D-RISK COMMUNITY INSIGHTS REPORT

D-RISK

INTRODUCTION

OVERVIEW OF D-RISK

CREATING THE WORLD'S LARGEST LIBRARY OF AUTONOMOUS VEHICLE EDGE-CASES

Autonomous vehicles (AVs) have the potential to radically reshape mobility. As well as improving road-safety, AVs provide a unique opportunity to improve service quality and accessibility for citizens, including those with disabilities, and those unable to drive.

For the AVs of tomorrow to operate safely, it is critical that developers today take account of the complexity of modern roads and their users and build technologies that are able to handle even the most unique and challenging of on-road scenarios. Building future AV service users into the development of this new technology is therefore vital.

To aid the development of safe AV services D-RISK is building the world's largest library of driving "edge cases". These are situations that aren't commonplace, they're unusual or unexpected but, as they could be dangerous, AVs will need to be able to handle them safely. By gathering data from multiple sources, including stories from the public, D-RISK is training autonomous vehicle artificial intelligence (AI) through a process of machine learning and simulation. Edge cases are being gathered from traffic cameras, road accident reports and through a crowdsourcing initiative with the UK public.

In this report we explore the findings of community research into public perceptions of AVs and we share the results of a unique approach to gathering real-world edge-cases to build the D-RISK edge-case library.

D-RISK Partners

D-RISK is a £3m Innovate-UK funded project focusing on the development of software for safe automated vehicles. It is a collaboration between dRisk.ai, Claytex, Imperial College London and DG Cities.

Figure 1: D-RISK simulation and edge case example

COMMUNITY RESEARCH

Deep public engagement is critical to the successful trialling and development of AV technologies. It is for this reason that D-RISK has worked closely with the public to capture their experiences and measure their views of key aspects of the developing technology and software. In this study we explore the following research questions:

- What is the general perception of AVs amongst the UK population? e.g., perceptions of safety, reliability, trustworthiness, predictability, and confidence of an AV in an emergency situation.
- What does the UK population believe is an 'appropriate' response for an AV to take within an edge-case scenario?

METHODOLOGY

We deployed a mix of quantitative and qualitative methods that consisted of an online survey and online focus groups.

ONLINE SURVEY

A survey was distributed online between December 2020 and February 2021. It was free and accessible via any digital device, promoted across social media. The survey was split into four sections:

- Video simulation of an edge case with a human and an AV response.
- Perceptions of AVs and ride-sharing services.
- Individual edge-case disclosure.
- General demographic questions.

FOCUS GROUP

Five online focus groups were conducted between March 2021 and May 2021, engaging 46 participants. Participants in this session took part in three activities:

- Rating of AV safety, trust and accessibility to capture perceptions and attitudes.
- Discussion following AV simulation video demonstration.
- Edge-case disclosure discussion

FINDINGS

Key findings

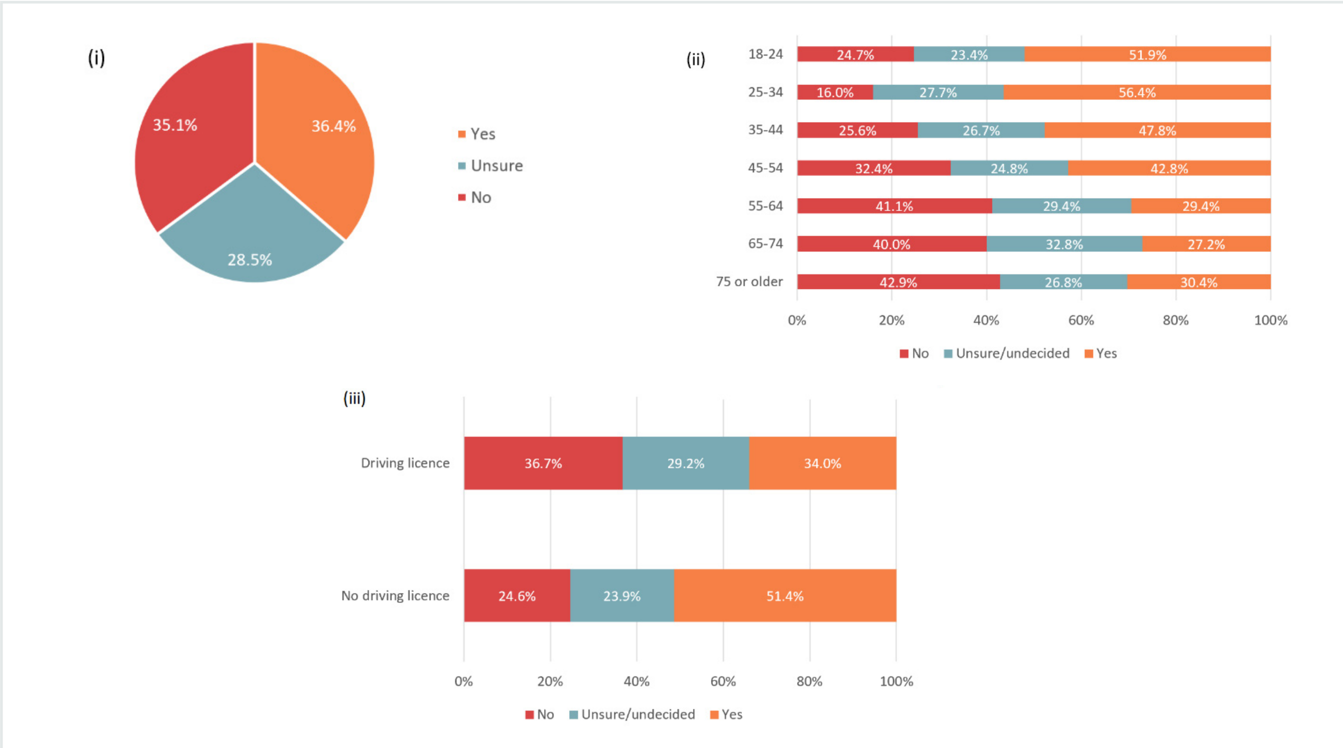
- Over a third (36.4%) are happy to ride in an AV tomorrow. Almost 3 in 10 (28.5%) are undecided so could be persuaded.
- There were significant differences in perceptions of AVs between younger and older groups with regards to safety and trust. In general, young people viewed AVs more favourably than their older counterparts.
- Overall, simulated human responses in the simulation exercise were judged to be more dangerous, more unpredictable, and slower than AV responses to the same scenario.

