road users (VRU) and automated vehicles (AVs)," in *Road vehicle automation 5*. Editors G. Meyer, and S. Beiker (Berlin, Germany: Springer International Publishing), 207–217. doi:10.1007/978-3-319-94896-6_18

Raats, K., Fors, V., and Pink, S. (2020). Trusting autonomous vehicles: An interdisciplinary approach. *Transp. Res. Interdiscip. Perspect.* 7, 100201. doi:10.1016/j.trip.2020.100201

Rhee, J., and Schone, M. (2010). *Toyota with "stuck" accelerator hits 94 MPH, driver rescued by California Highway patrol*. New York, NY, USA: ABC News. Available at: https://abcnews.go.com/Blotter/RunawayToyotas/toyota-stuck-accelerator-hits-94-mph-driver-rescued/story?id=10046912.

Sandt, L., and Owens, J. M. (2017). *Discussion guide for automated and connected vehicles, pedestrians, and bicyclists*. Kentucky, USA: Pedestrian and Bicycle Information Center.

Starke, G., Elger, B. S., van den Brule, R., and Haselager, P. (2022). Intentional machines: A defence of trust in medical artificial intelligence. *Bioethics* 36 (2), 154–161. doi:10.1111/bioe.12891

Tesla (2022). *Model Y*. Model Y. Available at: https://www.tesla.com/modely.

Thomas, D. (2010). *Toyota recalls 2.3 million vehicles over sticking accelerator pedal*. Cars.Com. Available at: https://www.cars.com/articles/toyota-recalls-2-3-million-vehicles-over-sticking-accelerator-pedal-1420663220496.

Umbrello, S., and Yampolskiy, R. V. (2022). Designing AI for explainability and verifiability: A value sensitive design approach to avoid artificial stupidity in autonomous vehicles. *Int. J. Soc. Sobotics* 14 (2), 313–322. doi:10.1007/s12369-021-00790-w

Utesch, F., Alexander, B., Fouopi, P. P., and Schießl, C. (2020). Towards behaviour based testing to understand the black box of autonomous cars, *Eur. Transp. Res. Rev.* 12 (1), 1–11. doi:10.1186/s12544-020-00438-2

Waung, M., McAuslan, P., and Lakshmanan, S. (2021). Trust and intention to use autonomous vehicles: Manufacturer focus and passenger control. *Transp. Res. Part F Psychol. Behav.* 80, 328–340. doi:10.1016/j.trf.2021.05.004

Xu, J., and Howard, A. (2020). "How much do you trust your self-driving car? Exploring human-robot trust in high-risk scenarios," in 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (Toronto, ON, Canada: IEEE Xplore Digital Library), 4273–4280. doi:10.1109/SMC42975.2020.9282866

Zerilli, J., Bhatt, U., and Weller, A. (2022). How transparency modulates trust in artificial intelligence. *Patterns* 3, 100455. doi:10.1016/j.patter.2022.100455