


Safety at the edge: a safety framework to identify edge conditions in the future transportation system with highly automated vehicles

Megan S Ryerson ¹, Carrie S Long,² Kristen Scudder,² Flora K Winston³

¹Department of City and Regional Planning & Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, Pennsylvania, USA

²Department of City and Regional Planning, University of Pennsylvania, Philadelphia, Pennsylvania, USA

³Center for Injury Research and Prevention, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA

Correspondence to

Dr Megan S Ryerson, Department of City and Regional Planning & Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104, USA; mryerson@design.upenn.edu

Received 2 May 2019
Revised 21 June 2019
Accepted 24 June 2019
Published Online First
16 July 2019

ABSTRACT

Automated driving systems (ADS) have the potential for improving safety but also pose the risk of extending the transportation system beyond its *edge conditions*, beyond the operating conditions (operational design domain (ODD)) under which a given ADS or feature thereof is specifically designed to function. The ODD itself is a function of the known bounds and the unknown bounds of operation. The *known bounds* are those defined by vehicle designers; the *unknown bounds* arise based on a person operating the system outside the assumptions on which the vehicle was built. The process of identifying and mitigating risk of possible failures at the edge conditions is a cornerstone of systems safety engineering (SSE); however, SSE practitioners may not always account for the assumptions on which their risk mitigation resolutions are based. This is a particularly critical issue with the algorithms developed for highly automated vehicles (HAVs). The injury prevention community, engineers and designers must recognise that automation has introduced a fundamental shift in transportation safety and requires a new paradigm for transportation epidemiology and safety science that incorporates *what* edge conditions exist and *how* they may incite failure. Towards providing a foundational organising framework for the injury prevention community to engage with HAV development, we propose a blending of two classic safety models: the Swiss Cheese Model, which is focused on safety layers and redundancy, and the Haddon Matrix, which identifies actors and their responsibilities before, during and after an event.

INTRODUCTION

Motivation and objective

From self-driving or highly automated vehicles (HAVs) to robotic arms for surgery, automation technology is trickling into the market and poised to transform safety-critical industries.^{1–3} Policymakers, practitioners and the public are eager to realise the benefits promised by the field of automation. Consider HAVs: transportation professionals support HAVs for their potential to improve traffic safety by fundamentally changing the roles of drivers and vehicles.^{4–6} To many, HAVs are a panacea for congestion, convenience and safety.^{4,7,8} Given the multiple crises within our transportation system—in 2017 alone, vehicle crashes killed over 37 000 individuals; transportation accounted for one-fourth of greenhouse gas emissions produced nationwide; and US commuters wasted 97 hours in

congestion^{9–11}—it is no surprise that a new mobility option that addresses many of the fundamental drawbacks of mobility has considerable momentum.

The momentum and excitement surrounding HAVs can be seen in how decision-makers from the federal level to the local level are preparing for their arrival. US Department of Transportation (USDOT) Secretary Chao cited HAVs as ‘one of the most important innovations in transportation history’.⁵ To usher in this transformative innovation, the USDOT established 10 federally designated HAV testing sites to serve as proving grounds for HAV technology, and, as of 2019, representatives from 23 states and DC have introduced or passed HAV adoption-related legislation.¹²

The zest with which HAVs are being developed and heralded by decision-makers is so acute that they may not be fully appreciating the limits and risks of technology; or worse, they may be inspiring HAV designers and the systems engineers responsible for testing and verifying the technology to work too quickly. The relationship between risk and innovation is direct—reducing risk can slow innovation; yet pushing technology too far, too quick can lead to designs that lack the necessary consideration of safety. How can we balance risk, robustness and innovation when working in safety-critical areas? With adoption of HAVs on the horizon, it is imperative to understand the *edge conditions*: situations that go beyond the reliable and accurate capability and limits of HAV technology. These situations can arise because the humans act in a way we do not expect, or because the technology (HAV and related components) functions (or malfunctions) in unexpected ways.

