Trials and tribulations:
Who learns what from urban experiments with self-driving vehicles?

Jack Stilgoe
Cian O'Donovan
Department of Science and Technology Studies,
University College London[1]

Cite as: Stilgoe, J. and O'Donovan, C. (2023) *Trials and tribulations: Who learns what from urban experiments with self-driving vehicles*? in Artificial Intelligence and the City Urbanistic Perspectives on AI. Edited By Cugurullo, F. Caprotti, F. Cook, M. Karvonen, A. McGuirk, P. and Marvin, S. Routledge. https://doi.org/10.4324/9781003365877-4

Abstract

Self-driving (or 'autonomous') vehicles (AVs) represent a profound test case for artificial intelligence in open, unpredictable contexts. Many cities have played host to trials of self-driving car technology, offering up their roads as a form of testbed. But what is on trial and what is really being learnt? This chapter reports on our experiences with AV trials. These trials, on their own, are a poor indication of possible urban futures. They are typically designed to give the impression that the technology is unconstrained, which means they are intentionally disconnected from questions of governance. Our conclusion is that, for trials to be made relevant for cities, planners and policymakers must play a more active role in shaping and learning from them.

> *"I sometimes felt that the government's enthusiasm for making this country a testbed for driverless technology would give us all the benefits that Christmas Island got from being a testbed for nuclear weapons. You get all the pain, but you don't necessarily get much of the gain."*
>
> – Former UK Department for Transport civil servant[2]

Introduction: Olympian dreams and concrete realities

An orange and white 'pod', described by one engineer as resembling a giant toaster, is making slow circuits of London's Olympic Park on a sunny September morning. The shell of the vehicle is the same as those that ferry passengers from a car park to a terminal in a dedicated concrete lane at Heathrow airport. Here, the 2019 version has upgraded software, radar, lidar, GPS and other sensors to aid its omniscience and now it is driving itself without rails. But the pods are still being defended, by a team of hi-vis-wearing marshalls. Although there's no driver sitting with the four passengers in the POD, an engineer is stewarding proceedings on her tablet-based control interface, walking among the marshalls behind. Inside, there's an on-board safety steward. The presence of the steward is contributing to the

[1] Jack Stilgoe ORCID 0000-0002-5787-2198. Cian O'Donovan ORCID 0000-0003-4467-9687

[2] From an interview conducted as part of a podcast for the Driverless Futures? project

underwhelming experience of the public passengers who have been offered a ride by marshalls along the way. The passengers seem nonplussed. At only four miles per hour, with frequent juddering halts interrupting progress, they'd likely have been quicker walking.

A new automation control system means the pod can leave the physical rails of Heathrow behind. But the vehicle is still far from roadworthy. That's why it's in the car-free Olympic Park, where transport bylaws and legal exceptions offer a simpler regulatory route to delivery compared with open highways. It's also the reason why the marshalls have been given a strict brief. Their job is to notify pedestrians and cyclists that the pod is approaching and, if necessary, encourage them away from the vehicle. If an "obstacle" cannot be removed, they must notify the safety steward. In the event of an emergency, the marshalls are required to assist the first responder (the safety steward riding in the pod) and a designated incident manager. The brief, like the formation the marshalls take up on the pathways, is defensive. Self-driving vehicles represent a high-profile test case for the use of artificial intelligence in the wild, with software in control of high-stakes and high-momentum hardware through unpredictable environments. But if this 'trial' is an experiment, it's a strange one, because it's designed at all costs not to fail. Given this is supposed to be a 'driverless', 'autonomous' vehicle, there are a surprising number of people involved.

The trial is part of a £4 million research and development consortium of public and private transport players. Private-sector engineers, including some of the team behind the original airport vehicles, are producing data and technical knowledge. Legal and insurance experts are building case studies for the industry. University transport studies researchers are interviewing participants on their experiences and attitudes. And public sector authorities are building relationships with all of the above. On one day during the trial, a team of cybersecurity researchers arrived and attempted to 'break-in' to the pod's systems. On another, a toy rat was attached to a remote control car in an attempt to play havoc with the pod's sensors. The sight of a robotic car in a public place is undeniably interesting to many different groups, but it's not clear what's at stake or for whom.