ANALYSIS

In total, the survey received 1,034 viable responses. A weight variable was calculated and applied to the data set to improve equivalence with the UK population by gender and age.[1] More information about the diversity of respondents can be found in the appendix.

DESIRE TO RIDE IN AN AV

Figure 2: Desire to ride in an AV (i) total, (ii) by age, (iii) by licence status (Base 1038 weighted)

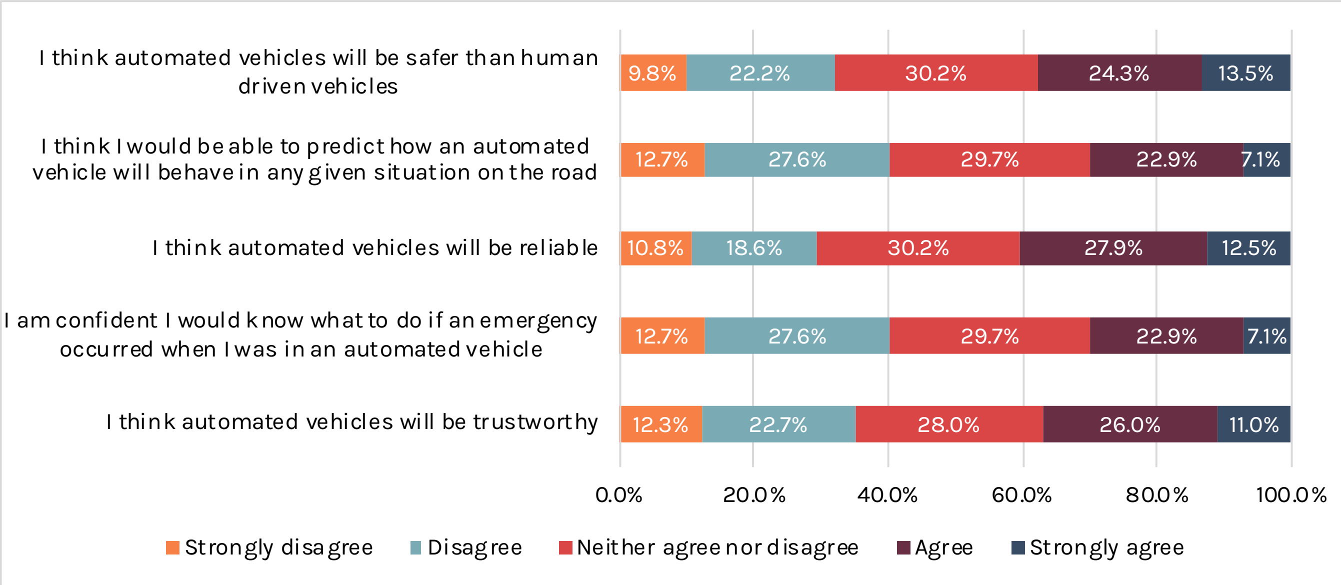


[1] A weighting variable was calculated through a process of raking (or proportional fitting) using the American National Election Study weighting algorithm ANESRAKE. The weights are calculated so that the survey marginals closely match UK population marginals for Gender and Age. More information can be found here: https://web.stanford.edu/group/iriss/cgi-bin/anesrake/resources/RakingDescription.pdf

Over a third (36.4%) of respondents were happy to ride in AVs in the near future, however almost three in ten (28.5%) were undecided. Those without driving licences were also more likely to want to ride in an AV (51.4%) than those with driving licences (34.0%).