To create policy and programme that optimise the benefits (while reducing the risks) of HAVs, safety practitioners and researchers must understand *what* edge conditions exist and *how* they may incite failure. HAVs have an operational design domain (ODD) which specifies under which conditions an autonomous driving mode can be performed safely; any possible vehicle operation outside of an ODD defines an ‘edge condition’. But the ODD itself is a function of the known and the unknown bounds of operation. The *known bounds* are those well defined by the engineers designing the vehicle and the policymakers defining operational policies; this includes, for example, the geofence or the weather conditions under which the vehicle can operate. The *unknown bounds* are those that arise based on the assumptions on which the system is designed.



© Author(s) (or their employer(s)) 2020. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Ryerson MS, Long CS, Scudder K, et al. *Inj Prev* 2020;**26**:386–390.

For example, consider anti-lock brakes (ABS) in conventional automobiles. These brakes are intended to improve safety by preventing the lockup of the vehicle wheel when a driver engages the brakes intensely.¹³ However, ABS increased the incidence of single-vehicle, run off the road crashes; automobiles equipped with ABS were found to be 24% more likely to be involved in crashes fatal to their own occupants than non-ABS versions of the same models.¹⁴ The reason for this is that drivers not trained specifically to use ABS found their braking profiles to be erratic when operating vehicles equipped with ABS (*known bounds*); moreover, drivers familiar with ABS chose to operate their vehicles in more dangerous conditions due to the presence of ABS (*unknown bounds*).¹³

In the following two sections, we explore the critical issues, gaps in the literature, and how two classic frameworks in the injury prevention community that identify possible safety failures and thus aid in evaluating possible countermeasures—the Haddon Matrix and the Swiss Cheese Model (SCM)—must evolve to incorporate edge conditions towards informing rapidly developing HAV technology standards with safety in mind. The Haddon Matrix blended with the SCM such that there is a redundancy across actors will assist the public health community in encouraging technology developers, transportation engineers and planners, and decision-makers to study the role and efficacy of safety improvements to fundamentally change how technology is understood.

Background on automated vehicle definitions

To understand edge conditions and the ways in which they might incite failure, let us first understand automation, and the realms in which automation exists. The Society of Automotive Engineers International (SAE International) distinguishes automation standards based on vehicle functionality, and, as of 2018, explicitly the role of the driver. A brief summary of the roles follows; more information can be found in SAE's 2018 revision of *Surface Vehicle Recommended Practice*.¹⁵

Conventional, manned vehicles on the road are considered SAE Level 1 or 2, indicating that many or all features are handled by a human driver. While SAE Level 2 category vehicles might assist with certain functions, such as correctional steering or providing warnings, the driver maintains full control. At SAE Level 3, the role of the driver transforms from active and full control to passive engagement; the autonomous driving modes scan the environment for threats with the expectation that a human driver can intervene. Vehicles classified as SAE Level 3 (and beyond) utilise systems such as light detection and ranging, front and rear-facing cameras, and road-monitoring algorithms to evaluate the road conditions and provide appropriate alerts and performance.

Both SAE Levels 4 and 5 are highly automated, differentiated only by the context within which they can operate. Level 5 is fully automated driving, sustained and without conditions or constraining factors—able to ‘operate on-road anywhere that a typically skilled human driver can reasonably operate a conventional vehicle’.¹⁵ In contrast, Level 4 vehicles are more restricted versions of Level 5 HAVs, such that they may be fully autonomous over specific geographical areas (specified by a geo-fenced area), roadway types or at certain speeds.¹⁶ At SAE Levels 4 and 5, once the vehicle is in motion the role of the driver is relegated entirely as the vehicle is expected to appropriately execute all driving tasks. However, the driver plays a role prior to the engagement of autonomous mode. These roles include installing updates to the vehicle software (Levels 4 and 5) and engaging

autonomous mode when conditions permit (Level 4). Drivers are expected to engage autonomous mode when the vehicle is inside the ODD.¹⁵ While the interaction between the human driver and the vehicle will be significantly reduced with Level 4 and 5, it will not go to zero, meaning that human unpredictability and misbehaviour can influence the safe operation of HAVs.