Interviews with some of the partners involved in the project suggested that learning was taking place, but it was happening behind the scenes. Asked about what the tests were really testing, one consultant involved in managing the project responded,

> "So, yeah, what am I allowed to say about that? I think it would be different if we were stood at a bar (laughs). So, of course, I know what [the project] is supposed to be doing. The vehicles are supposed to be able to operate amongst people in a pedestrianised area, and I think they also need to be able to operate on roads in some way… a university with six months and a few thousand pounds can make a vehicle drive itself… The real research is around how they work: are they up for 24/7 operation? Do they work on these kinds of narrow roads where they're very busy? Do they work on these kinds of roundabouts, and, given all that stuff, do people like them, will they get on them, and so on and so forth?" (Interviewee A)

As the project developed, it became clear that even these modest ambitions for "real research" would need moderating. A collaborator on the same project explained that the need to deliver a public demonstration had overtaken the initial research aims. "We said that we'd do it… so we need to do it. What are we really testing?... It's quite difficult to test the actual use case. Ideally, we want to be able to test the technology" but, he went on, "The challenge for us is that, given the technology constraints, we couldn't really deliver a trial which matched an obvious use case. So it's a fairly contrived route for the pods.” (Interviewee B). The question of whether large numbers of people would actually use the technology was complicated by the technology’s inability to cope with large numbers of people.

According to this interviewee, “there's been little interest from [the funders] in the outcomes of our project… I'm astonished, if I'm quite honest - the amount of money that's being put in - at how little interest there is from [the funders] to pull together the findings and the learnings across their investments” (Interviewee B). His colleague suspected that funders and policymakers were more interested in keeping up appearances: “I think they’re all trying to say on a worldwide stage, ‘Look. Look what we’ve done! How fantastic it is!’ So, I think a project that learns a lot and doesn’t show something with lots of video, and people at it, and, you know, press coverage, and such, is not all that interesting… they’re looking for that message to the world, ‘Hey, we are the best place to come and try and test these sorts of things’.” (Interviewee A)

In this chapter, we consider what cities might learn from public tests of self-driving vehicle technologies. We draw on research conducted as part of two projects *Driverless Futures?* and SCALINGS. These projects involved more than 50 interviews with innovators and stakeholders in the US, UK and elsewhere in Europe, public surveys in the UK and US, British public dialogue exercises and a series of technographic observations in the US and UK with organisations developing and testing the technology. Our engagement with self-driving vehicles starts from our position as researchers in Science and Technology Studies, so while we share the interest in emerging ‘smart’ cities , urban artificial intelligence (Cugurullo, 2020) and urban robotics (While, Marvin and Kovacic, 2021) that originates in geography and planning, our analysis has followed technological promises into urban contexts rather than starting in the city. As we argue below, a major issue with innovators’ claims about self-driving vehicles is that they are intentionally dislocated – disconnected from particular places. In seeking to understand how technological promises might come down to earth, the city has become a vital site for analysis.

As STS scholars, we are particularly interested in experimentation and the claims that innovators make about experiments. In this chapter, by focussing on experimentation, we investigate how public trials of the technology are framed by technology companies and the policymakers who support them. We see that ‘trials’ are more about public persuasion than technological testing. A constructive response to such urban testbeds might therefore be to ask for more experimentation, not less. Our aim in this chapter is not to reject experiments as useless, or as unethical, although there are reasons to be concerned about both the productivity and morality of such things. Instead, we argue that a democratic engagement with experimentation as a way of opening technological black boxes could offer new possibilities for governance.

Innovation and urban experimentation

The characterisation of technologies and technological progress in terms of experimentation has a long history among evangelists and analysts as well as critics of particular technological schemes. The language of trials, experimentation and laboratories invites both positive associations with future-making and concerns about uncertainties, unintended consequences and consent if citizens are unwilling or unwitting experimental subjects. The definitional politics of experimentation demand closer analysis, and we can build on two key insights from science and technology studies. The first point is that much innovative activity we might consider experimental is not publicly acknowledged as such. It is often in innovators' interests to pretend that a technology's effects are well-known. STS scholars (e.g. (Krohn and Weingart, 1987)) have highlighted the uncertainties and hidden experimentality of nuclear energy, and they have analysed public responses to corporate experiments with social media where ethical questions were not asked in advance (boyd, 2016). For innovators, to be explicitly experimental is to admit both uncertainty and ethical ambiguity.

The second insight from STS is that much activity that might be badged an 'experiment' is not very experimental, in the sense of it being closed to the possibility of surprise (Gross, 2021). Public 'tests' of a technology are often so controlled, and the possibility of failure so constrained, that they resemble performances or demonstrations rather than trials (Collins, 1988). Where new modes of governance are tried, including experiments in deliberative democracy (Laurent, 2011), policymakers often delimit the possibility of surprise or impact as a way to maintain control. The deployment, at least discursively, of the language of experiments and laboratories as a mode of urban development (Karvonen and van Heur, 2014) demands critical analysis in these terms. 'Experiments', rather than being a way of opening up urban governance, may just be a new discourse for control. AV trials do however, present an opportunity for democratisation, because they are at least partly public. There is an interesting reversal of the trend observed by Paul Leonardi (Leonardi, 2010), in which automotive testing became increasingly detached from the real world as computer simulation improved. Now, simulated possibilities of self-driving need to be exposed to the real world in order to test whether they are practically workable. Engels et al. (Engels, Wentland and Pfotenhauer, 2019) analyse the new enthusiasm for 'test beds', which

> "re-interpret what is meant by "laboratory" in that they do not test technologies in a separate space prior to use within society. "Living" labs rather test new sociotechnical arrangements by tentatively adopting the very technologies in question "as if" the involved technologies had been found safe and had entered the market already."

