Keele University

The Presentation and Perception of Automotive AI in Fictional and Non-Fictional Narratives

Alexander Farrell

Computer Science with Integrated Masters

Monday 29th April 2019

SCHOOL OF COMPUTING AND MATHEMATICS
Keele University
Keele
Staffordshire
ST5 5BG

I. Abstract

It is commonly misconceived that the automobile industry can be lethargic in its adaptation to new technologies, including artificial intelligence. However, in reality there are a number of new technologies that are deployed through vehicles that are then translated to other industries (Gusikhin, et al., 2007). This realisation correlates to the concerns that people may have regarding new technology and their subsequent acceptance.

Public acceptance can be influenced by a number of factors, one of which is the presentation and perception of technology in both fictional and non-fiction narratives. A study conducted by Cave, et al. identified 8 factors categorized as ‘hopes and fears’ that are commonly represented in narratives.

The main aim of this project is to evaluate the extent to which overpowering images of automotive AI can be mitigated by emphasising real, current application, or narratives of control and involvement. This idea stems from the recommendation of Cave, et al. who suggests the investigation of the impact of alternative narratives on public perception (2019).

II. Acknowledgements

I would like to that Dr Theocharis Kyriacou for being my project supervisor and for offering support and guidance whenever it was required. I would also like to thank all the students who took time out of their own studies to take part in the online survey and face-to-face interviews which provided the backbone for this project.

Contents

I. Abstract 1
II. Acknowledgements 2
Contents 3
1. Introduction 5
1.1. Aim 5
1.2. Objectives 6
1.3. Potential Limitations 7
2. Literature Review 8
2.1. Public Perception of AI 8
2.2. AI Presentation in the Media 10
2.3. Concern towards AI Systems 13
2.4. AI Discourse 14
2.5. Ethics of Responsibility 16
3. Methodology 20
3.1. Research Design 20
3.2. Online Survey 21
3.2.1. Sample 21
3.2.2. Content 22
3.2.3. Ethical Considerations 23
3.3. Face-to-Face Interview 23
3.3.1. Sample 24
3.3.2. Content 24
3.3.3. Ethical Considerations 25
4. Results 26
4.1. Online Survey 26
4.1.1. Awareness of Artificial Intelligence 26
4.1.2. Radar-Guided Cruise Control 27
4.1.3. Automated Braking Assistance 29

4.1.4. Collision Avoidance System 30
4.1.5. Automated Parking 31
4.1.6. Self-Driving 32
4.2. Face-to-Face Interview 33
4.2.1. Radar-Guided Cruise Control 34
4.2.2. Automated Braking Assistance 34
4.2.3. Collision Avoidance System 35
4.2.4. Automated Parking 35
4.2.5. Self-Driving 36
5. Discussion 38
5.1. Science-Fact vs Science-Fiction 38
5.2. Awareness vs Trust 38
5.3. Average Increase 39
6. Conclusion 41
6.1. Meeting the Aim 41
6.2. Meeting the Objectives 41
6.3. Overcoming Limitations 42
6.4. Further Work 43
7. References 44
8. Appendix 48
8.1. Appendix 1 – Online Survey Facebook Advert 48
8.2. Appendix 2 – Online Survey 49
8.3. Appendix 3 – Online Survey Consent Form 55
8.4. Appendix 4 – Interview Script 57
8.5. Appendix 5 – Interview Information Sheet 58
8.6. Appendix 6 – Interview Consent Form 60
8.7. Appendix 7 – Literature Review Structure (Hierarchy) 61
Project Plan 62
Ethics Form 65

Word Count: 10018

1. Introduction

It is commonly misconceived that the automobile industry can be lethargic in its adaptation to new technologies, including artificial intelligence. However, in reality there are a number of new technologies that are deployed through vehicles that are then translated to other industries (Gusikhin, et al., 2007). This realisation correlates to the concerns that people may have regarding new technology and their subsequent acceptance.

Public acceptance can be influenced by a number of factors, one of which is the presentation and perception of technology in both fictional and non-fiction narratives. A study conducted by Cave, et al. identified 8 factors categorized as ‘hopes and fears’ that are commonly represented in narratives. Their study explored what the UK believe AI is, and how much they subscribe to utopian and dystopian narratives.

1.1. Aim

The aim of the project is to evaluate the extent to which overpowering images of automotive AI can be mitigated by emphasising real, current applications, or narratives of control and involvement. This idea stems from the recommendation of Cave, et al. who suggests the investigation of the impact of alternative narratives on public perception (Cave, et al., 2019).

Cave explains that narratives presenting less-extreme conceptions, could investigate whether frightening and overpowering images of AI could be mitigated by emphasising real, current applications, or narratives of control and involvement. It is possible that

narratives allowing audiences to imagine an active role in the development or deployment of AI can instil a sense of empowerment.

1.2. Objectives

To ensure the aim of the project is thoroughly investigated, a systematic literature review will be conducted to explain factors such as ethics and trust that can influence a person's perception. This review will look to incorporate real-world issues and representations of automotive AI, as well as various fictional representations of current or future technologies. The intended outcome of this literature review is to identify a broad selection of concerns that the public may have regarding the deployment of automotive AI.

These concerns will then be ranked in terms of their level of identity amongst the public through the deployment of an online survey. It will be interesting to see the level at which the public is aware of these concerns as according to Fast and Horvitz, although media presence of AI has increased rapidly since 2009, discussions have been more optimistic than pessimistic (2017). However, their study reveals that pessimism towards specific concerns such as loss of control, ethics and the negative impact on work, has grown.

Once the most prominent concerns have been identified, the task will be to explore real-world applications that depict these concerns in a more positive light. In some cases, this may not be entirely truthful as there may be some concerns that relate to future technologies. Therefore, in this situation, the assumption is made that if a concern has been recognised by the public, there is likely to be a system in development or narrative that is working towards a real-world application relating to it.

The main stage of the investigation will look to investigate whether the concerns identified can be somewhat mitigated by exposure to the real-world applications or narratives that depict them in a more positive light. This will be achieved by conducting semi-structured interviews where subjects will be presented with an explanation of a particular concern or concern and then shown a real-world application or narrative to judge whether this exposure can decrease their pessimism towards this concern.

1.3. Potential Limitations

As mentioned above, it is likely there will be some concerns and concerns that relate to future technology to which a real-world example may be legitimately sourced. Even though the assumption made allows for fictional narratives, a hypothesis surfaces that the effect of fictional narratives will not be as strong as real-world applications.

From an ethical perspective, consideration must be made towards concerns expressed that could relate to such things as car accidents which could have recall for participants who might have negative associations with such incidents. To minimise this risk, a screening question or warning can be placed in the consent form for the survey and interview stage of the research.

2. Literature Review

To properly understand the concerns that people have towards automotive AI, one must understand the different aspects that can draw these concerns, including formal aspects such as legal and ethical issues, as well as more social aspects such as trust and perception. The research conducted within this project will also involve getting the thoughts and opinions of real subjects, therefore exploration of proper research methods is also imperative. Allowing development of understanding within these areas will create a strong foundation for the study, as well as offer support to subsequent conclusions.

2.1. Public Perception of AI

The presentation and perception of automotive Artificial Intelligence can be translated from literature consulting public perception of AI becoming a more integral part of everyday life. A recent publication emphasising the metaphor of "Scary Robots" (Cave, et al., 2019) explains the results of a nationally representative survey of the UK population on their perceptions of AI. The term 'nationally representative' allows the researcher to cover multiple variables such as age, gender, etc., as perceptions may differ depending on the demographic of the subjects. The significance of this research is highlighted as "how AI is perceived can have significant impact on how it is developed, deployed and regulated", meaning that if the public receive a negative presentation of AI from narratives, they are less likely to be encouraged by development with is making AI smarter and arguably more independent. Such perceptions, as stated by Dipple-Johnstone in 2018, has led to misplaced trust in AI, exposing people to

manipulation, privacy violation and loss of autonomy. In terms of deployment, exaggerated fears may lead to the abandonment of useful technologies such as those that aid medical diagnoses. Therefore, it is important to regulate for proper use of AI to develop applications with understanding of public concern (Fast & Horvitz, 2017).