CRITICAL ISSUES

The public, media and academic discourse surrounding automated technology often centres on potential benefits and impacts. Such discussions neglect critical issues: *What will the technology do when assumptions fail? What are the boundary conditions?* Momentum for innovation and human interference has the potential to push automated vehicles beyond their ODD limits.

Momentum for Innovation

HAV designers and test engineers, through a systems safety engineering (SSE) process for a new safety-critical device, would identify the ODD for the HAV and then test the device within and outside of that defined ODD.¹⁷ Yet, it is possible that momentum and excitement for innovation can contribute to either a tighter ODD than necessary (known bounds) or that the ODD excludes plausible scenarios and use cases that the engineering team did not consider (unknown bounds). While the process of identifying and mitigating risk of possible failures at the edge conditions is a cornerstone of SSE, SSE practitioners may not always account for the assumptions on which their risk mitigation resolutions are based. The role of product designer and systems engineer is not safety critical but instead one of innovation and detailed design. It is therefore quite possible that the goals of the public health community and the focus of HAV designers and engineers are misaligned, resulting in the emergence of safety-critical systems with latent safety issues.¹⁸

Consider, for example, the introduction of airbags. Airbags were designed with the assumption that passengers would sit in a normal, upright manner and, more importantly, all passengers would be the size of an adult male. Once deployed in vehicles, children and shorter women died sitting in the front seat; the airbag was not designed for them. Despite these deaths, airbags kept being installed in vehicles each day because the momentum of new technology deployment is incredibly hard to interrupt once it starts^{19 20}; this is a failure of unknown bounds. This scenario continued until groundbreaking research identified the relationship between airbags and child fatalities, and furthermore, achieved cross-sector collaboration to produce engineering, policy and education initiatives to eliminate the issue.²¹

With HAVs on the horizon, it is critical to better understand the limits of technology—to understand where edge conditions exist and how they may incite failure—before widespread HAV development and deployment. The safety of HAVs hinges on this understanding and on engineers' ability to identify possible use cases and fully examine the assumptions they make when defining these use cases. With automation, there is concern that the assumptions that are underlying the algorithms might lead to possible failures. As with airbags, designing HAV sensors to detect and respond to a standard human typology has been shown to cause significant safety failures. Consider that darker-skinned individuals are over 5% more likely to be struck by an automated vehicle than lighter-skin counterparts²²; an issue related to the recent admission of the Transportation Security Administration that their body scanners cannot clearly scan many black women's

hair, leading to black women being disproportionately held for additional screening.²³

The human operator may react differently than expected and dictated by HAV guidance

Human behaviour related to interaction with the HAV

Despite being limited, the safe interaction between the human and the automation software relies on HAV designers and engineers sufficiently incorporating the unpredictability of humans into their assumptions of how humans will interact with the vehicle. What happens if a passenger is constantly reprogramming the destination of a vehicle to the extent that they essentially control the route of the vehicle? Does this defeat the efficiencies of self-driving cars and cause a safety concern? Assuming people will properly use their autonomous vehicles is not supported by other technology industries; these industries have proven that planning for misuse is essential. It is not enough for automation technology to work in a law-abiding world, it must also have mechanisms to protect against misuse. Consider the serious issue of drivers operating vehicles while using their cell phones for text messages; both systems were developed separately yet today we struggle with the safety implications of their regular combined use.^{24 25}