We also measured five areas of AV perception to understand specific public perceptions of AV technology.

Figure 3: Public perceptions of AV



Base: 1038 (weighted)

SAFETY

Our data shows that perceptions of AV safety differ by age, with some group differences statistically significant. For example, those aged 25-34 were statistically significantly more positive about the safety of future AVs compared to human driven cars than older groups (55-64, 65-74, 75+).[2]

Perceptions of AV safety also differed by ethnicity when calculated as white versus non-white. People who identified as non-white were statistically significantly more likely to say they believed AVs will be safe compared to those who identified as white.[3]

We found a very weak positive relationship between risky road behaviour score and belief that AVs will be safer than human driven vehicles. The relationship was statistically significant.[4] We found no significant difference in participants' perception of AV safety between genders.

Participants in focus groups highlighted the potential for AVs to outperform human drivers in terms of safety:

[2]One way ANOVA was conducted for age groups: $F(6,1029) = 9.102$, $p < 0.01$, 25-34 ($M = 3.53$ $SD = 1.112$) compared to older ages, 55 – 64 ($M = 2.82$ $SD = 1.226$), 65 - 74 ($M = 2.83$ $SD = 1.131$), 75+ ($M = 2.79$ $SD = 1.137$)

[3]An independent samples T-Test was conducted between white and non-white: There was a statistically significant difference in the scores for white people ($M=3.04$ $SD=1.181$) versus non-white people ($M=3.49$ $SD=1.070$) for AV reliability; $t(1018) = -4.327$, $p = <0.001$

[4]$r = 0.166$, $N = 1038$, $p = < 0.001$

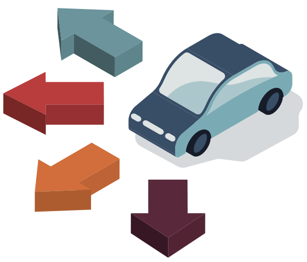
There might be potential for them to be safer than many drivers but it's going to require a lot more work. But surely the safety standard they're aiming for is zero collisions, and zero deaths, and until they get there there's more work to do. Focus group participant

Others believe that the safety of AVs is dependent on the context and situation it is operating in. Participants gave examples of situations where AVs may be safer than human driven vehicles:

I was raving about driverless cars. I think they're the greatest thing since sliced bread. But I need to say that I think that applies to motorways. I'm not so keen on AVs when it comes to built-up towns. I think it's going to be more efficient to have a human being, perhaps. Focus group participant

RELIABILITY

Perceptions of reliability matter if AVs are to be readily adopted by the public. Younger groups (18-24, 25-34, 35-44) were more likely to agree that AVs are reliable and were statically significantly different to older groups (45-54, 55-64, 75+), who were less likely to agree.[5] Improving perceptions of reliability may be important to build buy-in amongst older people.



PREDICTABILITY

Predicting how road users behave is an important skill for human drivers, and one which AVs are required to master. For human drivers the predictability of AVs is a key component of their willingness to adopt the technology. Our data shows that AV predictability differ across age groups. Younger people (25-34, 35-44) were more likely to agree that they could predict the behaviour of AVs and were statically significantly different to older people (55-64, 65-74, 75+), who were less likely to agree.[6] This illustrates that older member of the community may need additional support or assistance to build their confidence in predicting AV behaviour.

Focus group participants noted that AVs will need to manage unpredictable road users if they're to operate effectively:

You can't control how other people drive. How does the driverless car interact with learner drivers on the road? Or perhaps new drivers? How does the driverless car deal with that situation? Focus group participant.

When asked whether they could drive on roads where AVs were operating, one participant saw a major challenge being the transition period:

It's a scary thing, you know, it's rather like when Sweden changed over to right-hand side driving. It was an overnight success and I'm sure there was a hell of a lot of fear going on because people weren't understanding the difference. I think unless you've got everybody in driverless cars, it could be an issue.
Focus group participant

---

[5] One way ANOVA: $F(6,1029)= 11.708$, $p < 0.01$,, showed that there were significant differences between 2 groupings of age, with younger ages 18-24 (M = 3.41 SD = 1.115) 25 – 34 (M = 3.50 SD = 1.155), 35 – 44 (M = 3.42 SD = 1.015), seeing vehicles as more reliable than older groups 45 - 54 (M = 3.07 SD = 1.205), 55 – 64 (M = 2.80 SD = 1.219), 65 - 74 (M = 2.79 SD = 1.105), 75+ (M = 2.84 SD = 1.135)

[6] One way ANOVA: $F(6,1029)= 5.896$, $p < 0.01$, n2p = ? showed that there were significant differences between 2 groupings of age, with younger ages 25 – 34 (M = 3.11 SD = 1.124), 35 – 44 (M = 3.07 SD = 1.) being more confident at predicting AVs compared to older groups, 45 - 54 (M = 2.96 SD = 1.243), 55 – 64 (M = 2.67 SD = 1.171), 65 - 74 (M = 2.65 SD = 1.141), 75+ (M = 2.66 SD = 1.103).

