

SEEING LIKE A TESLA: HOW CAN WE ANTICIPATE SELF-DRIVING WORLDS?

JACK STILGOE
Department of Science and Technology Studies
University College London
j.stilgoe@ucl.ac.uk

Abstract: In the last five years, investment and innovation in self-driving cars has accelerated dramatically. Automotive autonomy, once seen as impossible, is now sold as inevitable. Much of the governance discussion has centred on risk: will the cars be safer than their human-controlled counterparts? As with conventional cars, harder long-term questions relate to the future worlds that self-driving technologies might enable or even demand. The vision of an autonomous vehicle – able to navigate the world’s complexity using only its sensors and processors – on offer from companies like Tesla is intentionally misleading. So-called “autonomous” vehicles will depend upon webs of social and technical connectivity. For their purported benefits to be realised, infrastructures that were designed around humans will need to be upgraded in order to become machine-readable. It is vital to anticipate the politics of self-driving worlds in order to avoid exacerbating the inequalities that have emerged around conventional cars. Rather than being dazzled by the Tesla view, policymakers should start seeing like a city, from multiple perspectives. Good governance for self-driving cars means democratising experimentation and creating genuine collaboration between companies and local governments.

Keywords: Tesla, self-driving cars, automotive autonomy, risk, governance.

TEST/DRIVE

I assume I am in the wrong place. The Tesla storefront, opposite an Apple store in a Denver mall, looks more like a place to buy videogames. Inside the shop’s one room there are two cars, some T-shirts and other merchandise, screens with publicity videos and little else. The cars are well designed and immaculately polished, but it is clear that much of what is on sale is invisible. The hardware is there to see, but the pitch is all about software.

The Tesla representative assures me that I am in the right place for my test drive. She walks me through the mall and outside, where the cars are plugged in and waiting in a corner of the parking lot. I sit in the car, a Tesla Model S, the cockpit of which is dominated by a giant touchscreen. The near-silence of the electric motor further contributes to my disorientation. Attempting to maintain a convincing impression of a potential customer, I explain that my real interest is in the

ISSN 2283-7949
GLOCALISM: JOURNAL OF CULTURE, POLITICS AND INNOVATION
2017, 3, DOI: 10.12893/gjcp.2017.3.2
Published online by "Globus et Locus" at www.glocalismjournal.net



company's much-publicised "Autopilot" function. I am an Autopilot novice. I ask if it would be OK to play, if only for a minute. The Tesla employees have already reassured me that they and thousands of other happy drivers are already making use of a technology that promises something mind-blowing – the ability to cede control of one's car to the car itself.

Once I drive onto the freeway, I am told by both my human companion and the car's graphics that it is now ok to engage Autopilot. I flick a lever under the steering wheel twice and the machinery takes over with an optimistic, ascending ring-tone. A message flashes up: "Please keep your hands on the wheel. Be prepared to take over at any time", but the human says it is fine to let go. The Tesla decides to accelerate towards the car in front before adjusting its speed to maintain what it considers a safe distance. Reassuring dashboard pictures show me the car and what it detects – road markings, other cars and kerbs.

I comment on how unnerving it is to sit, hands hovering above the wheel, foot floating next to the pedals, while my car steers itself at high speed. I am promised that, while the initial moment is a leap of faith, I will quickly learn to trust it. The paradox is that, legally, I have full responsibility for the vehicle even though I have delegated all meaningful tasks to a machine. The car bends its way perfectly through the easy corners, staying in the dead centre of its lane, turning just late enough to terrify its nervous pilot/passenger. To change lanes, I flick the indicator lever down. Once the car deems it safe, it moves left and adjusts its speed.

It is clear that there are three of us in this relationship: the Tesla representative, me and the car itself. The car's advocate and I talk about the car as if it isn't party to our conversation, but we are trying to second-guess its wisdom. The car, though not a person, seems to have a personality. I am reminded of Bruno Latour's (1992) maxim that "no human is as relentlessly moral as a machine", but I find it hard to work out what this machine wants. The technology is magically exotic, surpassing my prior expectations, and yet I find myself asking, despite its superlative powers, is it good enough? Which prompts a further question: Good enough for what?

As a prospective consumer, I am not expected to care how the Tesla does what it does. I am told that the car is not yet perfect¹. It is better at detecting cars than cyclists, for example. The Tesla sales people are offering a work-in-progress. For some customers, this is part of the appeal. They have been

told that, in the coming months and years, software upgrades will provide improvements to their cars. For those who are interested, there are online explanations of what the Tesla sees through its own cameras and how this is set to improve.

A few days after my test drive, a video appears on the Tesla web site offering a car's-eye view of an Autopiloted journey through a foggy Silicon Valley suburb. The film shows a driver's hands, but he is redundant. He is there, the text explains, only because the law requires it. His hands are in his lap and his feet are on the floor as the car drives itself. Feeds from three of the car's eight cameras show the things that the car detects, coloured by type: edges of roads, signs, objects in the car's way and other objects. The software is working out, in real time, what it is seeing and how it should respond.

Here is a bravura performance of autonomy enabled by machine learning. This is no sunny Coloradan freeway. The roads are narrow and winding and the weather makes for poor visibility. The streets are filled not just with other cars, but with pedestrians, cyclists and traffic cones. The car is demonstrating a linear robotic logic: sense, plan, act. Images from its cameras and sensors are classified based on the accumulated experience of a deep neural network – an on-board super-computer whose software is the product of extensive machine learning. The computer has been formally taught what some things are and what to do in certain circumstances. But it has also taught itself using data gathered by tens of thousands of similar Teslas. The car is following rules, many of which have never been programmed into it. The approach, normally called “deep learning”, is to feed the network with so much data that it can work out what matters in order to accomplish its task.