Human behaviour related to interaction with the information communication system

In order for HAVs to operate, the information communications network has to be fully functional, fast and protected against malware and disruption. The complex interface among machine components is realised in emerging vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I; together, these are referred to as vehicle-to-everything or V2X) technology.^{26 27} With V2V, vehicles can communicate directly with each other to share information about road conditions ahead. With V2I, vehicles communicate with roadway sensors, traffic lights and other infrastructure. The physical Information and Communications Technology (ITC) Network over which V2X messages will travel becomes a new, and critical, component of the physical environment, and also a target for interference. Like many other technological systems, humans may maliciously try to invade the ITC Security System, bringing the HAV system to a halt or interfering with safety-critical processes. Humans may also try to subvert the ITC Security System for non-malicious reasons; previous experience with the deployment of advanced IT suggests travellers may interfere with the sending of messages due to privacy concerns.²⁸ Similar to the recognition that the cyber security of national power grids is a critical issue, the security of HAV ITC systems also presents a threat to the safety and operation of HAVs.²⁹

GAPS IN THE LITERATURE

While the transportation engineering and planning community have taken on the topic of HAVs from a number of dimensions, we focus on safety here. The technical aspects of safety are well explored in the more technical literature; algorithms and sensors have been the focus of extensive research among scholars and automobile companies.³⁰⁻³² In addition, scholars have considered how HAVs will operate safely on test tracks and in the real world.³³ Moreover, in the behavioural economics world, experts have surveyed potential HAV adopters to confirm that the transparency of the technical system and the ability of the HAV to manage situations build consumer trust.³⁴ Recognising that ensuring the safety of HAVs is an interdisciplinary problem,

a multidisciplinary approach has been recommended by some scholars.³⁵

However, despite the focus in the research community, the community of planning in practice is unsure of how to plan for autonomous vehicles. In the urban planning literature, an interview of 15 metropolitan planning organisations representing the largest metropolitan areas of the USA found that none had included HAVs in their long range plan.³⁶ This stands in stark contrast to the zest with which states have passed laws allowing for the operation of HAVs. This means that HAVs will be permitted but not appropriately planned for; the policy is there but the infrastructure, traffic control strategies and issues at the municipal level that directly impact day-to-day safety are not.

Similarly, there exists a critical need to develop new data sources and surveillance towards achieving early detection of—and appropriate reaction to—dangerous conditions. The current frameworks for compiling crash data and developing safety strategies and interventions, as maintained by the National Highway Traffic Safety Administration (NHTSA), must be updated to account for HAV technology and possible failures. Consider the NHTSA Fatality Analysis Reporting System (FARS), a national census of all documented vehicle crashes that resulted in a fatality, and the National Automobile Sampling System General Estimates System (GES) which establishes a representative sample of fatal crashes towards shaping safety-critical policies, legislation and interventions.^{37 38} For each fatal crash documented within the NHTSA systems, FARS collects over 140 factors related to the crash event. These factors are comprised of vehicle factors, such as vehicle model, mileage and year; involved person(s) data, such as the age or other demographic information of crash victims, or injury severity for all involved motorists and non-motorists; and surrounding environment data, such as the roadway features, weather or lighting conditions at the time of the crash. These data are selected to account for crashes involving conventional vehicles. The data, statistics and sampling processes will need to be re-evaluated to account for contributory factors unique to HAVs. FARS analyses must expand the collected data elements to account for all ODD components and technology updates; likewise, the GES will need to account for these elements within its probability sampling. Accommodating new technology within these datasets and methodologies will be particularly challenging as HAV technology is (1) proprietary and (2) reliant on frequent updates which can change or blur the ODD. However, without this information, the NHTSA and the research built on FARS and GES might not be able to accurately examine and provide insight into HAV automation failures.

The injury prevention community has long taken on transportation safety and is beginning to discuss the urgency and value in studying vehicle automation.³⁹ A proposal to adapt the Haddon Matrix, a widely accepted tool and paradigm among safety experts used to assess factors involved in crashes, for Level 3 HAVs can be found in Ryerson *et al.*³⁹ Within transportation-focused fields of injury prevention, the Haddon Matrix explores the distinct roles of drivers, technologies or vehicles, and environmental factors, and identifies the impacts and interactions of these roles on crash events. In a Haddon Matrix representation of a typical automobile crash, various actors involved in vehicle safety (encompassing the driver, the vehicle and the social/physical environments) inhabit specific and well-defined roles scaled along the discrete phases of the collision timeline. The authors³⁹ reframe the matrix for Level 3 HAVs by delineating Conventional and Autonomous components of driver, vehicle, and physical and social environment factors.