When reviewing the literature, Cave, et al. discuss the perceptions towards government and researchers, and the awareness of the public to the risks and benefits of AI, as well as their opinion on the future shaping of the AI industry. This range of sources allows a more accurate representation of people from a range of socioeconomic backgrounds, although the paper tends to lend focus to government and researchers. Despite this, the author identifies a gap with the public perception, stating that although there is research available, more could be done to build upon it, particularly to do with the ethics surrounding AI, which has been identified as an important issue (Johnson & Verdicchio, 2017).

The hypothesis of the paper can be theorised as "To what extent do the UK public subscribe to utopian and dystopian narratives of AI?". To investigate this, the researchers approached an online market research panel, consisting of over 20,000 members, adhering to the random sampling method as the identity of the respondents was unknown. The personal data collected by the research panel allow for distinguishing between variances in population, however the risk is that you are limited to the people that have signed up meaning there is a possibility that some population demographics are not proportionally represented. To overcome this, the responses were weighted to provide a fairer picture, cut by group, gender and sociodemographic status. The overall sample size was 1078 which when calculated, is too small to give an accurate representation of the intended population, even with an optimal confidence interval, however due to the type of audience approached, a reasonable sample size

although theorised, could not be guaranteed. The survey produced consisted of questions presented in a set order with answers in a randomized order to minimize the possibility of people being influenced. The survey collected information about the subject's experiences with technology, along with questions to do with AI which queried current knowledge and evaluated claims made about AI which had been drawn from hopes and fears that underlie imaginative thinking about intelligent machines, with the latter employing a common Likert style measurement.

Results from the survey show that there is an overall negative view towards AI as subjects were identifying downsides from the utopian narratives as well as the dystopian ones. However, the interpretation of these results builds on research that has already been done, but do not offer proper examination of whether people's perceptions are influenced by narratives as the researchers do not show testing without the inclusion of narratives to uncover general perceptions. This is highlighted by suggestions of further work as well as examining the sources of narratives, studying the impact of alternative narratives and examining perception in other cultures. Overall, the paper concludes that based on their research, negotiating the deployment of new AI technologies requires contending with parts of the world where much of the population focus on the negative side-effects as appose to the benefits they could offer.

2.2. AI Presentation in the Media

Another publication by Fast and Horvitz analyses how public perception of AI has changed over time, focusing on views expressed in the New York Times over a 30-year period (2017). This is an exception to common research, as stated by Chuan, et al., mentioning that despite numerous speculations about AI, little empirical data exists regarding how AI is covered in the media (2019). The study outlines multiple variables,

including levels of engagement and optimism, and the hopes of concerns surrounding AI. These variables highlight the significance of the problem, stated as "If public expectations diverge too far from what is possible, we may court the smashed hopes that often follow from intense enthusiasm and high expectation". As well as being a unique take on the research, it has also been recognised as an important angle from which to research, to understand social obstacles regarding AI (Liliequist, 2018). These social obstacles are fuelled by the possibility of soon being capable of building dangerous AI systems, despite media coverage employing a more optimistic view (Obozintsev, 2018).

The study highlights other research into the impact of AI on society and different cultural perspectives from text corpora over long time periods, placed in a theoretical framework. From this research, the study identifies a gap which relates to whether optimistic or pessimistic viewpoints are more common in public imagination and how these visions have evolved over time. From this gap, hypothesis can be assumed that as a technology has more media presence, the views should become more optimistic. However, it can also be hypothesised that as more technology is critiqued, levels of pessimism could increase, bringing concerns such as loss of control to the forefront. This view correlates with those mentioned by Cyrus, et al, that levels of pessimism and optimism remain balanced, despite the increase in discussion (2018). The increase in discussion is also noticed by Cave and Heigeartaigh when talking about responsible development of AI and public perception, stating that "AI is the focus for a growing range of public concerns, as well as an increase in levels of optimism" (2018). Fast and Horvitz mention that news and social media can offer a powerful reflection of public attitudes and how they change over time. This point was highlighted by Manikonda and Kambhampati, who conducted an investigation on how the public perceives the

progress of AI by utilizing data shared on Twitter. Their results correlated with the study, concluding that general tweeters talking about AI appear to be more positive towards AI and its related topics (2017).

From their research, Fast and Horvitz outline four research questions covering topics such as the prominence of public discussion of AI, levels of optimism and pessimism over time as well as the associated ideas, and which concerns have diminished through time. The sample consisted of articles from the New York Times from January 1986 to May 2016, totalling more than 3 million articles. Using the 'judgement' sampling method, the sample is narrowed down to only include articles within the scope of analysis. Each article is then scraped using the 'BeautifulSoup' python package. It is evident that the historical research design method has been used as the data collected is from articles dating as far back as 30 years.

Analysis of the articles revealed that one topic mentioned towards the latter end of the timeline is that of concerns towards the permanent displacement of human workers, as discussed by Tussyadiah and Miller when stating that public discourse about the future of AI holds two opposing visions, also mentioning that discussions of fear and loss of control of Ai have persisted (2019). Tussyadiah and Miller also lend their conclusion towards explaining the two opposing views of AI impact on society by uncovering three underlying factors (benefitial, destructive, and risky AI). The views to negative impacts on workers as well as loss of control and ethical concerns are highlighted by Wirtz, et al. when discussing social acceptance and trust towards AI, stating that although social debate about the impact of AI on society has become more intense and tended to focus on the positives, these concerns have increased (2019).

2.3. Concern towards AI Systems

As mentioned previously, Fast and Horvitz highlights other research into the impact of AI on society. One such paper takes the viewpoint of the "Rise of Concerns about AI" (Dietterich & Horvitz, 2015). Generally, the paper summarises a vast improvement in scalability of AI techniques which has resulted in a wide-scale adoption of autonomous systems (Jha, et al., 2018). Outside of computer science, researchers have expressed concerns that AI systems could threaten the survival of humanity. One cause of this is technophobia which is particularly relevant when explaining people's apprehension towards the implementation of self-driving cars (Tussyadiah, et al., 2017; Tussyadiah, et al., 2018). Due to this, the paper suggests that investigation must continue and address concerns, regardless of their scale or distance in the future. This is important as consumers might resist a novel technology, leading to a potential avoidance of technological innovation (Hegner, et al., 2019).

The article identifies five classes of risk that are becoming salient as society comes to rely on autonomous and semi-autonomous systems to make high-stakes decisions. This is highlighted by Atkinson who talks about the AI community actively working to reduce these risks as they develop technology further (2016). These risks address concerns and threats of researchers by categorising them as "bugs", "cybersecurity", "Sorcerer's Apprentice", "shared autonomy" and "socioeconomic impacts". It is often considered that the reason people are hesitant about adopting AI systems, is because they are perceived as being too smart. However, Bundy states that this paradox can be resolved by observing the threats being caused by AI being too dumb, instead of too smart (2017). One category of risk that is mentioned relates to the threat of cyberattacks. This type of risk has been brought on by a societal-scale, grown by AI systems (2017). Cyberattacks are a major risk as criminals and adversaries are constantly attacking

computer systems with viruses and other forms of malware. Consequently, AI algorithms are just as vulnerable as any other form of software. Thus, consideration must be made for the new attack surfaces that these systems can expose in the near-future (Cave & Heigeartaigh, 2019). The overall discussion of AI and its social consequences by this paper as well as others has emphasized the progress made by AI research so far, and the categorization of risks has allowed views of additional progress that can be anticipated in the future (Brundage, 2016).

2.4. AI Discourse

Artificial Intelligence discourse is a topic debated by Johnson and Verdicchio in their study that looked to use a sociotechnical frame to examine discourse about futuristic AI (2017). At the top of their discussion they mention that the main ethical issue facing the AI research community is how AI research and product are responsibly conceptualized and presented to the public. The way the public engage with technological advances plays a crucial role in the acceptance of AI technologies, as the media fosters the fear of misunderstanding and overinflates the expectations of robotics and AI (Stokvis, 2018).