In the video, the car makes its way through a set of four-way stop junctions. These are often locations for very human interactions: confusion, eye contact, gesturing and negotiation. Drivers in other US states talk condescendingly of the “California stop”, a form of deviance in which an injunction to “stop” is routinely interpreted as “slow down”. Tesla's car knows about such things, but is reticent to mimic them. At some moments during the voyage its behaviour seems overly nervous and polite: it comes to a dead stop alongside a pair of joggers and slows long before junctions if there is a car ahead of it. But the performance is undoubtedly impressive.

As a public experiment (Collins 1988), the video is all but useless. It is a form of simulation and, as roboticists have

pointed out, “simulation is doomed to succeed” (Brooks and Mataric 1993). However, the advertisement’s story is clear and it is one of inevitability: here is an autonomous vehicle (AV); the technology may be a work-in-progress but the end is just around the corner.

GOVERNANCE BEYOND TROLLEY PROBLEMS

In May 2016, a Tesla Model S crashed in Florida. With his car in Autopilot mode, Joshua Brown died instantly. As far as anyone knows, neither he nor his car’s cameras saw the white truck that had driven across his carriageway. If the car’s radar sensors did see the truck, which was white against a white sky background, its computer chose not to take any action. The car, travelling at 76 mph, failed to slow down. As it went under the trailer, between the truck’s wheels, its roof was torn off. Data extracted from the car suggests that for 37 and a half minutes of Brown’s 40-minute drive, his car was in Autopilot and his hands were off the wheel. Even though Tesla would not admit their car was yet a self-driving car, this first fatality could be said to reveal more about the reality of self-driving cars than any number of video simulations (Stilgoe 2017).

Ethicists, lawyers and others have been quick to point out that, even if the vision of cars could be perfected, safety questions would not vanish. Self-driving cars have been a gift to applied ethicists, for whom the dilemmas of transport-related decision making have a ready-made thought experiment in the form of the “trolley problem”. For example, faced with a decision between crashing into a bus queue of five bystanders or ending its driver’s life, what should a car’s algorithms choose? Engineers have joined the debate on how, if at all, such dilemmas might be resolved through ever-improving algorithms. A 2016 study (Bonneton et al 2016) found that, while people would in principle like self-driving cars to make such decisions on utilitarian grounds, most would prefer being in a car that privileged its own driver. Such thought experiments imagine problems in terms of “quandary ethics” (JafariNaimi 2017), presuming that outcomes are calculable and that the relevant control mechanisms are within the car itself. In the real world, algorithms must work with incomplete information and imperfect control – this was the justification used by a senior Mercedes executive, who argued that its self-driving

cars would prioritise their own occupants (Taylor 2016). But governance concerns will also stretch well beyond questions of risk. As Kate Crawford and Ryan Calo (2016) observe,

The possibility of a driverless car weighing up “kill decisions” presents a narrow frame for moral reasoning. The trolley problem offers little guidance on the wider social issues at hand: the value of a massive investment in autonomous cars rather than in public transport; how safe a driverless car should be before it is allowed to navigate the world (and what tools should be used to determine this); and the potential effects of autonomous vehicles on congestion, the environment or employment (Crawford and Calo 2016).

The task, therefore, is to anticipate the futures that could surround self-driving cars: the futures that such cars might enable and the futures that those advocating such technologies might push for.

This paper is about the possible futures that self-driving car technology might help bring about. It is offered as a challenge to the narrative of autonomy and inevitability that has characterised much self-driving discourse. This plug-and-play story, in which the car is seen as able to get along with the world’s complexities as they are, without making additional demands, is a lie. Realising the purported benefits of self-driving technology will be an extensive process of modernisation. In order to accommodate the technology our roads and our lives will need to be made machine-readable. As the world around self-driving cars is upgraded in their image, some will be empowered and others disempowered. We should therefore look not just at risks, but also at the uneven distribution of benefits and the inequities that are exacerbated or created. The project is one of “anticipatory governance” (Guston 2004) or “responsible innovation” (Stilgoe and Guston 2016), seeking to replace the technological somnambulism (Winner 1986) that characterised humanity’s twentieth century relationship with the motorcar with a more alert alternative.

The conventional way of thinking about self-driving cars and the issues surrounding them takes a regular car and simply transplants a computer for a driver. Anticipatory governance means challenging such framings by considering the systems with which technologies may co-evolve. My view through the windscreen of the Tesla and the Tesla’s view through its own cameras are therefore unreliable glimpses of the future. The latter may even be a case of intentional misdirection. Rather than asking whether self-driving cars can beat humans at

their own game, we should instead look at how the game could change.

The ways in which roads and modes of transport push, pull and cajole different users are the result of (usually opaque) intentional politics or accidental oversight. Power lies in material arrangements as well as the rules of the road. The self-driving Tesla is not making any explicit political claims. The video advertises the car moving through the world rather than changing the world. However, the view from the Tesla is not just spatial. It is also an idealised view of the future.

SEEING LIKE A STATE

In 1998, James Scott published a scathing critique of twentieth century “high modern” attempts at utopian mega-engineering. Narrating, among other cases, collectivisation in the Soviet Union, homogenisation of trees in German forests, “villagization” in Sub-Saharan Africa, standardising of surnames in many countries and the imposition of Le Corbusier’s architectural purity upon Brazil’s new capital, Scott described the calamities of people in power imposing their view of what is best for their populations.