	Human	Vehicle		Physical Environment		Social Environment
		Software	Hardware	Information-Communication Networks	Transportation Infrastructure	
Pre-Event	Performance of regular vehicle software update; Human re-engagement*; Human choosing autonomous mode when truly safe*	Algorithms that can handle a high workload	HAV Computer and sensor capability; Sensors detecting ODD; Security of computer system; Procedure for outside of ODD; Vehicle design; Vehicle condition	Presence and security of connected vehicle infrastructure	Maintenance of road paint (or other identifying features); Road design; Weather Conditions	Market penetration of HAVs
During the Event	Brace position of human	Algorithm capacity to take control of the vehicle (or alert human to engage)	Vehicle alert systems; Presence of "minimal risk condition" to reach safety after failure; Advanced restraint systems; Size of vehicle; Vehicle crash rating	Speed of communications network to transmit messages	Presence of trees; Presence of guardrails	Mix of HAVs and conventional vehicles immediately proximate
Post-Event	Overall Health; Age of Victim	Adaptation of algorithms	Presence of GPS locator; Vehicle designed to withstand fire		Availability of competent EMS staff; Distance to trauma care; Availability of rehabilitation programs	Policies and funding support for emergency/medical response systems; Public support for trauma care; EMS training; Resources for psychological recovery from trauma

Figure 1 SCM/Haddon Matrix for Level 4 and Level 5 HAVs. (Note: Holes which are only relevant for Level 4 and not Level 5 are denoted with *). GPS, Global Positioning System; HAVs, highly automated vehicles; ODD, operational design domain.

Though Ryerson *et al*³⁹ move towards a new paradigm for safety, it does not encompass issues unique to Level 4 and Level 5 HAVs. The authors focus on a Haddon Matrix for Level 3 vehicles, which are vehicles where the driver must be present and ready to engage at any time. However, Levels 4 and 5 are quite different given that the driver, the vehicle and the environment are all directly participating in moving the vehicle in autonomous mode. As the Haddon Matrix defines—and delineates the roles of—the components, a Haddon Matrix for Levels 4 and 5 HAVs must allow for the consideration of how redundancy across the components can address safety concerns. In this way, contrasted with the Haddon Matrix for Level 3 vehicles, the Haddon Matrix for HAVs in Levels 4 and 5 must move towards a continuum instead of a matrix. A continuum, or shared responsibility for functions across multiple actors, allows for redundancy towards ensuring safety. Redundancy is the cornerstone of the SCM, originally pioneered by Reason.^{40 41} In the SCM, multiple layers of redundancy protect against the 'holes' of any given component or 'layer'. The application of this model to vehicle automation demands further research towards resolving the open-ended questions of: How will layers be defined, and what are the hidden-use cases and hidden assumptions input to the design of automation technology?

RECOMMENDATIONS

We suggest the creation of a framework that HAV designers and engineers and the public health and injury prevention community can use to assess the relationship between safety, automation and human behaviour. We propose a Haddon Matrix for Level 4 and Level 5 HAVs, depicted in [figure 1](#). Contrasted with traditional Haddon Matrices and the Haddon Matrix developed for Level 3 HAVs,³⁹ our proposed Haddon Matrix for Levels 4 and 5 incorporates the concept of layers from the SCM by introducing new layers, or components, and addressing the interdependency across layers.⁴² Haddon's original matrix included four actors involved in injury prevention for conventional vehicle transportation: the Driver, the Vehicle, the Physical Environment and the Social Environment. We propose the following: (1) reframing 'Driver' to 'Human'; (2) dividing both the Vehicle and the Physical Environment actors into two separate layers: Software and the Hardware for Vehicle; and Information Communications Network and Transportation Infrastructure for the Physical Environment. We identify these six layers explicitly such that we

can explore their specific countermeasures while accounting for the interdependency across layers.