Or, to use Linnet Taylor's phrase, "the experimentation taking place does not aim to test technology using people, but to test people using technology" (Taylor, 2021). Employed in this way, the idea of the testbed leapfrogs considerations that, as social scientists, we would regard as vital to do with the assessment, uptake and scaling of technology, in order to assert, via a proof-of-concept, the inevitability of an innovation. In this way, testbeds and 'living labs' become not just experiments but ways of performing innovation (Laurent and Tironi, 2015) to a public that is imagined as sceptical. This mode of persuasion becomes particularly important in the case of artificial intelligence, where innovators are likely to shift attention to systems' performance - what AI is able to do - rather than the processes through which it systems are created or operate – how and why AI works – because of a combination of technical opacity and corporate secrecy (Burrell, 2016).

Cugurullo (2018) uses the allegory of Frankenstein to critique current approaches to 'smart city' urban development that look to privatise experimentation and innovate for innovation's sake. We should remember Langdon Winner's (Winner, 1977) question: "what, after all, is Frankenstein's problem?" It is not, says Winner, that the experiment is performed, but that the doctor takes no care for the creature nor the wider meaning of his experiment. Perhaps, rather than demanding less experimentation, we should instead be asking experimenters to "love your monsters" Latour 2004), to take greater care of their creations in the context of their environments (see also (Cugurullo, 2021)). AVs could be an opportunity for a more careful model of experimentation in which, rather than innovation happening *to* citizens, it happens *for* them and possibly *with* them. In the UK, at least, the approach has been to foreground experimentation as a new approach to technology policy.

Laboratory as policy

British policy on AVs reflects a desire to capture both the imagined social benefits of the technology – safety, accessibility, efficiency and more besides – and as a slice of an imagined economic dividend. The concern over the nation's performance in a what policymakers see as a global 'race' to develop the technology is palpable. But the UK's policy approach reflects a neurosis that the country is no longer a manufacturing superpower, nor does it play host to the sort of giant technology companies that have come to dominate AI infrastructures. Its bet is therefore that it can be a laboratory for imported or homegrown novelty. In 2015, the Council for Science and Technology, a senior advisory body, recommended the establishment of a 'real-world lab' within the UK.[3] At the same time, the Government published a code of practice for testing AVs that attempted to free up would-be innovators to conduct tests.[4] A 2017 Act of Parliament sought to "ensure the next wave of self-driving technology is invented, designed and operated safely in the UK". In the years since, a series of projects supported by a new Centre for Connected and Autonomous Vehicles (CCAV) within the Government have put prototype technologies in public places and a handful of start-up companies have been given permission to test self-driving systems on public roads. (See (Hopkins and Schwanen, 2018) for further policy analysis).

The impact of this policy shift towards real-world demonstrations can be seen on the ground, inside British robotics labs, where engineers and others come together to test machines, infrastructures, safety systems and public acceptability. Analysis of CCAV-funded projects at one of the UK's largest public robotics labs shows a diversity of expertise being brought to bear. A logic of experimentation informed the lab's research and design of driverless vehicles (Michalec, O'Donovan and Sobhani, 2021), but capabilities to steer the research were kept firmly in hands of the automotive industry. Researchers on the ground were empowered to accelerate innovation, but could not steer, raising the question of where the direction of UK's innovation for AVs is being charted.

The UK's AV trial projects have in most cases included some narrowly-framed contributions from social science, asking questions to do with user acceptance (Stilgoe and Cohen, 2021). In addition, some have been studied by social scientists who are able to ask, at one stage removed, what the purposes and politics of such trials are (e.g. (McDowell-Naylor, 2018), (Marres, 2020) (Wu, 2022). The conclusion from scholars such as (Marres, 2020) is that because of the cautious, sanitised performance of such trials, nothing much is being learnt. As experiments have transgressed the lab's boundaries to be closer to the public, in living labs,

[3] Council for Science and Technology, Letter to the Prime Minister on how the UK can get the greatest value from the autonomous and connected vehicles industry. 23 July 2015 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/459521/cst-15-1-driverless-vehicles.pdf

[4] Department for Transport, 2015, The pathway to driverless cars: A code of practice for testing

test-beds and on public roads, the technological contingencies are often hidden (Engels et al., 2019; Marres, 2020; (Paddeu, Shergold and Parkhurst, 2020).