Scott explained how states seek to manage the social diversity that challenges their view. For governments to reshape the world they must make the world comprehensible or, in Scott’s (1998) terms, “legible”. The ways in which states see the world are therefore inseparable from the worlds they want to make. In some cases, these impositions succeed, constraining the possibilities of social life in the process. But in many others, as Scott describes, they fail. Human diversity and local expertise, seen through the high-modern lens as problems, can become pockets of resistance, armed with what Scott (1985) called “weapons of the weak”. Scott’s argument, if not anarchist then profoundly anti-authoritarian², offers a powerful way to think about emerging technologies, the worlds that are imagined around them and how their progress may be complicated by the real world.

With self-driving cars, governments have, at least in the US and at least for the time being, decided that they should not be substantially involved in governing the technology. They have been largely content to offer public roads as a laboratory and watch from a distance. This approach takes the suggested benefits of autonomous vehicles at face value and

delegates the imagination of futures to their innovators. When we consider the ways in which self-driving cars are already entangled in public infrastructure and will become more so in the future, this mode of governance appears irresponsible.

AUTONOMY AND ENTANGLEMENT

“It can’t find the lane markings! [...] You need to paint the bloody roads here!”. According to a Reuters reporter (Sage 2016), this outburst came from Volvo executive Lex Kerssemakers at the Los Angeles Auto Show in 2015. Kerssemakers was accompanying the city’s mayor Eric Garcetti on a test drive in a prototype self-driving Volvo XC90. When the car failed to drive the mayor as smoothly as expected, Kerssemakers deftly offloaded responsibility for technical failure onto the public sector: he blamed the roads. For Kerssemakers, the limitations of his self-driving Volvo were a function not of its naivety or myopia, but of the messiness of the outside world. The conversation, though lighthearted, was widely reported as embarrassing for California and its poorly-maintained infrastructure. Viewed differently, it offers anticipatory insight into how the ways in which self-driving cars see the world might contribute to reshaping the world.

As a way to resist the dazzling glare of new technologies, scholars in Science and Technology Studies have sought to turn critical attention towards the technological infrastructures that surround us and shape our everyday lives. Material infrastructures, digital infrastructures and the standards, rules and norms that shape their evolution are profoundly consequential, but they are designed to fade into the background (Lampland and Star 2009). The dullness of infrastructure is a sign of success (Star 1999). However, the politics of infrastructure mean that it can never be invisible for all. As Raquel Velho (2017) describes in her analysis of disabled transport users, infrastructures do not just “work”. They must be worked around or made to work in order to accommodate diversity. Wheelchair users on buses know, just as cyclists on poorly-designed roads do, that infrastructure is all too visible when its design excludes certain groups.

If representation affords power, then what can the view from the Tesla tell us about infrastructure? Tesla’s view is one of autonomy, in both technical and political senses. Its machinery is sold as an archetypal autonomous system, able to

self-govern in any circumstances and clearly outperform the human drivers that are blamed for more than 90 per cent of car crashes. Its claimed omniscience means that it makes no demands on the world. It can deal with real-world complexity, including imperfect lane markings, unpredictable cyclists and the misbehaviour of other drivers at stop signs. Everything the car needs is contained within the car and the data centres with which it communicates.

This detachment from the outside world lets Tesla argue for a libertarian approach to technology regulation: governments do not need to control the technology because the technology is in control of itself. As long as the outcomes are demonstrably better than the alternative, then the novelty of the technology's processes is unimportant. Elon Musk, the Tesla CEO who had previously joined the self-driving chorus bemoaning the illegibility of California's road markings, makes this case in the simplest possible terms: "Do the math", he says (Stilgoe 2017). If self-driving cars are safer than human drivers, then there is no cause for objection. This narrative is particularly convenient to a company such as Tesla that regards itself as a disruptor. The company is looking for approaches that cut through much of the regulatory infrastructure that currently enmeshes car companies in a web of complex responsibilities.

However, the "autonomy" of autonomous vehicles is a myth that disconnects self-driving cars from much of their own history. During the late twentieth century, it was assumed that, for self-driving cars to become possible, smart cars would need to communicate with equally smart highways (Wetmore 2003). This systemic view of the driverless dream has fallen out of fashion. The new story, enabled by rapid developments in machine learning over the last ten years and spurred on by the intervention of the US Defense Advanced Research Projects Agency (Darpa) in its "grand challenge" competitions between 2004 and 2007, is of heroic independence. The story is compelling, but self-driving cars are not self-contained, self-taught or self-sufficient (Stilgoe 2017). First, they are connected to one another, enabling what Tesla calls "Fleet Learning". When one car gathers data, all other Teslas can make use of improvements to their algorithms.

Secondly, while their ability to classify objects using visual, radar and lidar sensing is impressive, the real advantages are likely to be realised once vehicles are able to talk to other vehicles and to infrastructure (Dresner and Stone 2006; Shla-

dover 2009). The benefits of autonomous vehicles may be inversely proportional to their autonomy from one another. Engineering a conversation between cars and the world would be easier than engineering highly autonomous cars, but would involve the manufacturers of cars and software ceding control and negotiating futures with others.