There are countermeasures in each cell of the matrix, or actions the actor can take to improve or degrade safety before, during or after an event. In SCM language, these countermeasures could be more broadly thought of as a 'hole' or a weakness in the layer that can lead to a safety issue. If one of these countermeasures is not enacted, that layer is leaving a human vulnerable to injury due to a possible crash; however, enacting a countermeasure in another layer (plugging a hole) could help reduce either the possibility of a crash or the severity of that crash.

This blended SCM/Haddon Matrix allows the public health community to diagram scenarios to generate use cases which should become the foundation of HAV design and testing. Consider analysts diagramming the crash of a Level 4 HAV. They would consider (1) the human's role as presented in the matrix; (2) the role of the vehicle software and hardware, both independently and in concert; (3) the quality and security of the physical environment; (4) the influence of the social environment and (5) the relationship between components such as the vehicle hardware and ICT networks. In diagramming these components that could lead to a crash, could contribute to the severity of the crash during an event and could mitigate the effects of a crash, analysts will have identified the test case scenarios that need to be investigated to ensure HAV safety.

It is clear that a large, interdisciplinary community has a role in ensuring that crashes due to HAVs exceeding their ODDs are elevated in the discussion, planning, and engineering of HAVs. We encourage the injury prevention community to consider the countermeasures, or holes, as the beginnings of use cases and systems engineering requirements for HAVs. This Haddon Matrix provides a critical foundation on which the assumptions for human behaviour and technology function should be built to ensure safety is fully considered in the design and engineering process.

Those in the public health community must work alongside the engineering and design community to ensure safety in designs before deployment. In addition, researchers across fields should focus efforts on developing new data sources and how to grapple with the concept of technology that evolves over time, both in use and in technical capability.

The proposal to establish a new safety framework is critical to pre-empt safety issues arising from increased adoption of automated technology, including Levels 4 and 5 HAVs. A new framework aims to resolve the disconnect between innovation and risk and to allow for complete innovation while maintaining fidelity to safety considerations and needs. In doing so, we can prepare for HAVs without relying on the aftermath of failures—as with the introduction of airbags and deaths that followed—before we understand failures and how to prevent them.

FUTURE DIRECTIONS

Failure may happen in an expected way; it may occur if assumed technology capabilities exceed their true boundaries, or, it may fail in an entirely novel or unforeseen way. The paradigm proposed herein encourages the injury prevention community to approach failure through new lenses: the adaptation of the Haddon Matrix to incorporate concepts from the SCM facilitates new understandings of man-machine-environment interactions and safety implications. On a small scale, a revised model will equip engineers with the knowledge to strengthen each layer of redundancy within the technology. More broadly, safety-critical sectors will benefit from an innovative paradigm

for understanding and addressing technology and safety. Practitioners and policymakers must build on this framework and use it to explore the critical issues discussed above. To do so, partnerships between automobile manufacturers and those in the public health community must emerge to ensure that the edge conditions are fully explored and considered; these cross-sector collaborations will aid in the production of engineering, policy and education initiatives to eliminate safety issues due to known and unknown bounds.

Acknowledgements The authors would like to acknowledge the Mobility21 Transportation Center at the University of Pennsylvania and the Center for Injury Research and Prevention for their support. The findings and conclusions are those of the author(s) and do not necessarily represent the views of the University of Pennsylvania or the Children's Hospital of Philadelphia.

Contributors MSR led the organisation and framing of the manuscript, as well as the writing. CSL provided heavy writing and paper framing inputs, as well as led the development of Figure 1 (both in concept and design). KS provided research support and the development of the inputs into Figure 1. FKW provided expert guidance, shaping of the narrative and a critical review.

Funding This project was partially funded by Carnegie Mellon University's Mobility21 National University Transportation Center, which is sponsored by the US Department of transportation (69A3551747111).