TRUSTWORTHINESS

Trust is a critical factor that will influence the adoption and acceptance of AVs. Our survey and focus group data highlight some key challenges for developers to overcome. For example, we found different perceptions of trustworthiness across age groups. Younger groups (18-24, 25-34, 35-44) were more trusting of AVs and were statically significantly different to older groups (55-64, 65-74, 75+), who were broadly more negative and less trusting of the technology.[7]

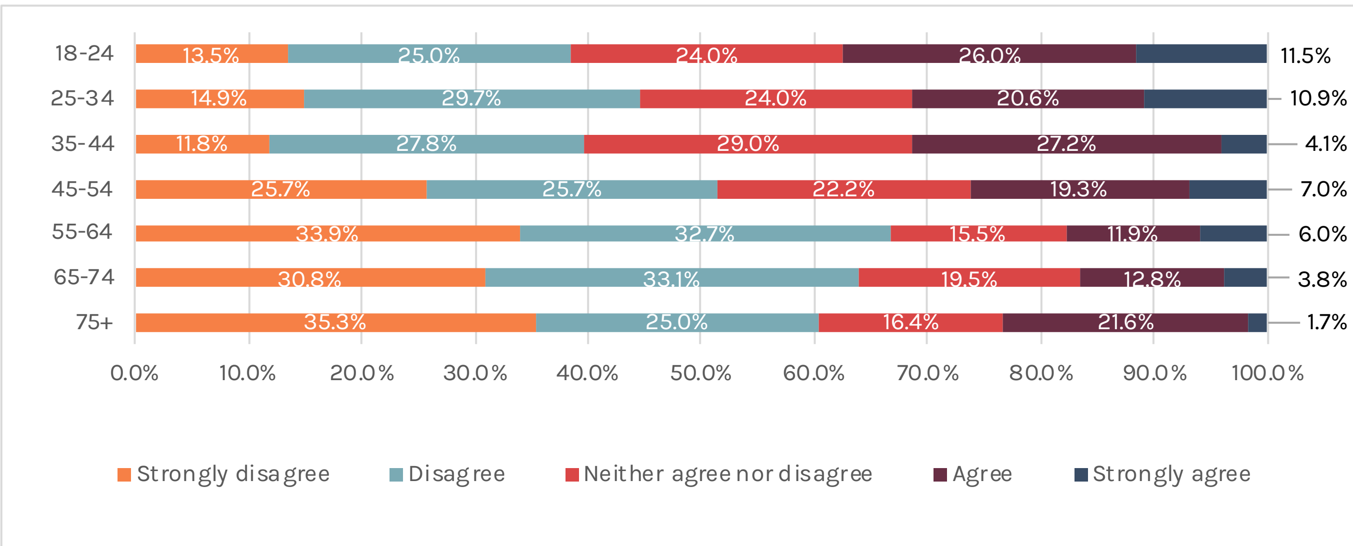
There was also a significant difference in perceptions of AV trustworthiness between ethnicities when calculated as white vs non-white. Participants who identified as non-white were statistically significantly more likely to report that they would trust AVs, compared to those who identified as white.[8]

> *"I'm just frightened to death of driverless cars. I drive a 125 scooter, I have a car, and I used to drive a massive one-toner work van, but the thought of driverless cars just fills me with horror. I wouldn't even like to be a passenger on a trial run. I just don't trust them. I don't know how they're supposed to work."* Focus group participant

INTENTIONS OF USING AVS IN THE FUTURE

Intentions to use AVs in the future are shaped by a number of factors. Our study explored how attitudes towards future-use of the technology differed across demographic groups.