Strictly speaking, there is no such thing as an autonomous system (Bradshaw et al 2013; Mindell 2015). For a thing to behave as if it is autonomous, the system needs to be constrained. With self-driving cars, this translates into a debate about the limits of a so-called “level 4” car, able to self-drive in certain circumstances³. When the National Transportation Safety Board investigated the fatal Tesla crash, they took the opportunity to remind drivers that Autopilot did not turn Teslas into true self-driving cars. However, they also criticised the company for doing “little to constrain the use of Autopilot to roadways for which it was designed” (quoted in Stilgoe 2017). The point is that any self-driving system will only work in certain circumstances and even in these circumstances there will be disagreement as to how good is good enough. Joshua Brown’s Tesla was only a “level 2” vehicle, and yet he and many others were using it as though it were a self-driving car. The confusion and controversy relates to the conditions in which a technology could be said to perform reasonably well. Waymo, a company that has done more testing than anyone, released a “safety report” (Waymo 2017) advertising its cars’ capabilities. But these were specified within an “operational design domain [...] including but not limited to roadway types, speed range, environmental conditions (weather, daytime/nighttime, etc.), and other domain constraints”. The easiest way to demonstrate that a self-driving car “works” is to narrow the conditions of its functionality. Voyage, another self-driving car company, has tested its cars within a retirement community, providing the dual benefit of a private, laboratory-like test track and photo-ops with visually- and mobility-impaired passengers.

The drawing and enforcement of technological limits by innovators and the resistance, testing and redrawing of lines by users will be a battleground for future self-driving experimentation. Car companies have discussed the possibility of “geo-fencing” their self-driving cars – preventing them from leaving a particular area – in order to delineate the small worlds of their applicability. This would allow car manufacturers to fit their cars to their likely environments. It makes

little sense to equip a car with all the requisite learning and sensors to find its way across a desert, as Darpa had initially demanded, when it is likely to spend 95 per cent of its driving time in Californian traffic. Debating the conditionality of self-driving would seem to be a climb down for companies offering all-powerful systems. There has been a subtle shift in rhetoric, from problematising the robot to problematising the outside world.

The third challenge to the autonomous ideal is that a self-driving car only makes sense when seen, like a regular car, as part of a complex socio-technical system that includes “road building, driver education programmes, gas stations, repair shops, manufacturers of spare parts and new forms of land use that spread out the population into the suburbs” (Nye 2006: 55). As car use has expanded, the system around it has had to grow too. In John Whitelegg’s words, “more money must be spent on roads, car-parking and all the associated infrastructure of dependency on motorised transport including the police and courts” (Whitelegg 1997: 18). The sociotechnical system of automobility, which enables and demands new forms of social life (Urry 1999), extends far beyond the car. “Autonomous” vehicles will be not only be no less autonomous than their conventional counterparts, they will also depend on new connections with private infrastructures of data as well as public infrastructures, norms and rules.

A further critique of the narrative of autonomy comes from Winner (1977) and others in science and technology studies who argue that, while they may give the impression of being inevitable and out-of-control, technologies are products of human work and human values. Behind the auto-didactic façade, it is notable that image classification and “machine” learning still require substantial human drudgery (Both 2014; Bradshaw 2017). The politics of technology, often disguised by innovators, need to be systematically unearthed. Winner’s (1977: 323) conclusion that “technology in a true sense *is* legislation” has been given a digital update by Lessig (1999) that should inform analysis of self-driving cars: “code is law”. For driving, today’s algorithms could become tomorrow’s “rules of the road” (Both 2016). The first step towards the governance of algorithms is therefore to pay attention to how algorithms themselves govern (Introna 2015).

Once we reject the story of autonomy, we can more clearly anticipate the politics that may come with the emergence of technology. We can ask who is likely to benefit and what new

sources of risk and injustice might arise. The view from the Tesla that we are shown is one in which the world's complexity is designed into a system that is intelligent enough to handle any eventuality. But there are pressures to design out this complexity rather than factor it into increasingly sophisticated algorithms.

DESIGNING IN / DESIGNING OUT

In the first decades of the twentieth century, when humans and cars began to mix regularly, roads had not yet been designed around cars. A film shot in San Francisco just before the 1906 earthquake shows a typical situation. The camera is fixed to the front of a cable car. It records pedestrians, carriages (horse-drawn and horseless), cable cars and streetcars moving at different speeds and angles, narrowly avoiding one another⁴. Aside from the cable cars, whose responsiveness is constrained by their rails, the road users are mutually interactive. Neither the material infrastructure nor the social infrastructure of norms, rules or standards is particularly imposing. Depending on one's view of urban transport, the scene is either chaos or prototypical "shared space" (Hamilton-Baillie 2008).

In the early years of the twentieth century, US cities sought to take advantage of the benefits of motorcars. Most social concerns about this rarefied technology related to its safety. British policymakers responded by enacting a "red flag law", demanding that each car travelled at walking pace, accompanied by flag-wavers to warn pedestrians. Many US cities determined that the use of streets should be rapidly reconfigured in favour of cars. Local authorities sought to modernise streets by shaming pedestrians away from roads. Thus, the term "jaywalking" was born and a new misdemeanour created (Norton 2007).

When it comes to the governance of emerging technologies, the choices of analogy and precedent are vital. Whether we see cars as mundane objects like bicycles or complex systems like trains frames their regulation and draws lines of responsibility (Jain 2004; JafariNaimi 2017). The privileging of cars, and their subsequent regulation as risk objects, has meant that other, softer parts of transport systems have shouldered more than their fair share of blame. As with other technologies, it becomes easy to problematise the public, whether

as ignorant critics, error-prone users or lackadaisical bystanders.