The literature discussed covers concepts such as autonomy, future of AI and the AI artefacts of today. The term ‘autonomy’ is used as a metaphor that refers to different types of computational behaviour, however a multiplicity of meanings for autonomy can lead to miscommunication, as mentioned by Johnson and Verdicchio who use this argument to discuss the potential of AI and the concept of triadic agency (2017). The study uses this literature as a foundation to their argument for a reframing of AI discourse that avoids the pitfalls of confusion about autonomy. Their proposal requires AI discourse to recognize two distinct entities that are computational artefacts and AI

systems. The paper defines a computational artefact as an artefact whose operation is based on computation. In terms of autonomy, by delegating the execution of operations needed to reach specific goals, human effort is freed of the burden. Furthermore, the paper defines an AI system as one that consists of a computational artefact together with human behaviour to make the artefact a useful and meaningful entity. Overall, their discussion concludes with the claim that AI research and researchers will be better served and will provide better understanding of AI to the public by framing the discourse in this way, as the concerns that people have about autonomous systems have a lot to do with human actors in AI systems rather than solely on the computational artefacts contained within them.

Another study by Omohundro looks to use formal methods to create provably safe but limited autonomous systems. Their research identifies that military and economic pressures are driving the rapid development of autonomous systems. They mention that systems are designed to approximate rational economic agents and rational systems exhibit universal drives towards self-preservation, replication resource acquisition and efficiency. It is argued that these drives will lead to dangerous behaviour if not countered and the current computing environment would be very vulnerable to this kind of system.

The discussion of autonomous systems begins with the fact that autonomy is an imminent technology. The risk of the pressure discussed above is that these systems will be rushed and released incomplete, resulting in potentially dangerous consequences. There is discussion for the development of autonomous systems to be rationalised, meaning that unnecessary complexity will be removed meaning more focus will be applied to the aspects of these systems that are crucial to proper execution. The rationalisation of systems exhibits universal drives, as discussed previously, which

concurs with the assumption that AI will decide to turn against its 'masters'. Prediger states that although these systems are designed to serve the goals encoded in them, it is extremely likely for AI agents to converge to these drives, should they be striving to maximize an objective (2017). Furthermore, to develop an intuition about the drives of rational systems, it is useful to consider a simple autonomous system with a concrete goal. The universal drives described will influence the actions of the AI and may result in behaviour not anticipated by the designers (Ebhardt, 2018).

The solution presented by the study, referred to as "The safe-AI scaffolding strategy", involves the safe development of autonomous systems that are used in the construction of more powerful and less limited successor systems. This strategy calls for the development of powerful intelligent systems that are highly constrained. Pohl states that next generation systems will be built on the previous generation with the proof of safety required for each subsequent generation (2015).

2.5. Ethics of Responsibility

As autonomous vehicles become more commonplace, the need for ethical responsibility to be placed on an involved party, becomes crucial. Hevelke and Nida-Rümelin discuss that the assignment of blame poses both legal and moral questions due to the variance of parties involved (2015). The main parties discussed are the manufacturers and the users. More recently, the concept of responsibility and moral phenomenology of using self-driving cars has been studied by Coeckelbergh who state that discussions concerning ethics of self-driving cars should not be restricted to either general responsibilities related to the use of self-driving cars or questions regarding the behaviour of the car in response to objective features of abstract traffic situations (2016). For the purpose of the study, the assumption is made that it is possible to design

autonomous vehicles which cause fewer and less severe accidents than cars steered by the average driver. This assumption is made because if autonomous cars prove less safe than human-operated vehicles, the question asked within the paper are rendered moot.

Before responsibility can be reasonably assigned, the courts must determine the circumstances in which each party can be at fault (Borenstein, et al., 2019). The paper states that holding the manufacturer responsible would be the most obvious solution as any flaw in the system that could cause an accident should be known about as otherwise they would be selling a defective product. There is also the question as to whether the development of autonomous cars should be promoted in the first place. This risk of designing the liability to promote development is that high liability risks may weaken the incentive of manufacturers to innovate, which could slow improvements to safety (Taeihagh & Lim, 2018). The paper states that "it is a fallacy to take the real consequences of a decision into account when confronted with probabilistic phenomena". Instead weight should be lent to the probability of the consequence of a decision as appose to the actual consequence itself. On the other hand, JafariNiami argues that self-driving cars will disproportionately affect the population such as those of lower socioeconomic status (2018).

There is reference to concern that any security gain derived from the introduction of autonomous cars would constitute a trade-off in human lives. Ethical issues of groups who might be more at risk mean that some pre-programmed reactions in case of accidents might use some groups as means to an end. This point is raised by Coca-Vila who states that these programmes may result in the inevitability of the breaking of regulations (2018). If autonomous vehicle development is to be promoted, it is important to design car manufacturer tort liability in such a way that this is helped along. However, the legal liability for autonomous vehicles is difficult to govern as

there are very few legal documents for autonomous vehicles and even fewer that discuss the concept of liability in traffic accidents (Li, et al., 2018). Although there is very little in the way of formal liability documents that can be referenced should a case surface, the paper states that it is a form of defamation if a person is held to blame if they never had a chance to intervene. This is highlighted by Gogoll and Uhl, who mention that as technology progresses, situations arise where the division of work and supervision during a task is neither needed nor wanted (2018).

When considering assigning blame to the user of the autonomous vehicle, the paper explores whether it would be morally permissible to impose liability on the user based on duty to pay attention to the road and traffic and to intervene when necessary to avoid accidents. Horl, et al. build on this by adding that this is a cause for question as both the driver and third persons outside of the car are a severe risk (2016). Another alternative design discussed is an approach where the person in charge of the autonomous vehicle has no duty of interfering, however is considered morally responsible for possible accidents. The risks associated with this method, from a financial perspective, are that as the assignment of responsibility becomes fuzzier, there will be a likely increase in insurance premiums (Thomopoulos & Givoni, 2015), meaning that people will have to pay more money to insure their vehicles (autonomous or manual).

Throughout the paper, the author highlights many arguments both for and against autonomous vehicles, although lacks equal focus with quantitative empirical research (Karnouskos, 2018). However, this issue is commonplace amongst similar research, and is identified as a gap for further research to take place. Although more philosophical discussions of responsibility have taken place, the research into ethics of autonomous vehicles must make headway in the near future before the technology can be understood

to the point where it can be fully implemented (Kumfer, et al., 2016). Although the paper concludes that current regulatory systems can be suitably modified for autonomous vehicles, London and Danks argue that expanding and generalising the scope of existing regulatory agencies is decidedly unsuited to the task of regulating autonomous vehicles and so propose a novel regulatory system, modelled on those for regulating medical interventions (2018).

3. Methodology

3.1. Research Design

The research conducted in this study is best suited to descriptive research methods. This type of study is concerned with the functional relationships between variables, hypothesis, and the development of generalizations across populations (Salkind & Rasmussen, 2007). Descriptive research involves identification of attributes of a particular phenomenon based on observations or exploration of correlations between two or more phenomena (Williams, 2007).

The benefits of conducting a study under this design method is that any subject involved is being observed in a completely natural environment. Descriptive research is often used as a precursor to more quantitative research designs with the overview giving valuable pointers as to variables that are worth testing. Also, providing the limitations of this research method are understood, descriptive methods can be useful tools in developing a more focused study.

However, the limitations of using descriptive research methods are that the results obtained cannot be used to discover a definitive answer or disprove a hypothesis, as well as being unable to replicate these results due to the use of observational methods. Also, the descriptive function is heavily dependent on instrumentation such as interviews and surveys, for measurement and observation.

In summary, a descriptive research design method is valid for researching specific subjects and as a precursor to more quantitative studies. Whilst there are some concerns

to the validity, if the limitations are understood by those involved in the study, this method can be an asset to the project (Shuttleworth, 2008).