Figure 4: Intentions to use an AV in the future, by age



Base: 1037 (weighted)

[7] One way ANOVA: F(6,1029)= 13.125, p < 0.01, showed that there were significant differences between 3 groupings of age, with younger ages 18-24 (M = 3.35 SD = 1.128), 25 – 34 (M = 3.41 SD = 1.191), 35 – 44 (M = 3.29 SD = 1.043), being more trustworthy of AVs, compared to older ages, 45 - 54 (M = 2.96 SD = 1.243), 55 – 64 (M = 2.67 SD = 1.171), 65 - 74 (M = 2.65 SD = 1.141), 75+ (M = 2.66 SD = 1.103)

[8] An independent samples T-Test was conducted between white and non-white: There was a statistically significant difference in the scores for white people (M=2.94 SD=1.193) versus non-white people (M=3.48 SD=1.064) for trustworthiness; t(1018) = -5.152, p = <0.001

Younger groups (18-24, 25-34, 35-44) were statically significantly more likely to agree that they would, when compared to older groups (55-64, 65-74, 75+).[9] We also found that people who identified as disabled were less cautious than non-disabled people about riding in an AV in the near future.[10]

When asked about the future potential of AV services, some participants recognised a risk in the complexity of modern towns and cities, and questioned whether or not the technology could operate within them:

> *"I just think the environment we're trying to use these in is just far too complicated. Aircraft have air corridors at set heights, but you think on roads you've got a combination of multiple road users, changes to road layout, weather conditions. The actual task is immense so to make it work maybe you need to try and simplify the road layout, but then you'll end up putting everything on rails wouldn't you."*
>
> *Focus group participant*

Other participants recognised the potential value of AV technology to improving how they spend their time, for example those who commute by car could see themselves gaining back valuable leisure or work time in which they could be more productive:

> *"There's a difference between essential journeys and driving for pleasure. I probably do a mixture of the two, I go out because I enjoy driving, but I also have the necessary journeys whilst I was at work, commuting. If I could've put over commuting to the vehicle, then perhaps I could've been using the one hour's drive each way more productively if the vehicle were safe to do that. I could've continued working as a lot of commuters on trains do."*
>
> *Focus group participant*

---

[9] One way ANOVA: $F(6,1029)= 9.814$, $p < 0.01$, showed that there were significant differences between 2 groupings of age, with younger ages 18-24 ($M = 2.98$ $SD = 1.229$) 25 – 34 ($M = 2.83$ $SD = 1.22$), 35 – 44 ($M = 2.85$ $SD = 1.081$) differing when compared to older groups. 55 – 64 ($M = 2.23$ $SD = 1.211$), 65 - 74 ($M = 2.25$ $SD = 1.141$), 75+ ($M = 2.29$ $SD = 1.207$). The 45 - 54 ($M = 2.56$ $SD = 1.251$) group was not statistically significantly different to either group.

[10] T-test: Disabled ($M = 2.74$, $SD = 1.320$); Not disabled ($M = 2.51$, $SD = 1.180$); $t(1036) = 2.765$, $p = 0.06$; Effect size $(d) = 0.192$

PERCEPTION TEST: HUMAN AND AV SIMULATION

Blind testing

By doing a blind test we were able to test whether prior knowledge of the vehicle controller (AV or human) influenced assessments of the vehicle behaviour.

We tested three pairs of simulation videos, each pair representing a common road scenario. In one video the participant observed a simulation based on a human response to a real-world scenario re-constructed from UK traffic camera data. In the second video the participant watched an AI response to the same scenario.

The participant was blind to which response they were observing. Participants then rated the vehicle's behaviour on five scales: safety, predictability, avoidance capability, decision speed and humanity between zero and ten.[11] Below we detail the mean responses from participants:

SUDDEN STOP SCENARIO

A video of a sudden stop scenario in which a vehicle has to brake suddenly to avoid collision was shown to participants. The results are below:

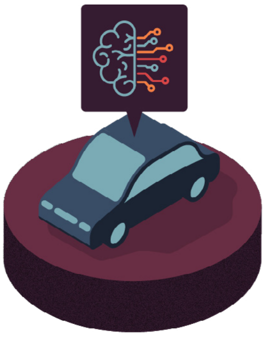
Figure 5: Sudden stop scenario (mean scores, e.g., 0 = Unsafe, 10 = Safe)

Base: Safety n = 299**, Predictability n = 293**, Avoidance capability = 289**, Decision speed = 284**, Humanity = 273. ** = <0.001 significance.

Our findings show that the public were more positive towards the AV's response, with safety, decision speed and avoidance capability being the biggest differences between videos.

The data highlights that individuals were unable to tell the difference between an AV and a human response, and that they were generally more positive about the behaviour of the AV in the simulation exercise.

[11] Participants scored the video on the following scales: Dangerous (0) – Safe (10), Unpredictable (0) – Predictable (10), Incorrect avoidance (0) – Correct avoidance (10), Slow decision speed (0) – Fast decision speed (10), AI (0) – Human (10).