In addition to the car and its driver, a functioning transport system demands myriad systems of social control that vary widely across cultures⁵. The social infrastructure of automobility is just as important as the material infrastructure. Things that appear automatic require substantial social organisation (for example, the functioning of airline autopilots depends on the control of airspace as well as the control of the aeroplane). The forms of social control that handed streets over to cars in the US were wide-ranging, combining laws with carmakers' public relations efforts (Norton 2007). Indeed, according to the Library of Congress, the typical street view from the front of the cable car in the 1906 San Francisco movie is nothing of the sort. The cars in the film are not as numerous as they at first seem. They circle around the camera, moving in and out of shot in order to give the impression of a busy, car-heavy city. In 1906, cars were still relatively rare in San Francisco. The film is a PR stunt⁶.

Self-driving carmakers recognise that what works in Silicon Valley simulations will not work for the rest of the world. Engineers are working to improve their algorithms by designing in ever-greater granularity of human complexity and diversity. This means adding sensors, processing power and cost to the car itself. Behind the scenes, the performance of autonomy requires a private data infrastructure of machine learning and mapping that compensates for a hard-to-read world. In areas where self-driving cars are likely to emerge most profitably, companies are building detailed 3D maps and retuning existing digital maps to be read by machines rather than humans. While the Tesla's vision is presented as a form of augmented reality, it is increasingly reliant on radar. And other companies use more expensive lidar. The maps that work with these sensors are stylised and uninterpretable to human eyes. Since the uptake of the car, urban infrastructures have largely been designed to suit cars and human drivers. The risk is that the maps, roads and laws of the future will be made to empower machine rather than human navigation. Like the Mars Rover analysed by Vertesi (2012), the "view from somewhere" presented by the Tesla's cameras is part of a bigger "view from everywhere" project. The difference is that, in Tesla's case, the grander ambitions are downplayed to suit the company's strategy for a self-driving future.

Self-driving car engineers would claim that they are going to extraordinary lengths to design into their systems the full diversity of circumstances that their machines might encounter, such as variability in environment and human behaviour. However, we can see, following Scott (1998), the temptations that exist within the high modern worldview to instead design human complexity and diversity *out*. The urge is not to learn from or with social complexity, but to correct it. We can imagine that the power of this arch-modernist vision will exert substantial pressure on the public sector and civil society. If the benefits of self-driving cars are seen as unarguable, then states will find themselves under pressure to make their environments fit the needs of self-driving cars, at some cost to the public and with little consideration of matters of equity. Scott (1998) focussed on how modernist schemes aim to make social life “legible” (i.e. predictable and controllable), and how people have resisted such attempts. For self-driving cars, we can anticipate a new demand: that of making the world and its inhabitants machine-readable.

The marketing of self-driving cars as flawless is likely to reflect badly on the condition of highways. The world has, according to one estimate, 64 million miles of road, of which only 18 million miles are paved (CIA Factbook 2017). Those that could be said to be “designed” have been designed with human perception in mind. And the rest might be judged by constraints including their particular uses, the types of vehicles that travel on them and what is affordable. If we estimate that upgrading roads in rich countries costs around a million dollars per mile, we can begin to see the price for realising the self-driving proponents’ vision of a world without 1.2 million car deaths every year. For “autonomous vehicles” to work as promised, parts of the world will need to be upgraded, at great expense, to match the expectations of a car’s sensors.

At the moment, only a few engineers (e.g. Ng and Lin 2016) would admit to such a requirement. DiClemente and colleagues (2014) conclude that “the conversion to a fully automated road infrastructure will be one of the most momentous challenges that humanity will face in the Twenty-First century”. The imagined upgrades would not just be to our roads, but also to our attitudes, including our willingness to place our trust in new technology. People inside and outside cars would need to adapt to accommodate the technology, but this challenges the autonomous ideal.

The Tesla view sees other road users as passive design constraints. People become just another part of the backdrop, to be interpreted just like a stop sign. A pedestrian is just an object to be avoided. Their motivations are only important inasmuch as they are likely to determine what happens next. Early social research suggests that, in their encounters with self-driving cars, other road users are, unsurprisingly, likely to behave as active, knowing agents. They may take advantage of the over-cautiousness of a car's algorithms or play games with its sensors (Rothenbücher et al 2016). Some pedestrian and cyclist campaigners have spotted an opportunity to use the deference of AVs to reassert their rights (Connor 2016). And engineers (e.g. Evtimov et al 2017) have begun to speculate on the possibility of self-driving car systems being hacked, perhaps by disgruntled former truckers with "adversarial machine learning" (Garfinkel 2017). Even though the precise form of such algorithmic resistance might be impossible to predict, Scott's analysis would point to some form of inevitable opposition.

In a more mundane way, self-driving cars would also ask new things of their users, while allowing existing driving skills to atrophy. As has already become clear from the Tesla crash and its aftermath, self-driving cars need to be used with care, just as aeroplane autopilots do (Stilgoe 2017). Drivers cannot just sit back and let the car take over⁷. And even if innovation does eventually make humans no more than passive passengers, some will resist. Commentators are already starting to imagine the ways in which the norms and rules surrounding transport will be put under pressure by automation. One writer, tracing a scenario in which the safety benefits of automotive automation persuade policymakers to confiscate conventional cars, suggests the need for an equivalent of the National Rifle Association to assert drivers' rights (Roy 2016).