The ifs and buts of autonomous vehicles

The trial of the pods in the Olympic Park and others like it are symptoms of a pattern that characterises AV innovation and policy. The narrative of a self-driving future extends the idea of 'autonomy' far beyond its narrow engineering sense (Tennant and Stilgoe, 2021). The story told by innovators is of direct substitutability between human and machine drivers, giving the impression of always-and-everywhere autonomy, able to plug-and-play in urban environments. This narrative serves multiple purposes: it is 'disruptive' while also being comfortable; it suggests the possibility of rapid innovation; and, crucially, it also maintains that the future is in tech companies' hands rather than depending upon others, such as regulators or infrastructure providers. Urban trials typically seek to enact the narrative, by emphasising the 'autonomy' of vehicles rather than the connectivity on which they actually depend. For urban policymakers, however, such a narrative is unhelpful because it seeks to strip the technology of context. If cities are to make good decisions about self-driving technologies, they need to understand the technologies' limits, the conditions that allow for their safe and smooth operation, their infrastructural needs and their likely relationships with other actors (Stilgoe, 2017).

Close observation of the development of AVs reveals a collection of ifs and buts that define the technology (see Tennant and Stilgoe 2021 for more detail). As well as the people that protect AVs in trial situations (including safety drivers and remote operators able to take over during technological wobbles), most vehicles depend on some form of digital high-definition map. While most rely on cameras and lidar sensors to observe the world, developers admit that smart infrastructure communicating directly with the world would improve the operation of their systems. And, though they may not admit it, the successful operation of any road user, robotic or otherwise, depends on the cooperation of others. AVs have trouble with unpredictable pedestrians, but their systems are built to see pedestrians just as they see other objects they might encounter, as passive threats rather than active agents with whom they have mutual relationships. These ifs and buts – connections, conditions and relationships – that define 'autonomous' vehicles are rarely foregrounded in trial situations, even though they are precisely the things that would be useful for people wishing to assess the potential for the technology to make a positive difference within (parts of) a city.

The UK has seen a strategic attempt to deliver on the early ambition to build a "real-world lab". An organisation called Zenzic has smartened up a set of existing infrastructures and rebranded it as "Testbed UK". In London, the Smart Mobility Living Lab (SMLL) forms part of the network. Described in one presentation as offering an AV test everything "from the nursery slopes to the black run", the SMLL argument has been that London should be a laboratory not because it is well-suited to AVs but precisely because it is a hard case. An appreciation for geographical contingency hints at a recognition that AV systems are likely to vary between locations, but the more commonly deployed argument is a 'Sinatra strategy' ("If I can make it there, I'll make it anywhere"), a logic that has also been used by companies testing in Arctic Norway (Ryghaug *et al.*, 2022). This narrative, in which any issue can be resolved through the gathering of more data, can be seen as an attempt to escape from the attachments that actually define the technology (Tennant and Stilgoe, 2021). So if we take a definition of urban experiments as "political acts of defining and creating niches, by influencing both the social norms and physical spaces in which new social practices (can) emerge and thrive" (Savini and Bertolini, 2019), p. 832), 'autonomous' vehicles are an odd fit.

For all the talk, however, the SMLL is making life easy for AVs. It is serving up 15 miles of monitored urban roads, "instrumenting" the environment with connected infrastructure, offering a high-definition digital map of the whole testbed, controlling public access to certain areas and blanketing the area with a 5G network to enable constant communication (Naqvi 2021). It claims to be a 'living lab', putting innovation to the test, but, as Naqvi (2021) identifies, it is not clear how anything would fail its tests. As with other technologies of AI, innovators are, most of the time, able to reconfigure 'failure' as just one more argument for more data.

The 'lab' is clearly more an instrument of industrial strategy than of technology assessment. A 2020 innovation report from CCAV uses variations of the phrase *world-leading* 33 times in framing UK self-driving vehicle innovation over only 12 pages.[5] Zenzic's *UK Connected and Automated Mobility Roadmap*, produced in 2019, takes the same line. "We are a world leader" is the concluding clause in its vision statement for 2030. It seems that the desire to perform and win on a world stage, to an audience never actually specified, is a central goal, rather than a productive spillover, of British AV policy. The UK policy of seeking to encourage and organise experimentation becomes clearer when compared with some of the approaches taken in other places.