VEHICLE OVERTAKING SCENARIO

A video of a vehicle overtaking scenario, in which the vehicle overtakes a bicycle, was shown to participants. The results are below:

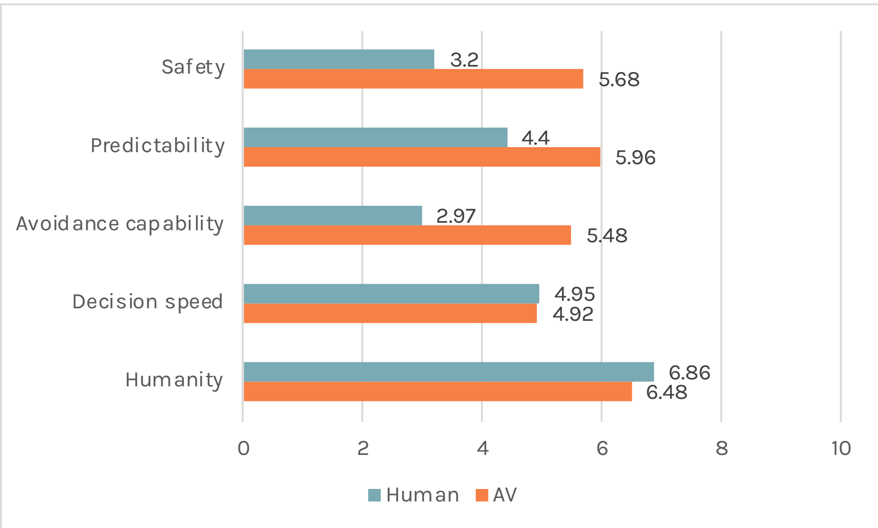
Figure 6: Vehicle overtaking scenario (mean scores, e.g., 0 = Unsafe, 10 = Safe)

Base: Safety n = 341**, Predictability n = 337**, Avoidance capability = 328**, Decision speed = 319**, Humanity = 315**. ** = <0.001 significance.

VEHICLE TURNING RIGHT

Finally, a video of a vehicle turning right in which the vehicle must wait for oncoming traffic to pass and a pedestrian to cross was also shown to participants. The results are below:

Figure 7: Vehicle turning right (mean scores, e.g., 0 = Unsafe, 10 = Safe)



Base: Safety n = 359**, Predictability n = 352**, Avoidance capability = 343**, Decision speed = 303**, Humanity = 345. ** = <0.001 significance.

KEY FINDINGS

Our data highlights several important findings:

- The human response in all three scenarios was judged to be more dangerous than the AV response. This difference was significant.
- The AV response was judged to be slightly more predictable than the human response in all videos. In some edge case scenarios this difference was significant (vehicle sudden stop and vehicle turning right).
- In all videos the AV response was judged by participants to have made a more correct avoidance manoeuvre than the human response. This difference was significant and substantial.
- The AV response was judged to involve faster decision making than the human response in two of the three scenarios. The difference in the two scenarios was significant and substantial.
- Both AV and human responses were judged by participants to be more similar to a human response than an AV response. However, the score was only significantly different from average (5) in one scenario.

We found that participants were unable to differentiate between the AV response and the human simulated response. This illustrates potential for using simulations to both build public awareness and demonstrate the capability of AV artificial intelligence.

How simulations can benefit AV development.

D-RISK combines patented knowledge graph technology with advanced sensor-realistic environment simulations, state-of-the-art computer vision, and behavioural testing to identify new scenarios that push AVs to the limit of their capabilities.

The project is identifying the scenarios that AVs struggle with today, and using them to retrain AVs to react more safely. Trials using the D-RISK training programme have already resulted in AVs reacting six times more quickly, and with twice the confidence, to high-risk situations.

EDGE-CASE LIBRARY

Do you have an edge-case to share?

We are continuing to build the edge-case library in 2022. Members of the public can continue to disclose their edge cases via the D-RISK website.

Visit www.drisk-project.org to find out more.

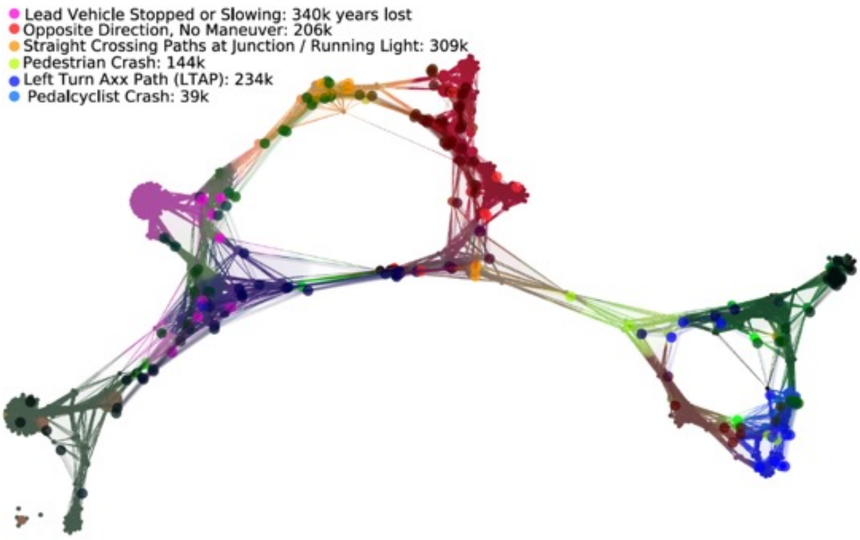
The survey we developed also collected stories from the public to help build the edge-case library. By asking four simple questions we were able to gather useful insights that were used by the D-RISK machine learning algorithm to create new scenarios for testing and development.
The four disclosure questions we asked were:

- Describe the situation (e.g., what vehicles or road users were involved? what was the weather like?)
- What did you do?
- What did others do?
- What was the outcome?

In total 411 edge-cases were supplied by participants. Almost a third (30%) were near-miss scenarios which are not captured through traditional reporting methods, such as accident reports. The public D-RISK edge-cases were interpreted within the drisk.ai software and added to the knowledge graph of potential edge-cases, alongside those being developed from traffic cameras and reported incidents.

Below we illustrate how the D-RISK knowledge graph represents the data we collected.

Figure 8: An example of the D-RISK knowledge graph



DISCUSSION

Our data provides some important findings for the developers of autonomous vehicle technologies, policy makers and the wider public.

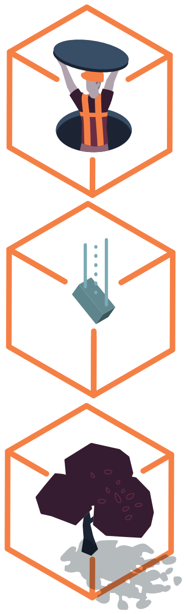
Over a third (36.4%) of respondents are willing to use an AV. Almost three (28.5%) in ten are unsure. We found a willingness to use AVs in a majority of responses, but a large proportion do not wish to use the technology. The large number of unsure respondents highlights the potential for persuading undecided members of the public to use the technology. More work however is needed to understand the barriers to adoption for those negative towards the technology.

AV responses were judged by the public to be more acceptable than human responses. The AV simulations we tested consistently performed more favourably in terms of safety and speed of response in comparison to a simulated 'typical' human response for the same scenario. This is encouraging for the development of AV software and illustrates the value of public demonstrations of AV simulations as a method for building acceptance and understanding of the technology.

General perception of AVs is low but differ across groups. Our data shows that although there is an appetite for AVs amongst some groups, general perception of AV trustworthiness, predictability, reliability and safety are low. These differences highlight some important areas for future research and policy engagement. Deeper public engagement which is targeted towards specific groups is needed to support public acceptance.

Younger age groups were generally more positive about AVs and their introduction in the future. They were also more likely to find AVs trustworthy, predictable, reliable and safe, compared to older groups.

Willingness to use an AV in the future was more positive amongst disabled participants in comparison to non-disabled populations. This may in part be due to recognition of the possibility a driverless car to offer increased accessibility, and therefore a readiness to trust in the technology for the benefits it could deliver in the future.



CONCLUSION

It is vital that the public is supported to play an active role in the development of future AV technologies. Our public engagement research shines a light on directly engaging the public in AV innovation and shows that there are groups that require tailored education, information and support to build their knowledge of AV technologies and the AI that operates them.

As well as providing vital data to improve the quality of services and reduce software bias, the experiences of road users can help to ensure the development of safe and trusted technologies. Public outreach and engagement are therefore an important tool for overcoming bias, reducing risk and ultimately ensuring that the deployment of new AV technologies is for the benefit of everyone.

ABOUT THE AUTHORS

DG CITIES

DG Cities is an urban innovation consultancy, specialising in helping clients harness the power of technology and data to transform our towns and cities. Within the D-RISK project, DG Cities is ensuring that public opinions are heard and shape the development of AV technology. DG Cities has invited the public to share their views on what is 'appropriate' behaviour for automated vehicles through online surveys and focus groups.

DG Cities also lead the work building the public edge-case library to ensure that real-life experiences are programmed - and appropriate responses generated - within AV software.

Ed Houghton, Head of Research and Service Design at DG Cities
Ed is a thought leader in systems-thinking, system resilience, and AI in different contexts. He is a mixed-methods researcher who specialises in evidence-based policy and practice development. Ed leads the research and service design practice at DG Cities.

Hiba Alaraj, Project Manager at DG Cities
Hiba brings to DG Cities a background in master planning, project management and sustainability within the private sector. She has expertise in the energy and sustainability aspects of smart city innovation and is currently working on net-zero carbon, social sustainability and mobility related projects.