The self-driving car will not be able to deliver its utopia alone. In addition to making the world machine-readable, the scale of wider changes is likely to be substantial, and modernisation will impose burdens in terms of public investment and social changes – new laws and norms that will constrain some social choices while opening up others. The view from the Tesla, which seems modest at first, has more far-reaching consequences. It is a view of the world and a view of the future direction of innovation. It is a technology-first view, starting with technical possibilities and postponing consideration of

wider consequences. Given the potential for world-changing implications, what might the alternative views be?

SEEING LIKE A CITY

The view taken by Brasilia's high modernist architects when they planned the new city was determinedly top-down (Scott 1998). The utopia they envisaged failed to materialise. The purposes that they ascribed to places did not fit people's lives. Their officially designated public spaces ended up as dead, empty spaces (Holston 1989). The plan was packed with imagination, but it was drawn up without any attention to people's everyday lives: all "Vision"; no vision.

Critics of urban modernism have pointed out that the purposes of cities and their inhabitants are multiple and impossible to demarcate. For Jane Jacobs (1961, in Scott 1998), the city is emergent: "Intricate minglings of different uses are not a form of chaos. On the contrary, they represent a complex and highly developed form of order" (see also Hinchcliffe et al 2005). This is not to argue that planners should just leave cities alone, but to recognise that they are mixed public, private, economic, social and technological systems and plan accordingly. Cities are places of innovation, rather than places to which innovation happens.

A city is not one thing, so seeing like a city means taking multiple perspectives. In terms of transport, it also means putting technology in its place. In this paper, I have explored some of the implicit constitutional arrangements sitting underneath motivations for self-driving car development. In most cases, these would not be admitted by the self-driving innovators themselves. They would frame their work in terms of simple problems such as road safety, and claim that their philosophy of innovation was mere disruption. However, scratch a disruptor and you often find a utopian. The self-driving car is, on closer inspection, being sold as a vehicle for social improvement (Bilger 2013), offering benefits not just for safety but also for accessibility, equality, congestion, urban design and the future of work. The futures on offer, from innovators themselves or from enthusiastic policy consultants, are radical, expressing with admirable certainty that, for example, there will be 80 per cent fewer cars on the road (Claudel and Ratti 2015) or 585,000 lives saved and \$7 trillion gained (Strategy Analytics 2017).

In this paper I have attempted to sketch some ways in which we might anticipate the ramifications of such visions for human diversity, equity and justice. The view from the Tesla is one of technological autonomy. But this view is misleading. In its emergence, the technology is far from autonomous. It is already imbued with ideals about the world in which it should drive. And it will, despite its claims, make demands on the world around it. It will bring new infrastructures of its own and demand improvements to roads, laws and public behaviours. It will, while diminishing our existing capacities as drivers, demand new skills and new responsibilities. In doing so, it will risk exacerbating inequalities and running over human diversity. The privatisation of machine learning in cars (Stilgoe 2017) could, if left uninterrogated, lead to a de facto privatisation of public transport.

All of this is not to say that such concerns have been ignored in the debate about self-driving. Governments are reactively expressing concern about the impact that self-driving cars may have on cities, public transport systems and drivers' employment. But these things are conventionally characterised as "second order consequences" of technology. The technology itself remains black-boxed. The task is now to connect such concerns with the constitution of technology itself. The investment and volume of debate swirling around competing self-driving players is such that identifying alternatives is not particularly hard. The things imagined as "second order impacts" are in many cases enabling conditions for particular technological visions. As with conventional cars, different assumptions about technology and transport across cultures and jurisdictions will lead to very different governance arrangements for self-driving cars. The hope is that such arrangements are at least partly intentional rather than somnambulant. Encouraging policymakers to see like a city rather than a Tesla may be the first step towards responsible innovation in self-driving cars.

THE EXPERIMENTAL STATE

In an article written near the end of his term in office, President Obama was cautiously optimistic about self-driving cars, claiming that they would mean, "Safer, more accessible driving. Less congested, less polluted roads. That's what harnessing technology for good can look like. But we have to get

it right” (Obama 2016). There is little imagination of what else it would take, beyond the self-driving car itself, to “get it right”.

Successful technologies do not just enter the world. They promise to change it. And the world speaks back. Driving is similarly conversational. When driving, we treat other road users as active participants rather than mere background. The project of developing a workable self-driving car, able to navigate a range of environments and contexts, is therefore far harder than just mimicking human perception. The design challenge stretches well beyond the car, to include the worlds in which self-driving cars will operate (Blyth et al 2016). Taking this challenge seriously means radically rethinking current modes of experimentation and testing.

Learning for self-driving cars is currently privatised. Car companies are doing the innovating and enthusiastic local governments are expected to respond, perhaps through inviting testing on their roads. But what if the conversation between innovators and governments were genuinely collaborative? What if, instead of government merely acting as leaseholders on the laboratory, they helped define the experiment?

In most developed countries, national governments are in thrall to the advertised possibilities offered by Tesla, Waymo and other developers of self-driving cars. The dominant governance concern is with capturing the economic and social opportunities on offer. However, more imaginative local governments recognise that the view from the Tesla is partial. They know that a trajectory of innovation that starts in Silicon Valley cannot just be transported into their transport systems. In places with more established traditions of public transport such as London and Gothenburg, the pace of innovation seems less frenetic, but cities are pushing back to articulate their vision and experiment on their own terms.

NOTES

¹ The car that I drove/drove me contained first generation Autopilot hardware. When the subsequent generation of hardware, described as “full self-driving hardware” was first introduced, Autopilot’s performance was diminished as the car had to relearn how to drive with its new set of sensors.