3.2. Online Survey

An online survey was released to gather information about peoples understanding of a collection of clever systems that are present in vehicles. Although discussed in a previous chapter that both science-fiction and science-fact would be tested, a decision was made to include only systems that are currently present. It is also important to note that due to the abstract nature of the definition of 'artificial intelligence', the use of a range of clever systems allows for a broad spectrum of understandings to be adhered to.

The benefits of producing a survey and publishing it online are that there is an immediate increase of population access (Schmidt, 1997). It is also quicker and easier for participants to complete the survey as they can access it from their mobile phone at any time, providing they have access to the internet.

The survey was produced using 'Google Forms' as this is automatically accessible to Keele students through their University accounts. As well as providing a slick and easy platform to create surveys with multiple input types, Google Forms also includes automatic rendering of charts and graphs for quantitative data, removing the need to create them personally, avoiding the risk of human error when interpreting the data.

3.2.1. Sample

The survey was released to students at Keele University, through social media (see Appendix 1). From an ethical perspective, choosing students from Keele was the easiest solution as it required the least amount of ethical clearance. The sample is generalised over the UK population of 66.85m. Using a sample size calculation with an

optimum confidence interval of 1.359, the estimated required sample size is around 5200. The current size of the sample approached is 5217 (as of 26th April 2019), therefore is enough for this generalization. However, the number of respondents was only 27. Therefore, the results for this survey will be used to show proof of concept, and not as definitive evidence to prove hypotheses.

3.2.2. Content

The survey (see Appendix 2) begins by asking the respondents whether they are aware of the concept of artificial intelligence, along with an optional field for them to give a brief description of what they think artificial intelligence is, to gauge their understanding. Although this section is not technically related to the purpose of the study, it provides another angle to the study which explores whether the level of trust is affected depending on whether the subject is aware of artificial intelligence.

For each of the systems mentioned in the survey, the respondent is given the name, along with a brief description of the system, including any other names that the system may be known as. The survey asks whether they are aware of the system, offering 'yes', 'no' and 'not sure' as the options. They are then asked whether they would trust the system if it was present in their car. It is important to note that as not all respondents will hold a driving license, some results will be from the perspective of a passenger. The question is answered using a common 5-point Likert-style scale ranging from 'strongly agree' to 'strongly disagree'. Finally, the respondent is asked where they first became aware of the system, with the options being 'science-fiction', 'science-fact' and 'first-hand experience'. There is an option for 'other' responses, however it is anticipated that answers that would fit the other categories will be present in this section due to their perceived ambiguity.

3.2.3. Ethical Considerations

At the start of the survey, a 'terms and conditions' page (see Appendix 3) is displayed to the respondent which discusses the purpose of the research as well as the potential risks of taking part. It is stated that the respondent is free to opt-out of the survey at any point should this be necessary. Finally, the respondent is informed about how their information is stored and used upon completion, where they are told that their response is completely anonymous as no personal information is collected by the researcher. To ensure the 'terms and conditions' are read and understood, a checkbox asking whether participation is agreed, is displayed at the bottom of the page and is a required field before the next page of the survey can be accessed.

In terms of the contents of the survey, advice from the ethics committee at Keele suggested that changing the word 'fear' to 'concern' would further minimise the chance of people having negative recalls whilst answering the questions.

3.3. Face-to-Face Interview

After the online survey had gathered sufficient data, participants were invited to take part in a face-to-face interview to test whether exposure to a real-world example of a clever system in action, would change their initial perception in terms of how much they would trust the system if it were present in a vehicle they were in.

Interviews are a qualitative research method (Dudovskiy, 2017) with the format of the interview employing that of a 'semi-structured' nature. The questions will prompt an answer in the form of a 7-point Likert Scale (Likert, 1932). This will produce a quantitative result for analysis, as well as comments which can be references when explaining individual results.

3.3.1. Sample

The interview was intended for a wide range of subjects with all different levels of knowledge and age groups. Unfortunately, due to time constraints, only 5 subjects were able to complete the interview in its entirety. Although verbal agreements had been made with a larger number of subjects, should these interviews have taken place, insufficient time would have remained for analysis of the results.

However, the subjects that were interviewed did have different levels of knowledge of AI and the systems they were presented with, as well as a substantial age range, which allowed for a different angle to be taken on the results which explored the levels of trust in clever systems of a generation who are naturally perceived to be less open to adopting new technologies. Nonetheless, due to the limited number of results subjects interviewed, the results are to be considered as proof of concept of the hypothesis instead of concrete evidence of its viability.

3.3.2. Content

The structure of the questions for each system remained the same to facilitate consistency and unbiased responses (see Appendix 4). Subjects were first asked of their awareness of the system in the form of a short description of what they understood about the system. In the event that the subject did not know what the system was or was unable to provide sufficient explanation, a brief definition was provided by the interviewer.

Subjects were then asked to judge whether they would trust the system if it were present in a vehicle they were travelling in. This question would give an initial trust level for each subject with newer, more advanced systems expected to achieve lower results before viewing the video.

During the video viewing, the interviewer would watch for social cues from the subject to see whether they are visibly appearing to gain trust in the system. The videos chosen for viewing are of real experiments involving the systems in question in order to provide the best proof, apart from physically experiencing the system first-hand. After the video viewing, subjects were asked the same question regarding their trust in the system to see whether they would be any change.

3.3.3. Ethical Considerations

Before the interview took place, each subject was given an information sheet (see Appendix 5) to read which included information about why the research was taking place, what would happen if they decided to take part and the risks that could arise should they take part. This consent from was completed with support from the Keele Ethics Committee to ensure that all eventualities were covered.

After the information sheet was fully understood, each subject was asked to fill out and sign a consent form (see Appendix 6) to formalise their agreement to participate in the interview. The consent from included agreements such as their understanding of the voluntary nature of the interview, data collection and usage terms and the agreement that the interview would be audio recorded to facilitate more detailed analysis of results.

4. Results

Upon completion of the online survey and the interview process, the results were collected and analysed. There are multiple ways that the results from the online survey can be analysed, some of which are not included in this report if deemed not suitable or out of scope of the purpose of the study. Audio recordings of the interviews were captured and used as reference where applicable to aid the understanding of the researcher to an answer given by a participant.

4.1. Online Survey

In total, 27 responses were collected from the sample pool the survey was released to. Although the intended sample adheres to the minimum required sample, the actual sample size does not provide a reasonable presentation of the views of the generalised population. Therefore, for this study, the result analysis will be used as a proof of concept for the hypotheses.

4.1.1. Awareness of Artificial Intelligence

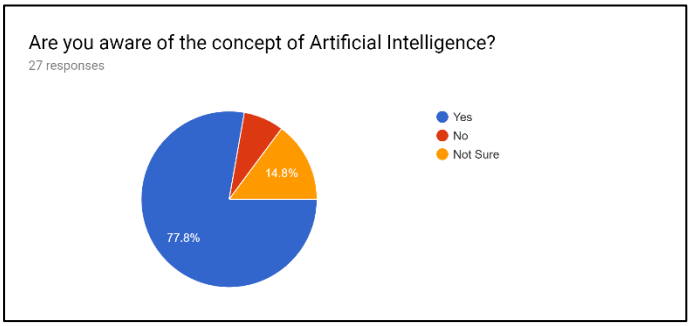


Figure 1: Online Survey - Awareness of AI

The first section was produced to gain understanding of the awareness of artificial intelligence by the participants. Due to the abstract nature of the definition of artificial intelligence, a definition was not provided. In total,

77.8% of participants were aware of AI, with 14.8% saying they were unsure and 7.4% saying they were not aware. The results of this question were as expected given the increase in media presence of AI in recent years.

When asked to provide a brief description of their understanding of AI, 55.6% of participants chose to answer. It is the understanding of the researcher that some participants had extensive knowledge of computer science and subsequently of artificial intelligence. This understanding is clear in the responses, for example one participant mentioned that AI could be described as "systems mirroring biological neural networks which are capable of processing data and 'learning' from the data provided".