APPENDIX
DEMOGRAPHICS

Base: 1034 respondents

Figure 9: Age

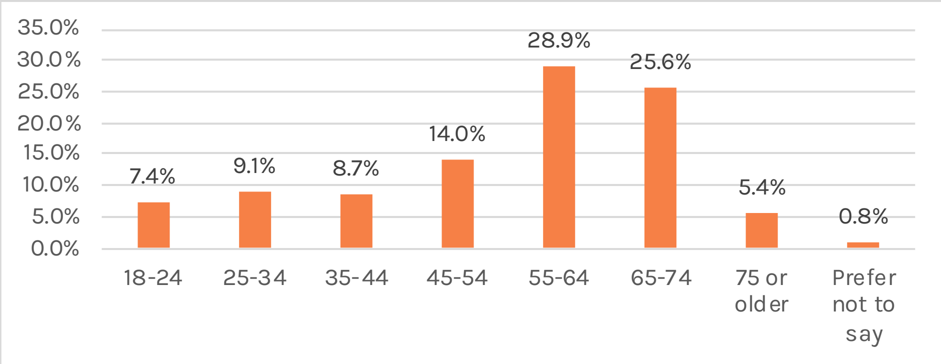


Figure 10: Gender

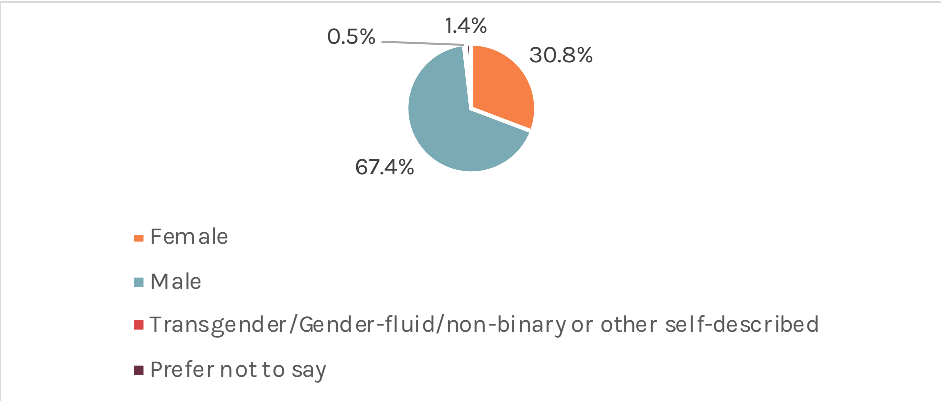


Figure 11: Ethnicity

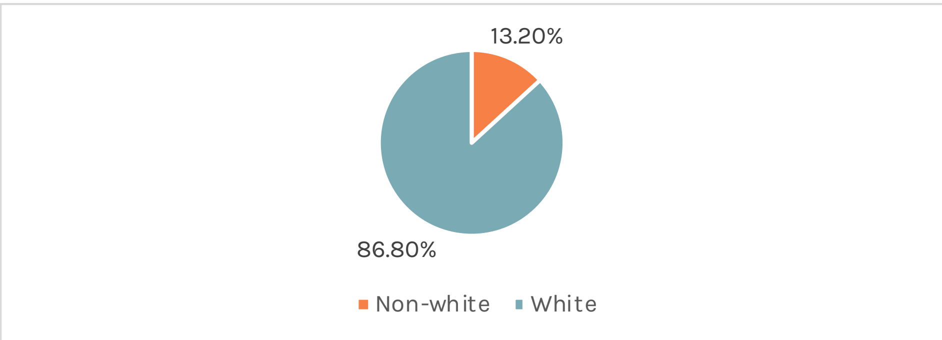
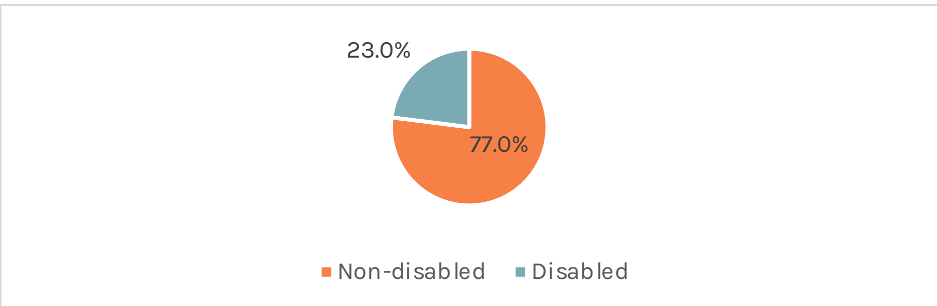


Figure 12: Disability status



D-RISK

TO FIND OUT MORE ABOUT D-RISK VISIT

WWW.DRISK-PROJECT.ORG