Detached and dislocated experimentation

In the US, Federal and local governments have offered less direct funding, but have been even more enthusiastic, opening up their streets to AV testing with few constraints. As of 2022, Waymo, a company that began life as Google's self-driving car project, reports that its vehicles have driven 20 million 'autonomous driving miles' on public roads and 20 billion miles in simulation. In Phoenix, Arizona, thousands of customers have been driven by 'Waymo One', a taxi service with nobody in the driver's seat. What began as a trial has become a de facto deployment of an AV system (in a limited area and only in favourable weather conditions). In 2022, Waymo began operating their cars in San Francisco, where another company, Cruise, has also started running self-driving cars with no backup driver, albeit only at night, when the roads are quieter. The electric car company Tesla is running a different sort of experiment. After charging some customers thousands of dollars to purchase 'full self-driving' as an optional extra, with the vague promise that the company would do its best, via software updates, to deliver such a feature, Tesla has allowed a subset to download a version – 'FSD Beta' – that enables users to help the company debug its software.

The results of these experiments are unclear. Waymo are growing in confidence and allowing users to record and publicly comment on their rides, but most of the data that would be of interest to cities is kept secret. One bureaucratic attempt at collaboration, designed by the California Department of Motor Vehicles, demands that AV companies report 'disengagements', moments of switching between automated and manual mode. Though criticised by some because of a lack of clarity about what counts as a disengagement, the

---

[5] Centre for Connected and Autonomous Vehicles, 26 October 2020, https://www.gov.uk/government/publications/connected-and-automated-vehicles-in-the-uk-2020-information-booklet

initiative has at least allowed some form of external learning (Sinha *et al.*, 2021) (Favarò *et al.*, 2017).[6]

One major analysis of AV trials around the world has explored and explained such tests according to their purposes, participants and the technologies involved (Dowling and McGuirk, 2020). This audit of what was, as of 2019, 135 AV trials concluded that, while some cities are playing more active roles in reshaping AV tests, many tests are not real tests at all, while the learning from others is intensely privatised. McAslan and colleagues, (2021) start their analysis of US AV trials from the perspective of city governments and find a common disconnect between AV developers' and cities' aims, reducing the trials' value in informing local transport strategies.

As trials have proliferated, showing more interest in performance than learning, one could argue that the real lessons have come from the moments of unarguable technological failure. An early misadventure on the part of Uber resulted in the death of the first bystander from an AV (Macrae, 2021) (Stilgoe, 2019). Uber's self-driving R&D arm subsequently became one of many loss-making operations to be acquired by a competitor. Other crashes involving Tesla (Stilgoe, 2018) have also been investigated and revealed not just the limits of the technology, but the governance challenges that lie ahead in rolling out safe and accountable mobility systems.

We should ask what all of these trials, tests and experiments add up to. Do they provide real-world evidence of the technology at work in diverse contexts? Do they show the range of companies and actors involved? Do they provide support for the idea that *the* technology is inevitable – a question of when, not if? Or does quantity have a quality of its own? Are these trials just a means of amassing data for machine learning or a way to impress the public and policymakers by, as Uber put it, "crushing miles"?[7] If the experiments were genuinely collaborative ones, we (and the companies involved) might see them as a source of social learning, but their performative role in the 'race for autonomy' obscures the potential for reflection.

City authorities should recognise the paradox produced by the dominant narrative of autonomy that surrounds self-driving vehicle innovation. The technology will be defined not by its purported 'autonomy', but by its attachments to a byzantine sociotechnical system that comprises infrastructures, sensors, safety drivers, data-labellers, emerging regulations, digital maps, other road users and more besides (Tennant and Stilgoe 2021). The more the technologies develop, the clearer the attachments become. Social research with members of the public has revealed that non-experts recognise unavoidable attachments ((Cugurullo *et al.*, 2021) (Stilgoe and Cohen, 2021)), but these are downplayed by innovators. Companies developing technologies and testing them on local streets are more tied to particular places than they would care to admit, but their eyes are on the distant horizon.

---

[6] The California DMV originally drafted regulations that demanded substantial collaboration between companies and the state on the terms and conduct of the social experiment. California announced in March 2017 their intention to relax these controls. Instead, they would put the onus on manufacturers to declare that their cars were safe and fully-insured and trust that the legal system would work out questions of liability and unintended consequences. The threat of being sued is imagined to be the strongest lever.

[7] Quote taken from 'I'm the operator', Wired magazine, 8 March 2022, https://www.wired.com/story/uber-self-driving-car-fatal-crash/

The narrative of autonomy is a story of an artificial intelligence learning to drive like a human, but better, at which point the world is its oyster. The ambition, as with other AI technologies, is to develop a platform that can be sold as software with monopolistic rents. Given the excitement about these technologies, it remains perplexing that none has a clear route to market. For instance, even a cursory economic analysis shows the fragility of the assumptions underpinning proposed business models for 'robotaxis' (Nunes and Hernandez, 2020).