Other responses gave the impression that limited knowledge of AI was present, such as the response provided by one participant which stated that AI could be "something being operated by computer programming, for example a tram at the airport". This response highlights the ambiguity of the definition of AI as depending on the definition explored, the example system provided could be considered as and AI system, where as other definitions may solicit a different opinion.

4.1.2. Radar-Guided Cruise Control

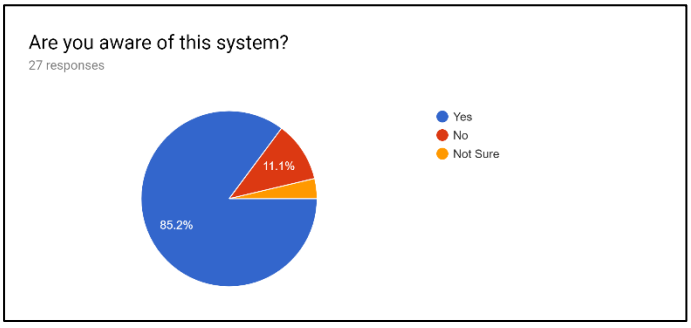


Figure 2: Online Survey - Radar-Guided Cruise Control 1

When asked about their awareness of this system, 85.2% of participants said they were aware of the system, 11.1% were not aware and 3.7% were unsure whether they had heard of the system. It is possible that this result is slightly skewed as some

participants that said they were aware, could have assumed that they were being asked their awareness of ‘cruise control’ instead.

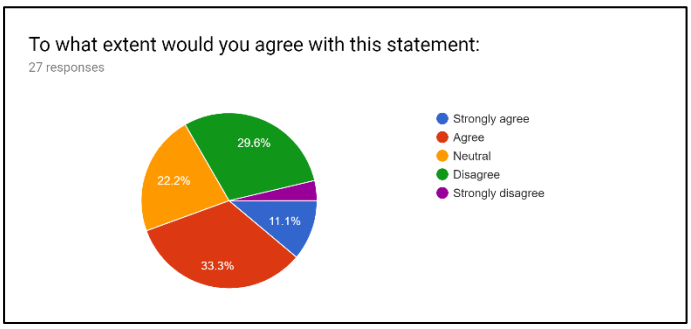


Figure 3: Online Survey - Radar-Guided Cruise Control 2

Participants were then asked to what extent they agree with the statement “I would trust this system if it were present on my car”. For this system, the results were fairly mixed. The most common result at 33.3% is that for ‘agree’, which is the second most positive response, closely followed by ‘disagree’ at 29.6%. With this result, a possible cause for the mixed response could be the understanding of the term ‘radar-guided’, along with its explanation which showed basic levels of autonomy with regards to controlling the vehicle speed depending on the velocity of the car in front.

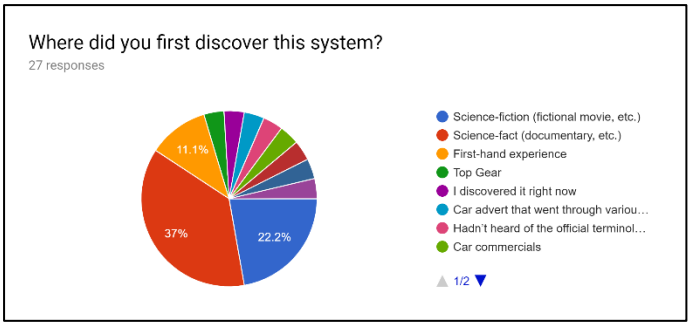


Figure 4: Online Survey - Radar-Guided Cruise Control 3

Finally, when asked about their first discovery of this system, the most common answer was through ‘science-fact’ at 37%, followed by ‘science-fiction’ at 22.2%. There were 8 responses considered as ‘other’ (29.6%), however on review of these responses, it has become apparent that many of them would fit into the pre-provided categories. Nevertheless, due to the absence of the context to which this system was portrayed, for the sake of the study, these will be considered individually, and therefore omitted from the results.

4.1.3. Automated Braking Assistance

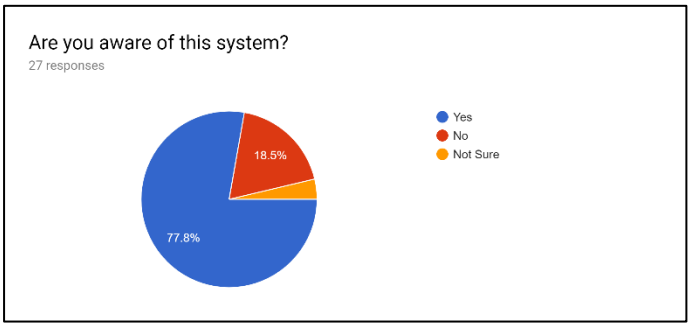


Figure 5: Online Survey - Automated Braking Assistance 1

When asked about their awareness of this system, 77.8% of participants said they were aware of the system, 18.5% were not aware and 3.7% were unsure whether they had heard of the system. This result was as expected as it is a system that has been around for a long time and is becoming more commonplace in cars nowadays.

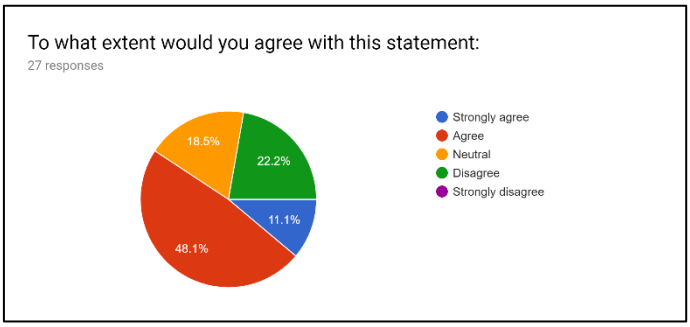


Figure 6: Online Survey - Automated Braking Assistance 2

Participants were then asked to what extent they agree with the statement "I would trust this system if it were present on my car". For this system, the results were overall quite positive with 59.2% giving a positive response and only 22.2% of participants offering a negative response. This result is as expected as this system was described as a system that is there as a 'last resort' and it should not be a system that is relied upon to stop the vehicle in a necessary situation.

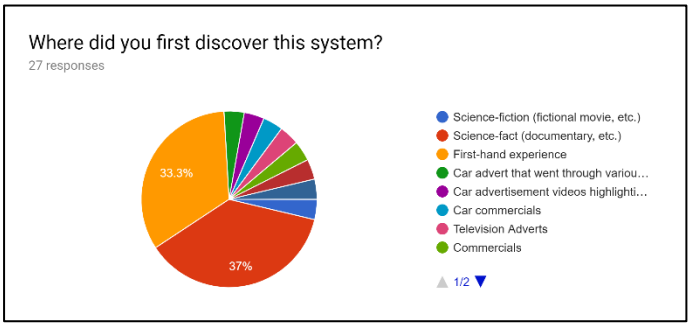


Figure 7: Online Survey - Automated Braking Assistance 3

Finally, when asked about their first discovery of this system, the most majority of responses either real-life experience or science-fact with a combined 70.3% of participants offering

these responses. As with the previous question, answers given under the category of 'other' can be considered as related to other categories, however due to not knowing the context to which this system was portrayed, it is not possible to reasonably place these answers in other categories without asking the participants individually as to their reasoning.

4.1.4. Collision Avoidance System

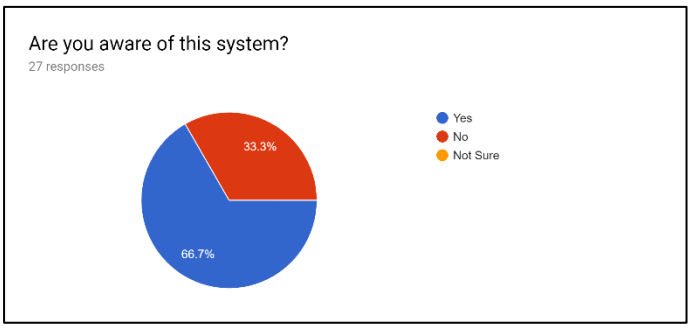


Figure 8: Online Survey - Collision Avoidance System 1

When asked about their awareness of this system, 66.7% of participants said they were aware of the system and 33.3% said they were not aware. This result is slightly different to what was expected as collision avoidance is a relatively new system and one that can be considered as an upgrade on automatic braking systems.