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iD

Megan S Ryerson <http://orcid.org/0000-0002-8843-3286>

REFERENCES

- Dolph E, Krause C, Oleynikov D. Future robotic systems: microrobotics and autonomous robots: A comprehensive textbook. In: *Robotic-Assisted minimally invasive surgery*, 2018: 329–35.
- Liu P, Guo Q, Ren F, et al. Willingness to pay for self-driving vehicles: influences of demographic and psychological factors. *Transp Res Part C Emerg Technol* 2019;100:306–17.
- Peters BS, Armijo PR, Krause C, et al. Review of emerging surgical robotic technology. *Surg Endosc* 2018;32:1636–55.
- Eugensson A, Brännström M, Frasher D. *Environmental, safety legal and societal implications of autonomous driving systems*. Proceedings of the International Technical Conference on the Enhanced Safety of Vehicles (ESV), Seoul, 2013.
- National Highway Traffic Safety Administration. Automated driving systems: a vision for safety, 2017. Available: https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/13069a-ads2_0_090617_v9a_tag.pdf
- Riggs W. *Disruptive transport: Driverless cars, transport innovation and the sustainable city of tomorrow*. Routledge equity, justice and the sustainable City series. New York, NY: Routledge, 2019.
- Fagnant DJ, Kockelman K. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations for capitalizing on self-driven vehicles. *Transp Res Part A Policy Pract* 2015;77:167–81.
- Stern RE, Cui S, Delle Monache ML, et al. Dissipation of stop-and-go waves via control of autonomous vehicles: field experiments. *Transp Res Part C: Emerg Technol* 2018;89:205–21.
- INRIX. INRIX global traffic scorecard, 2019. Available: <http://inrix.com/scorecard/> [Accessed 4 Mar 2019].
- National Highway Traffic Safety Administration. Traffic safety facts research note: 2017 fatal motor vehicle crashes: overview. (traffic safety facts research note dot HS 812 603); 2018. <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812603>
- Porter CD, Brown A, DeFlorio J, et al. Transportation energy futures series: effects of travel reduction and efficient driving on transportation: energy use and greenhouse gas emissions, 2013. Available: <https://doi.org/10.2172/1069182>
- Autonomous vehicles state bill tracking database. (n.d.). Available: <http://www.ncsl.org/research/transportation/autonomous-vehicles-legislative-database.aspx> [Accessed 5 Mar 2019].
- Farmer CM. New evidence concerning fatal crashes of passenger vehicles before and after adding antilock braking systems. *Accid Anal Prev* 2001;33:361–9.
- Farmer CM, Lund AK, Trempe RE, et al. Fatal crashes of passenger vehicles before and after adding antilock braking systems. *Accid Anal Prev* 1997;29:745–57.
- SAE International. Surface vehicle recommended practice: taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles, 2018. Available: <https://webstore.ansi.org/standards/sae/sae30162018> [Accessed 27 Feb 2019].
- Fraade-Blanar L, Blumenthal MS, Anderson JS, et al. *Measuring automated vehicle safety: forging a framework*. Santa Monica, CA: RAND Corporation, 2018.
- Blanchard BD, Fabrycky WJ. *Systems engineering and analysis*. 5th Edition. Upper Saddle River, NJ: Prentice Hall, 2010.
- Hernández-Orallo J, Martínez-Plumed F, Avin S. *Surveying Safety-relevant AI characteristics*. The AAAI's Workshop on Artificial Intelligence Safety (SafeAI), Honolulu, Hawaii, USA, 2019.
- Kent R, Viano DC, Crandall J. The field performance of frontal air bags: a review of the literature. *Traffic Inj Prev* 2005;6:1–23.
- Kahane CJ. Lives saved by vehicle safety technologies and associated federal motor vehicle safety Standards, 1960 to 2012 – passenger cars and LTVs – with reviews of 26 FMVSS and the effectiveness of their associated safety technologies in reducing fatalities, injuries, and crashes. (report No. dot HS 812 069). Washington, DC National Highway Traffic Safety Administration, 122; 2015.