Many of the technological demonstrations are in-principle rather than in-practice, designed to give the impression that a truly self-driving car is either already with us or 'just around the corner'. The line connecting current AV experiments to urban futures is faint, and neither innovators nor national policymakers are doing much to help clarify it. Innovators are more interested in solving narrowly defined technical problems than in addressing complex social needs – ends that are necessarily defined by others. Even as evaluation of roadworthiness, AV trials often reveal just how far their contexts are from a traditional driving test (see (Stilgoe, 2021)). While they may reveal little in terms of technological or mobility possibilities, these trials might at least demonstrate how claims to the relevance and proximity of the technology should be taken lightly.

For cities, the important political economic issues will be to do with whether the technology merely follows the money, exacerbating existing economic and transport injustices, or whether it can be emancipatory for particular groups who currently have few mobility options. Cities will want to know whether the technology will be able to pay its own way, demand public subsidies or, as with other platform technologies, need to be supported by a data-hungry advertising model. Cities should ask whether the technology can be reconfigured for local needs or whether it has path dependencies that enforce a take-it-or-leave-it standard. Questions about users, business models, winners and losers seem to have been deliberately postponed until after the technology is shown to 'work'. But the reality is that systems are tightly constrained by what their engineers call 'operational design domains'. If they 'work' within their own comfort zones, they may be incapacitated outside them. Diverse, variegated cities, may find that self-driving technologies suit only small areas. The problem may be that the potential benefits for road safety and opening up mobility access are greatest where the technical and economic challenge is hardest - in areas with poor infrastructure. The technologies are inextricably attached to their local environments, but have little to say to local decision makers wondering what the future of mobility looks like.

Conclusion – Opening up experimentation

A year after we observed the orange and white pods navigating the footpaths of East London's Olympic Park, the consortium returned for a final week of trials. This time the fleet of pods came installed with a remote safety stewarding system. According to the project's promotional material this freed up a seat in the pod resulting in "a larger passenger capacity and achieved a world first in operations of an autonomous vehicle without a safety driver on board in a public environment".

The strapline used on a summary video was clear about the intent and achieved outcome of the trials which focussed on "building regulatory and market confidence in autonomous pods". This hints at the recognition of some attachments, even as it neglects others.

Documents released in the project's final months emphasised prospective frameworks for verifying and validating emerging technologies in real-world settings. They pointed to the need for safety scenarios designed for testing amongst the public, not just in simulations, and cyber security reports emphasised the potential of future methodologies. Undoubtedly then, the trials produced *something*. However, much of what was being learnt, about the technology's limits and its likely relationships with others, did not fit the dominant narrative of innovation. Publicly, therefore, the narrative that what was needed was more of the same remained strong.

In the main, the knowledge produced through these trials is directed inwards towards consortia and industry interests, illustrating and occasionally strengthening relationships between players already in the AV sector. Outputs such as verification and validation frameworks provide justificatory arguments for ever more testing and ever more data gathering. But on their own they do little to tell us why or how these technologies contribute to public life.

Where the expertise, interests and knowledge of the public were included in the trials they were constrained. Experiments designed to test trust are a case in point. Social scientists surveyed more than 300 of the pod's passengers throughout the project, investigating whether, for example, facing forwards or backwards, or riding with a safety steward made a difference to how much riders trusted the technology. These results are perhaps useful for future design iterations, but they reveal almost nothing about the trustworthiness or accountability of emerging transport infrastructures.

These trials, and policymakers' enthusiasm for staging them, have important ramifications for debates about the governance of AI in and by cities. First, by asserting control over the conditions and outcomes of public trials and framing the future deployment of technologies as inevitable, AV trials preclude the possibility of failure. Second, by foregrounding interactions and attachments between members of the public and a discrete technology such as the pods, they are bypassing meaningful public engagement at a systemic level. These issues pose risks to the governance of urban technological systems in the long term. Unreflexive trials, even though they are temporary, help cement their subjects in urban space. If the danger is that we sleepwalk into technological change (Winner 1977), public trials can have a soporific effect, even though they hold the potential to alert their experimenters and others to important contingencies.

The lessons extend beyond governing the technology stack and into the city's political system. Trustworthiness, accountability and authority rest on the possibility of failure in experiments. In robotics, the trustworthiness and safety of autonomous systems are verified and validated against established and standardised knowledge in any given application domain such as city streets. But this kind of technological governance works only if someone is accountable for validating the veracity of the scientific (i.e. experimental) knowledge claims. Accountability in a broader political sense then rests on the vulnerability of political leaders to exposed failures of knowledge claims. In other words, the option of removing politicians from office is the quid pro quo demanded by citizens for allowing complex digital systems play such a pervasive role in our lives. Experiments or trials that foreclose the possibility of failure entirely risk diminishing the trustworthiness of experts who build and maintain autonomous systems and the politicians who ultimately govern them on our behalf. In this sense, the most important revelation from urban experiments with AVs is the scale of the work still left to do. Opening up the testing of AV technologies to make them relevant to urban decisionmaking will require a radical expansion of experiments' framing, apparatuses and participants. In the UK and in many other places, cities are already notionally involved in AV testing, but they have so far been reluctant to reframe experimentation so that questions of failure, technological limts and other infrastructural dependencies become the focus. If the results of such tests are going to be relevant for urban policy, such 'experiments' urgently need to become more experimental.