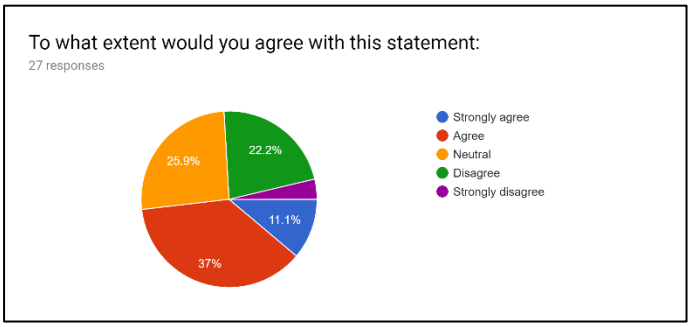


Figure 9: Online Survey - Collision Avoidance System 2

Participants were then asked to what extent they agree with the statement "I would trust this system if it were present on my car". For this system, the results were quite mixed with only 48.1% giving a positive response. The variance in these results could be down to a large portion of the participants not being aware of the system prior to taking the survey.

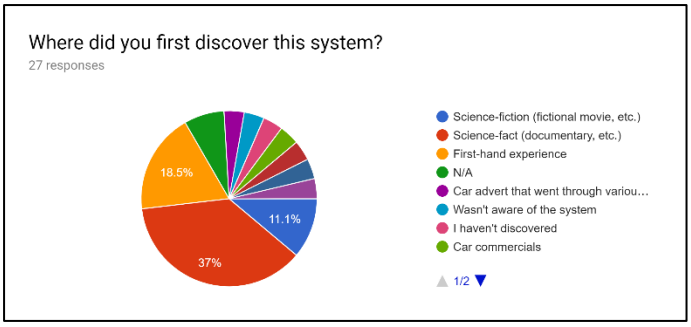


Figure 10: Online Survey - Collision Avoidance System 3

Finally, when asked about their first discovery of this system, 37% recalled hearing about this system in science-fact, with 18.5% saying they had first-hand experience of the system. Only 11.1% say they first heard about the system in science-fiction. Other answers included recalling the system being explained in car commercials which, although can be considered as science-fact, if the system is displayed as an animation instead of a real-life example, the is ambiguity as to the categorisation of this as science-fact.

4.1.5. Automated Parking

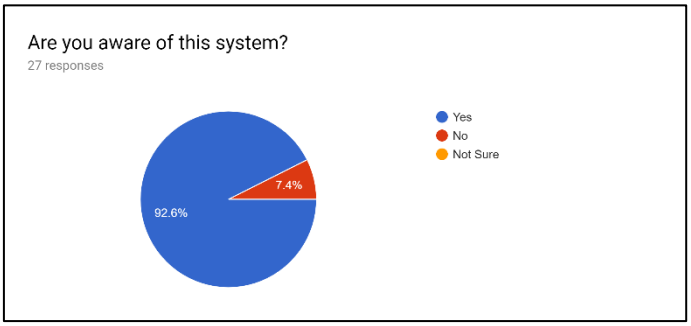


Figure 11: Online Survey - Automated Parking 1

When asked about their awareness of the system, 92.6% said they were aware of the system with only 7.4% saying that they hadn't heard of it. This makes it the most well-known of all the systems asked about in this survey which is somewhat as expected, due to the system being around in different forms for many years.

Participants were then asked whether they agree with the statement "I would trust this system if it were present on my car". This system provided the most positive response of all with 74.1% of participants saying that they would either 'agree' or 'strongly agree' with the statement. The reason for this result can be put down to the speed at

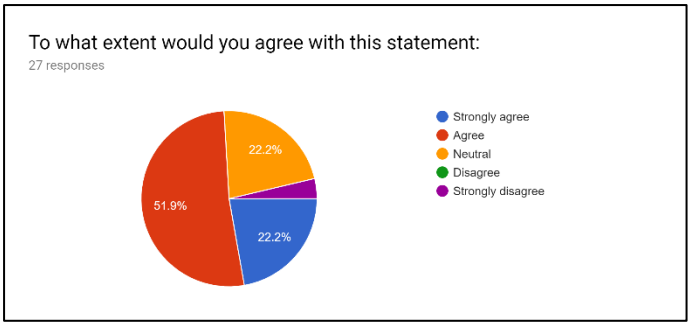


Figure 12: Online Survey - Automated Parking 2

which this system would operate at, as well as the known ability to override and take manual control of the system, should the driver of the vehicle wish to do so.

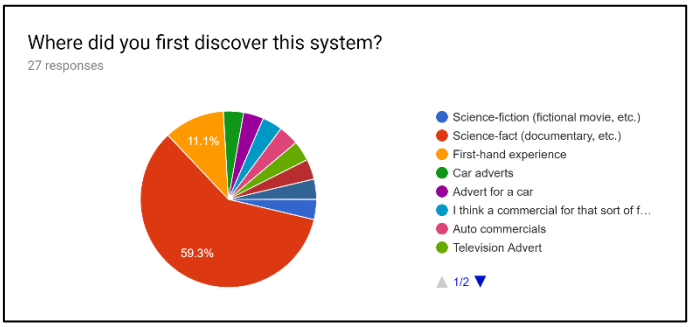


Figure 13: Online Survey - Automated Parking 3

Finally, when asked about their first discovery of the system, 59.3% recalled viewing this system in science-fact. As with the previous system, the majority of the 'other' answers refer to commercials that they have seen either on television or online. As stated, without being aware as to the context of the advert, it is unreasonable to place these answers within the 'science-fact' category.

4.1.6. Self-Driving

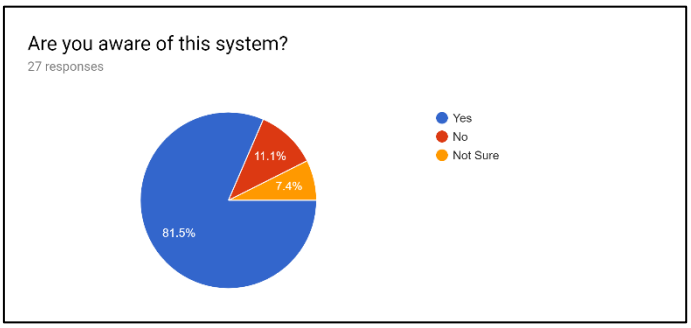


Figure 14: Online Survey - Self-Driving 1

When asked about their awareness of the system, 81.5% said they were aware and 11.1% said they weren't. This result is as was expected due to the media presence both fictional and non-fictional, being arguably the largest of all the systems discussed in this study. 7.4% of participants were unsure whether they had heard of the system, however this could be down to the number of different terms used to describe the

concept of 'self-driving' and depending on the sources to which the participant has been exposed to this system.

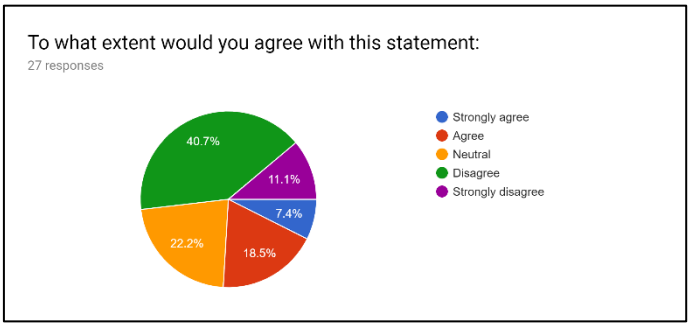


Figure 15: Online Survey - Self-Driving 2

Participants were then asked whether they agree with the statement "I would trust this system if it were present on my car". The response to this question was the most negative of all systems presented with 51.8% of participants offering a negative response. Positive responses only accumulated to 25.9% of the overall sample. This result was to be expected due to ideas discussed earlier in this paper about 'loss of control'.