- Durbin DR, Kallan M, Elliott M, et al. Risk of injury to restrained children from passenger air bags. *Traffic Inj Prev* 2003;4:58–63.
- Wilson B, Hoffman J, Morgenstern J. Predictive inequity in object detection. ArXiv:1902.11097 [Cs, Stat], 2019. Available: <http://arxiv.org/abs/1902.11097>
- Medina B, Frank T. TSA agents say they're not discriminating against black women, but their body scanners might be. ProPublica, 2019. Available: <https://www.propublica.org/article/tsa-not-discriminating-against-black-women-but-their-body-scanners-might-be>
- Caird JK, Johnston KA, Willness CR, et al. A meta-analysis of the effects of texting on driving. *Accid Anal Prev* 2014;71:311–8.
- Inselberg, E. United States patent No. US8401578B2, 2013. Available: <https://patents.google.com/patent/US8401578B2/en>
- Harding J, Powell GR, Yoon R, Fikentscher J. Vehicle-to-vehicle communications: readiness of V2V technology for application. (report No. dot HS 812 014). Washington, DC National Highway Traffic Safety Administration; 2014.
- National Highway Traffic Safety Administration. Federal motor vehicle safety standards; V2V Communications., PUB. L. No. 82 Fr 3854-4019; 2017.
- Fries RN, Gahrooei MR, Chowdhury M, et al. Meeting privacy challenges while advancing intelligent transportation systems. *Transportation Research Part C: Emerging Technologies* 2012;25:34–45.
- Sun C-C, Hahn A, Liu C-C. Cyber security of a power grid: state-of-the-art. *International Journal of Electrical Power & Energy Systems* 2018;99:45–56.
- Abbas H, O'Kelly ME, Rodionova A, et al. A driver's license test for driverless vehicles. *Mechanical Engineering* 2017;139:S13–S16.
- K.N, N., Kelkar P, Kumar D, et al. *Hardware optimizations for anytime perception and control*. 2015 IEEE Real-Time Systems Symposium, 2015:380.
- Mangharam R, Abbas H, Behl M, et al. *Three challenges in cyber-physical systems*. 8th International Conference on Communication Systems and Networks (COMSNETS), Bangalore, 2016:1–8.
- Kaira N, Paddock SM. Driving to safety: how many miles of driving would it take to demonstrate autonomous vehicle reliability? *Transportation Research Part A: Policy and Practice* 2016;94:182–93.
- Choi JK, Ji YG. Investigating the importance of trust on adopting an autonomous vehicle. *Int J Hum Comput Interact* 2015;31:692–702.
- Koopman P, Wagner M. Autonomous vehicle safety: an interdisciplinary challenge. *Vehicle Safety: An Interdisciplinary Challenge*. *IEEE Intelligent Transportation Systems Magazine* 2017;9:90–6.
- Guerra E. Planning for cars that drive themselves: metropolitan planning organizations, regional transportation plans, and autonomous vehicles. *J Plan Educ Res* 2016;36:210–24.
- U.S. Department of Transportation. Fatality analysis reporting system (FARS) analytical user's manual, 2018. Available: <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812602>
- National Highway Traffic Safety Administration. National automotive sampling system general estimates system analytical user's manual, 1988-2015 (No. DOT HS 812 320; p. 463). Washington, D.C: 2018.
- Ryerson MS, Miller JE, Winston FK. Edge conditions and crash-avoidance roles: the future of traffic safety in the world of autonomous vehicles. *Inj Prev* 2019;25:76–9.
- Reason J. *Human error*. New York, NY, US: Cambridge University Press, 1990.
- Reason JT. *Managing the risks of organizational accidents*. London: Routledge, 1997.
- Jeffer L, Berta W, Lingard L, et al. Learning from near misses: from quick fixes to closing off the Swiss-cheese holes. *BMJ Qual Saf* 2012;21:287–94.