² Scott (2012) would later call this view “seeing like an anarchist”.

³ The Society of Automotive Engineers’ widely-adopted typology of autonomy levels delineates responsibility between human and machine. Level one is a car with have one automated driving function, such as adaptive cruise control. At Level two cars, the machine controls the steering and acceleration but the human is responsible.

In level three cars, the human is able to take their eyes off the road while the car is in control, but expected to take back control when the situation demands. At level four, the car is fully-self-driving in certain locations.

⁴ The film, produced by the Miles brothers and titled *A Trip Down Market Street*, is widely available on YouTube.

⁵ Melissa Cefkin, personal communication.

⁶ Film: *A trip down Market Street before the fire*, Library of Congress, <https://www.loc.gov/item/00694408>, accessed 15 Aug 2017.

⁷ Within self-driving car innovation, there is substantial disagreement on the wisdom of so-called “level 3 autonomy”, in which cars and humans swap driving responsibilities.

REFERENCES

- W.E. Bijker, T.P. Hughes, T.J. Pinch (eds.) (1987), *The social construction of technological systems: New directions in the sociology and history of technology* (Cambridge: MIT Press).
- B. Bilger (2013), *Auto correct*, in “The New Yorker”, 25 November, available at: www.newyorker.com/magazine/2013/11/25/auto-correct (accessed 12 March 2017).
- P.L. Blyth, M.N. Mladenovic, B.A. Nardi, H.R. Ekbia, N.M. Su (2016), *Expanding the design horizon for self-driving vehicles: Distributing benefits and burdens*, in “IEEE Technology and Society Magazine”, 35,3, pp. 44-49.
- J.F. Bonnefon, A. Shariff, I. Rahwan (2016), *The social dilemma of autonomous vehicles*, in “Science”, 352(6293), pp. 1573-1576.
- G. Both (2014), *What drives research in self-driving cars? (part 2: surprisingly not machine learning)*, in, “The CASTAC Blog”, available at <http://blog.castac.org/2014/04/what-drives-research-in-self-driving-cars-part-2-surprisingly-not-machine-learning/> (accessed 11 March 2017).
- J.M. Bradshaw, R.R. Hoffman, D.D. Woods, M. Johnson (2013), *The Seven Deadly Myths of “Autonomous Systems”*, in “IEEE Intelligent Systems”, 28, 3, pp. 54-61.
- T. Bradshaw (2017), *Self-driving cars prove to be labour-intensive for humans*, in “Financial Times” 9 July 2017.
- R.A. Brooks, M.J. Mataric (1993), *Real robots, real learning problems*, in “Robot learning”, pp. 193-213.
- CIA Factbook (2017) Field Listing: Roadways, <https://www.cia.gov/library/publications/the-world-factbook/fields/2085.html>, accessed 23 August 2017.
- M. Claudel, C. Ratti (2015), *Full speed ahead: How the driverless car could transform cities*, in “McKinsey and company”, <http://www.mckinsey.com/business-functions/sustainability-and-resource-productivity/our-insights/full-speed-ahead-how-the-driverless-car-could-transform-cities>, accessed 23 August 2017.
- H.M. Collins (1988), *Public experiments and displays of virtuosity: The core-set revisited*, in “Social studies of science”, 18, 4, pp. 725-748.
- S. Connor (2016), *First self-driving cars will be unmarked so that other drivers don't try to bully them*, in “The Guardian”, 29 October, available at https://www.theguardian.com/technology/2016/oct/30/volvo-self-driving-car-autonomous?CMP=share_btn_tw, accessed 12 March 2017.
- K. Crawford, R. Calo (2016), *There is a blind spot in AI research*, in “Nature”, 538, pp. 311-313.
- J. DiClemente, S. Mogos, R. Wang (2014), *Autonomous Car Policy Report* (Pittsburgh: Carnegie Mellon University).
- K. Dresner, P. Stone (2006), *Traffic intersections of the future*, in “Proceedings of the national conference on artificial intelligence”, 21, 2, pp. 1593-1596) Menlo Park, CA; Cambridge, MA; London; AAI Press; MIT Press; 1999.
- I. Evtimov, K. Eykholt, E. Fernandes, T. Kohno, B. Li, A. Prakash, D. Song (2017), *Robust Physical-World Attacks on Machine Learning Models*, in “arXiv preprint” arXiv:1707.08945.