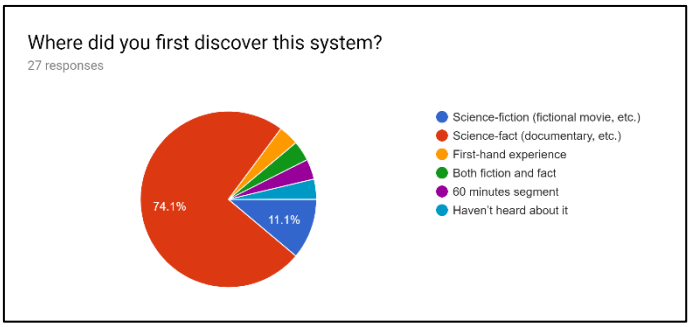


Figure 16: Online Survey - Self-Driving 3

Finally, participants were asked about their first discovery of the system, with 74.1% of participants recalling from science-fact. 11.1% of participants recalled seeing this system in science-fiction. This result was not what was expected as self-driving technology, although coming to light in vehicles produced at present (e.g. Tesla Autopilot), there is still argument to suggest that the idea of self-driving vehicles in their true form is in the realms of science-fiction.

4.2. Face-to-Face Interview

The semi-structured interview was designed to see whether exposure to a real-life example of a clever automotive system in action will change how someone

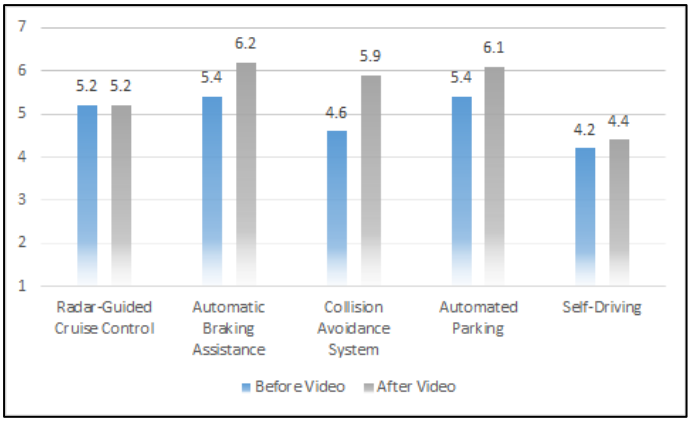


Figure 17: Interview - Trust Measurement

perceives it in terms of whether they would trust it (see figure 17). Due to time constraints, only 5 subjects were acquired, therefore the results of the interviews will show proof of concept of the hypothesis, as appose to providing definitive proof for or against it.

4.2.1. Radar-Guided Cruise Control

Initially, the trust level was relatively high with an average response of 5.2. The lowest response given was 4, however when asked what their understanding of the system was, they were unable to give a clear answer and an explanation had to be given to them by the interviewer.

After watching the video showing how the system worked, the average trust score remained the same. However, one subject 4 explained that they were slightly more trusting of it having seen it in action, with their score going up from 4 to 5. On the other hand, subject 5 was slightly less trusting having watched the video with their score going down from 5 to 4. After being asked to explain why they had reduced their score they summarised that it appeared that the driver of the vehicle in the video was hesitant as they were hovering over the controls, in case the system failed.

4.2.2. Automated Braking Assistance

Initially, the trust level was relatively high with an average response of 5.4. The lowest response given was 3 with the subject stating that they are used to having control of braking so would be nervous about leaving this to the system. One point that it

important to mention here, pointed out by subject 1, is that they put their trust at a 7 “assuming the breaks are up to standard with the system”. As this is a point regarding the mechanical performance as appose to the performance of the system, the subject omitted this from their thought process when giving a trust score.

After watching the video showing how the system worked, the average trust score raised to 6.2. The biggest increase came from subject 3 who stated that after seeing the system working, it was clear that the vehicle was able to sense obstacles and react quickly, perhaps better than a human would.

4.2.3. Collision Avoidance System

Initially, the trust level was quite low with an average response of 4.6. The lowest score of 4 was given by subjects 2 and 3. The reasoning for giving this score was based on lack of knowledge (subject 2) and the fact that it is a newer technology (subject 3).

After watching the video showing how the system worked, the average trust score raised to 5.9 which is the largest increase of any system tested. The largest increase came from subject 4 who went from a trust score of 5, to a score of 7 after watching the video. The subject did not give any verbal response as to why they increased their score, however from their tone of voice when giving there answer, as well as their reactions throughout the video, the visual cues seemed to suggest a more open attitude to trusting this system.

4.2.4. Automated Parking

Initially, the trust level was relatively high with an average response of 5.4. The lowest score of 3 was given by subject 4. Their reasoning for this score was because

although the speed of the manoeuvre would be quite low, the car would be in complete control.

After watching the video showing how the system worked, the average trust score raised to 6.1. The biggest increase came from subject 4 who added that whereas other systems have a better reaction time than a human, when parking this is not as crucial and is something that a human is perfectly capable of doing.

4.2.5. Self-Driving

Initially, the trust level was the lowest of any of the systems with an average response of 4.2. The lowest score of 2 was given by subject 4 stating that they were not too trusting of the system as it is relatively new and somewhat untested. Similarly, subject 1, who gave a score of 5 stated that the reason they would not give a higher score is because 'automated driving' involves multiple clever systems working at once.

After watching the video showing how the system worked, the average trust score raised to 4.4 which was the smallest increase of all the systems tested. The most significant increase came from subject 2 who increased from a score of 5 to 6. Their reasoning was that from the video, it was reassuring being able to see what the car is monitoring and how it is recognising different objects that could be a risk.

Interestingly, subject 1 dropped their score from 5 to 3.5. When asked their reasoning for this, they stated that the vehicle appeared to hesitate and was unsure as to the lane it needed to be in. They also mentioned that the vehicle stopped lot sooner at a junction that perhaps a human would which, although not a problem for the driver, may cause issues for a vehicle following who may not be aware that the vehicle is being piloted by an 'autopilot' system and would not be able to predict the earlier breaking of the vehicle, subsequently risking a rear-end incident. Finally, subject 1 mentioned that

particularly in America where they have a rule that allows a car to turn right on a red light, even if this is programmed into the system, there may be some situations where this is prohibited, and a sign is in place to notify the driver of this. However, if this sign is out of view of the cameras being used by the system to monitor its surroundings, the system may miss this and assume it can proceed which could cause a collision.

5. Discussion

Once the results had been collected and analysed, the hypotheses could be discussed and answered. As well as hypotheses outlined at the beginning of the study, other useful analyses were discovered.

5.1. Science-Fact vs Science-Fiction

	Science-Fact	***Science-Fiction***
Radar-Guided Cruise Control	*3.5*	*2.3*
Automated Braking Assistance	*3.3*	*2*
Collision Avoidance System	*3.7*	*2.7*
Automated Parking	*4.1*	*3*
Self-Driving	*2.6*	*2*

Figure 18: Online Survey - Trust Comparison 1

An interesting angle to consider analysing the results from the survey is to consider the difference in trust level depending on whether the participant had discovered the system through science-fact or science-fiction. As can be seen in figure 18, across all systems, the trust level was higher with science-fact that it was with science-fiction. This was to be expected as it is reasonable to assume that exposure to a real-life example would envelop more trust that of a science-fiction example.

5.2. Awareness vs Trust

	Aware	***Not Aware***
Radar-Guided Cruise Control	*3.3*	*2.3*
Automated Braking Assistance	*3.5*	*3.2*
Collision Avoidance System	*3.6*	*2.7*
Automated Parking	*3.8*	*3.5*
Self-Driving	*2.6*	*3.3*

Figure 19: Online Survey - Trust Comparison 2

Another interesting angle to consider analysing the results from the survey is to consider the difference in trust level depending on whether

the participant is aware of the system prior to taking the survey. As can be seen in figure 19, the trust level was higher with all systems when the participant was aware, apart from 'self-driving'. This could be due to this system having arguably the largest media presence of all the systems with particular reference to the 'Autopilot' system being developed by Tesla.