Acknowledgements

This research was supported by three funded projects: SCALINGS (Horizon 2020 grant 788359), Driverless Futures? (Economic and Social Research Council grant ES/S001832/1) and RAILS (Engineering and Physical Sciences Resarch Council grant EP/W011344/1). Thanks are due to our many collaborators on those projects, particularly Carlos Cuevas Garcia, Chris Tennant and Nuzhah Miah.

Copyright notice

References

boyd, danah (2016) 'Untangling research and practice: What Facebook's "emotional contagion" study teaches us', *Research Ethics*, 12(1), pp. 4–13. https://doi.org/10.1177/1747016115583379.

Burrell, J. (2016) 'How the machine "thinks": Understanding opacity in machine learning algorithms', *Big Data & Society*, 3(1), p. 2053951715622512. https://doi.org/10.1177/2053951715622512.

Collins, H.M. (1988) 'Public Experiments and Displays of Virtuosity: The Core-Set Revisited', *Social Studies of Science*, 18(4), pp. 725–748. https://doi.org/10.1177/030631288018004006.

Cugurullo, F. (2018) 'Exposing smart cities and eco-cities: Frankenstein urbanism and the sustainability challenges of the experimental city', *Environment and Planning A: Economy and Space*, 50(1), pp. 73–92. https://doi.org/10.1177/0308518X17738535.

Cugurullo, F. (2020) 'Urban Artificial Intelligence: From Automation to Autonomy in the Smart City', *Frontiers in Sustainable Cities*, 2. Available at: https://www.frontiersin.org/article/10.3389/frsc.2020.00038 (Accessed: 22 June 2022).

Cugurullo, F. (2021) *Frankenstein Urbanism: Eco, Smart and Autonomous Cities, Artificial Intelligence and the End of the City*. Routledge.

Cugurullo, F. *et al.* (2021) 'The transition to autonomous cars, the redesign of cities and the future of urban sustainability', *Urban Geography*, 42(6), pp. 833–859. https://doi.org/10.1080/02723638.2020.1746096.

Dowling, R. and McGuirk, P. (2020) ‘Autonomous vehicle experiments and the city’, *Urban Geography*, 0(0), pp. 1–18. https://doi.org/10.1080/02723638.2020.1866392.

Engels, F., Wentland, A. and Pfotenhauer, S.M. (2019) ‘Testing future societies? Developing a framework for test beds and living labs as instruments of innovation governance’, *Research Policy*, 48(9), p. 103826. https://doi.org/10.1016/j.respol.2019.103826.

Favarò, F.M. *et al.* (2017) ‘Examining accident reports involving autonomous vehicles in California’, *PLOS ONE*, 12(9), p. e0184952. https://doi.org/10.1371/journal.pone.0184952.

Gross, M. (2021) *Ignorance and Surprise: Science, Society, and Ecological Design*. MIT Press.

Hopkins, D. and Schwanen, T. (2018) ‘Automated Mobility Transitions: Governing Processes in the UK’, *Sustainability*, 10(4), p. 956. https://doi.org/10.3390/su10040956.

Karvonen, A. and van Heur, B. (2014) ‘Urban Laboratories: Experiments in Reworking Cities’, *International Journal of Urban and Regional Research*, 38(2), pp. 379–392. https://doi.org/10.1111/1468-2427.12075.

Krohn, W. and Weingart, P. (1987) ‘Commentary: Nuclear Power as a Social Experiment—European Political “Fall Out” from the Chernobyl Meltdown’, *Science, Technology, & Human Values*, 12(2), pp. 52–58. https://doi.org/10.1177/016224398701200206.

Laurent, B. (2011) ‘Technologies of Democracy: Experiments and Demonstrations’, *Science and Engineering Ethics*, 17(4), pp. 649–666. https://doi.org/10.1007/s11948-011-9303-1.

Laurent, B. and Tironi, M. (2015) ‘A field test and its displacements. Accounting for an experimental mode of industrial innovation’, *CoDesign*, 11(3–4), pp. 208–221. https://doi.org/10.1080/15710882.2015.1081241.

Leonardi, P.M. (2010) ‘From Road to Lab to Math: The Co-evolution of Technological, Regulatory, and Organizational Innovations for Automotive Crash Testing’, *Social Studies of Science*, 40(2), pp. 243–274.