- S. Garfinkel (2017), Hackers Are the Real Obstacle for Self-Driving Vehicles, in "MIT Technology Review", 22 August 2017, <https://www.technologyreview.com/s/608618/hackers-are-the-real-obstacle-for-self-driving-vehicles/>, accessed 23 August 2017.
- D.H. Guston (2014). *Understanding "anticipatory governance"*, in "Social Studies of Science", 44, 2, pp. 218-242.
- D.H. Guston, D. Sarewitz (2002), *Real-time technology assessment*, in "Technology in society", 24, 1, pp. 93-109.
- B. Hamilton-Baillie (2008), *Shared space: Reconciling people, places and traffic*, in "Built environment", 34, 2, pp. 161-181.
- S. Hinchliffe, M.B. Kearnes, M. Degen, S. Whatmore (2005), *Urban wild things: a cosmopolitical experiment*, in "Environment and planning D: Society and Space", 23, 5, pp. 643-658.
- J. Holston (1989), *The modernist city: An anthropological critique of Brasília* (Chicago: University of Chicago Press).
- L.D. Introna (2016), *Algorithms, governance, and governmentality: On governing academic writing*, in "Science, Technology, & Human Values", 41, 1, pp. 17-49.
- J. Jacobs (1961), *The Death and Life of Great American Cities* (New York: Random House).
- N. JafariNaimi (2017), *Our Bodies in the Trolley's Path, or Why Self-Driving Cars Must Not Be Programmed to Kill*, in "Science, Technology, & Human Values", 0162243917718942.
- S.S.L. Jain (2004), "Dangerous instrumentality": the bystander as subject in automobility, in "Cultural Anthropology", 19, 1, pp. 61-94.
- M. Lampland, S.L. Star (eds.) (2009), *Standards and their stories: How quantifying, classifying, and formalizing practices shape everyday life* (Ithaca: Cornell University Press).
- B. Latour (1992), *Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts*, in W.E. Bijker, J. Law (eds), *Shaping technology/building society: Studies in sociotechnical change* (Cambridge: MIT press).
- L. Lessig (1999), *Code: And other laws of cyberspace* (New York: Basic Books).
- D.A. Mindell (2015), *Our robots, ourselves: Robotics and the myths of autonomy* (New York: Viking Adult).
- M.M. Moore, B. Lu (2011), *Autonomous vehicles for personal transport: A technology assessment*, available at https://papers.ssrn.com/sol3/papers2.cfm?abstract_id=1865047.
- A. Ng, Y. Lin (2016), *Self-driving cars won't work until we change our roads – and attitudes*, in "Wired", 15 March, available at <https://www.wired.com/2016/03/self-driving-cars-wont-work-change-roads-attitudes/> (12 March 2017).
- D.E. Nye (2006), *Technology matters: Questions to live with* (Cambridge: MIT Press).
- A. Roy (2016), *We Need An NRA-Type Lobby for Human Driving*, in "The Drive", 14 November 2016, <http://www.thedrive.com/opinion/5979/we-need-an-nra-type-lobby-for-human-driving>, accessed 23 August 2017.
- A. Sage (2016), *Where's the lane? Self-driving cars confused by shabby U.S. roadways*, in "Reuters", 31 March 2016, <https://www.reuters.com/article/us-autos-autonomous-infrastructure-insig/wheres-the-lane-self-driving-cars-confused-by-shabby-u-s-roadways-idUSKCN0WX131>, accessed 4 December 2017.
- J.C. Scott (1985), *Weapons of the weak: Everyday forms of resistance* (New Haven: Yale University Press).
- J.C. Scott (1998), *Seeing like a state: How certain schemes to improve the human condition have failed* (New Haven: Yale University Press).
- J.C. Scott (2012), *Two cheers for anarchism: Six easy pieces on autonomy, dignity, and meaningful work and play* (Princeton: Princeton University Press).
- S.E. Shladover (2009), *Cooperative (rather than autonomous) vehicle-highway automation systems*, in "IEEE Intelligent Transportation Systems Magazine", 1, 1, pp. 10-19.
- S.L. Star (1999), *The ethnography of infrastructure*, in "American behavioral scientist", 43, 3, pp. 377-391.
- J. Stilgoe (2017), *Machine learning, social learning and the governance of self-driving cars*, in "Social Studies of Science", <https://doi.org/10.1177/0306312717741687>.



- J. Stilgoe, D. Guston (2016), *Responsible research and innovation*, in U. Felt, R. Fouché, C.A. Miller, L. Smith-Doerr (eds.) (2016), *The Handbook of Science and Technology Studies* (Cambridge: MIT Press).
- Strategy Analytics (2017), *Accelerating the Future: The Economic Impact of the Emerging*, report for Intel, June 2017, <https://newsroom.intel.com/newsroom/wp-content/uploads/sites/11/2017/05/passenger-economy.pdf>, accessed 23 August 2017.
- M. Taylor (2016), *Self-Driving Mercedes-Benzes Will Prioritize Occupant Safety over Pedestrians*, 7 October 2016, <http://blog.caranddriver.com/self-driving-mercedes-will-prioritize-occupant-safety-over-pedestrians/>, accessed 1 December 2017.
- J. Urry (1999), *Automobility, car culture and weightless travel: a discussion paper* (Lancaster: Department of Sociology, Lancaster University).
- R. Velho (2017), *Fixing the Gap: an investigation into wheelchair users' shaping of London public transport*, PhD thesis, University College London.
- J. Vertesi (2012), *Seeing like a Rover: Visualization, embodiment, and interaction on the Mars Exploration Rover Mission*, in "Social Studies of Science", 42, 3, pp. 393-414.
- J. Wetmore (2003), *Driving the Dream, The History and Motivations Behind 60 Years of Automated Highway Systems in America*, in "Automotive History Review", Summer 2003, pp. 4-19.
- Waymo (2017), *Safety Report: On the Road to Fully Self-Driving*, <https://storage.googleapis.com/sdc-prod/v1/safety-report/waymo-safety-report-2017.pdf>, accessed 4 December 2017.
- J. Whitelegg (1997), *Critical Mass: Transport, Environment and Society in the Twenty-first Century* (London: Pluto Press).
- J. Wilsdon, R. Willis (2004), *See-through science: Why public engagement needs to move upstream* (Demos).
- L. Winner (1977), *Autonomous technology: Technics-out-of-control as a theme in political thought* (Cambridge: MIT Press).
- L. Winner (1986), *The Whale and the Reactor: A Search for Limits in an Age of High Technology* (Chicago: University of Chicago press).