Love Your Monsters | The Breakthrough Institute (no date). Available at: https://thebreakthrough.org/journal/issue-2/love-your-monsters (Accessed: 7 June 2022).

Macrae, C. (2021) ‘Learning from the Failure of Autonomous and Intelligent Systems: Accidents, Safety, and Sociotechnical Sources of Risk’, *Risk Analysis*, n/a(n/a). https://doi.org/10.1111/risa.13850.

Marres, N. (2020) ‘Co-existence or displacement: Do street trials of intelligent vehicles test society?’, *The British Journal of Sociology*, n/a(n/a). https://doi.org/10.1111/1468-4446.12730.

McAslan, D. *et al.* (2021) ‘Pilot project purgatory? Assessing automated vehicle pilot projects in U.S. cities’, *Humanities and Social Sciences Communications*, 8(1), pp. 1–16. https://doi.org/10.1057/s41599-021-01006-2.

McDowell-Naylor, D. (2018) *The Participatory, Communicative, and Organisational Dimensions of Public-Making: Public Engagement and The Development of Autonomous Vehicles in the United Kingdom*. PhD Thesis.

Michalec, O., O'Donovan, C. and Sobhani, M. (2021) 'What is robotics made of? The interdisciplinary politics of robotics research', *Humanities and Social Sciences Communications*, 8(1), pp. 1–15. https://doi.org/10.1057/s41599-021-00737-6.

Nunes, A. and Hernandez, K.D. (2020) 'Autonomous taxis & public health: High cost or high opportunity cost?', *Transportation Research Part A: Policy and Practice*, 138, pp. 28–36. https://doi.org/10.1016/j.tra.2020.05.011.

Paddeu, D., Shergold, I. and Parkhurst, G. (2020) 'The social perspective on policy towards local shared autonomous vehicle services (LSAVS)', *Transport Policy*, 98, pp. 116–126. https://doi.org/10.1016/j.tranpol.2020.05.013.

Ryghaug, M. *et al.* (2022) 'Testing Emergent Technologies in the Arctic: How Attention to Place Contributes to Visions of Autonomous Vehicles', *Science & Technology Studies* [Preprint]. https://doi.org/10.23987/sts.101778.

Savini, F. and Bertolini, L. (2019) 'Urban experimentation as a politics of niches', *Environment and Planning A: Economy and Space*, 51(4), pp. 831–848. https://doi.org/10.1177/0308518X19826085.

Sinha, A. *et al.* (2021) 'A Crash Injury Model Involving Autonomous Vehicle: Investigating of Crash and Disengagement Reports', *Sustainability*, 13(14), p. 7938. https://doi.org/10.3390/su13147938.

Stilgoe, J. (2017) 'Seeing Like a Tesla: How Can We Anticipate Self-Driving Worlds?', *Glocalism: Journal of Culture, Politics and Innovation* [Preprint], (3). https://doi.org/10.12893/gjcpi.2017.3.2.

Stilgoe, J. (2018) 'Machine learning, social learning and the governance of self-driving cars', *Social Studies of Science*, 48(1), pp. 25–56. https://doi.org/10.1177/0306312717741687.

Stilgoe, J. (2019) *Who's Driving Innovation?: New Technologies and the Collaborative State*. Springer Nature.

Stilgoe, J. (2021) 'How can we know a self-driving car is safe?', *Ethics and Information Technology* [Preprint]. https://doi.org/10.1007/s10676-021-09602-1.

Stilgoe, J. and Cohen, T. (2021) 'Rejecting acceptance: learning from public dialogue on self-driving vehicles', *Science and Public Policy*, p. scab060. https://doi.org/10.1093/scipol/scab060.

Taylor, L. (2021) 'Exploitation as innovation: research ethics and the governance of experimentation in the urban living lab', *Regional Studies*, 55(12), pp. 1902–1912. https://doi.org/10.1080/00343404.2020.1826421.

Tennant, C. and Stilgoe, J. (2021) 'The attachments of "autonomous" vehicles', *Social Studies of Science*, p. 03063127211038752. https://doi.org/10.1177/03063127211038752.

While, A.H., Marvin, S. and Kovacic, M. (2021) 'Urban robotic experimentation: San Francisco, Tokyo and Dubai', *Urban Studies*, 58(4), pp. 769–786. https://doi.org/10.1177/0042098020917790.

Winner, L. (2001) *Autonomous technology: technics-out-of-control as a theme in political thought*. 9. printing. Cambridge, Mass.: MIT Pr.

Wu, X. (2022) 'Fast forward: technography of the social integration of connected and automated vehicles into UK society'. https://doi.org/10.7488/era/2320.