5.3. Average Increase

	Average Increase
Radar-Guided Cruise Control	*0*
Automated Braking Assistance	*0.8*
Collision Avoidance System	*1.3*
Automated Parking	*0.7*
Self-Driving	*0.2*

Figure 20: Interview - Average Increase

An area that has already been discussed to an extent is that of the average increase (see figure 20) in trust of each system after watching the video. For radar-guided cruise control, the video appeared to have no effect, although with more subjects, this may heed a different outcome.

The self-driving video example produced a small increase of 0.2. Although this shows that the video had a minute effect on the trust of the subjects, it is the opinion of the researcher that there are general reservations about the performance of self-driving cars that will only be rectified over time as the system develops and becomes more reliable.

The automated parking video produced an increase of 0.7. This system had the highest initial trust score, most likely due to the speed at which the manoeuvre takes place, is very low meaning that there is very little risk involved, as well as more time for the driver to react, should they feel the system is not performing as it should.

Similarly, the automated braking assistance video produced an increase of 0.8. This system had the equal highest initial trust score, most likely due to the system being portrayed as an assistance to the driver, rather than a system where the driver relies on its performance for safety. The average increase for this system may have been higher,

if the video used to test the system, had the vehicle travelling at a higher speed than 10 mph. However, the results show that exposure to a real example of this system can envelop slightly more trust in someone.

Finally, the video used to show off a collision avoidance system gave the highest increase with an average of 1.3. One factor that may have come into effect here is the awareness result from the online survey which shows that this system had the lowest awareness level of all the systems tested. This system has multiple terms that are used to refer to it, therefore there is the possibility of an element of realisation that the subject is aware of the system, once they are shown it in action.

6. Conclusion

6.1. Meeting the Aim

At the beginning of the project, the aim was set out to evaluate the extent to which overpowering images of automotive AI can be mitigated by emphasising real, current applications, or narratives of control and involvement. Overall, this project shows proof of concept that exposure to real-life examples of clever systems can increase a person's trust in that system. Predominantly, the system at the forefront of the media, self-driving, is one that will benefit from further examples as appose to a single video, however the results from the investigation suggest that this video did increase the trust of the system to a degree.

6.2. Meeting the Objectives

The first objective set out was to conduct a systematic literature review. As shown by Appendix 7, links between research papers were established and noted. There is also evidence of critical analysis of papers through in-depth reading of each paper and how it links to other papers. The feeling is that the conduction of a systematic literature review has been met and has provided an invaluable asset and foundation to the investigation and subsequent completion of the project.

The second objective was to rank a set of concerns based on level of identity by the public through the deployment of a survey. Although not explicitly mentioned in this report, further analysis of the results would allow for a ranking of the concerns. In terms of the reference to Fast and Horvitz who mention that pessimism towards concerns such

as loss of control have increased (2017), the results show that the more control is given to a clever system, the less open to acceptance the public becomes, with 'self-driving' systems being the least accepted system, even after exposure to a real-life example of the system in action.

The next objective concerns the exploration of real-world applications that depict concerns in a more positive light. As concerns have been expressed in the form of clever systems in vehicles, this objective was easier to complete. A search was performed on each system to find a real example of it in action, with results offering animations and tv adverts. However, for the sake of the project, consistency was adopted and the video for each system depicted an unbiased test of each system which, as far as the researcher is aware, is not influenced by the manufacturer or sponsors of the systems.

The final objective was to test whether the concerns that had been raised about the systems throughout the project, could be mitigated through expose to the real-world examples identified by the previous objective. The results show that this hypothesis was met in theory, due to limited sample size. However, as far as proof of concept is concerned, exposure to real-world examples of clever systems that depict various concerns about artificial intelligence, can indeed mitigate the overpowering images present in the media.

6.3. Overcoming Limitations

The limitations highlighted at the start of the project were handled so that they did not have a detrimental effect on the investigation or subsequent results. In terms of being able to source real-life examples due to the concerns being expressed as clever systems depicting various concerns people have with AI, this limitation was eradicated as it was much easier to find examples of actual systems as appose to finding examples

of specific concerns which may only relate to systems that exist in science-fiction. However, the hypothesis that was drawn up based on this limitation that "the effect of fictional narratives will not be as strong as real-world applications", was somewhat answered during the online survey which shows that trust levels were not as high with discovery through fictional narratives.

The second limitation which referenced the ethical concern relating to negative recall of such things as car accidents when presented with certain scenarios, was reduced. The suggestion of a screening question before commencing the online survey, and the inclusion of a consent form for the interview, were both implemented meaning that the participants were aware of the risks before starting, and the researcher was covered, should any of these risks come to light.

6.4. Further Work

Future studies into the perception of automotive AI in the media would strengthen the results of the hypotheses expressed in this paper. One possible angle of research is that of an expanded sample area. The results of this study are generalized over the population of the UK, however due to limited sample size, these results are only suitable as proof of concept. Further work should include a much larger initial sample size to minimize the risk of falling short of optimal target sample to suitably represent the population it is generalized for.

7. References

Atkinson, R., 2016. “It’s going to kill us!” and Other Myths About the Future of Artificial Intelligence. *Information Technology & Innovation Foundation,* pp. 8-11.

Borenstein, J., Herkert, J. & Miller, K., 2019. Self-Driving Cars and Engineering Ethics: The Need for a System Level Analysis. *Sci Eng Ethics,* Volume 25, pp. 383-398.

Brundage, M., 2016. *Modeling Progress in AI.* Phoenix, AAAI.

Bundy, A., 2017. Smart Machines are Not a Threat to Humanity. *Communications of the ACM,* 60(2), pp. 40-42.

Cave, S., Coughlan, K. & Dihal, K., 2019. *"Scary Robots": Examining Public Responses to AI.* Honolulu, Association for the Advancement of Artificial Intelligence.

Cave, S. & Heigeartaigh, S., 2018. *An AI Race for Strategic Advantage: Rhetoric and Risks.* New Orleans, AAAI.

Cave, S. & Heigeartaigh, S., 2019. Briding near- and long-term concerns about AI. *Nature Machine Intelligence,* Volume 1, pp. 5-6.

Chuan, C.-H., Tsai, W.-H. S. & Cho, S. Y., 2019. *Framing Artificial Inteligence in American Newspapers.* Honolulu, AAAI.

Coco-Vila, I., 2018. Self-driving Cars in Dilemmatic Situations: An Approach Based on the Theory of Justification in Criminal Law. *Crim Law and Philos,* 12(1), pp. 59-82.

Coeckelbergh, M., 2016. Responsiblity and the Moral Phenomenology of Using Self-Driving Cars. *Applied Artificial Intelligence,* 30(8), pp. 748-757.

Cyrus, W. N. Y., Justin, A. Y. K., Qing, C. J. & Bryan, L. W. Y., 2018. *Artificial Influence A Study of the Informational Social Influence of Human and AI Experts,* Singapore: Singapore University of Technology and Design.

Dietterich, T. & Horvitz, E., 2015. Rise of Concerns about AI: Relfections and Directions. *Communications of the ASM,* 58(10), pp. 38-40.

Dudovskiy, J., 2017. *Interviews - Research Methodology.* [Online] Available at: https://research-methodology.net/research-methods/qualitative-research/interviews/
[Accessed 28 April 2019].

Ebhardt, J., 2018. The Threat of Artificial Superintelligence. *Exigence,* 2(1), pp. 1-8.

Fast, E. & Horvitz, E., 2017. *Long-Term Trends in the Public Perception of Artificial Intelligence.* San Francisco, Association for the Advancement of Artificial Intelligence.

Gogoll, J. & Uhl, M., 2018. Rage against the machine: Automation in the moral domain. *Journal of Behavioural and Experimental Economics,* Volume 74, pp. 97-103.

Gusikhin, O., Rychtyckyj, N. & Filev, D., 2007. Intelligent systems in the automotive industry: applications and trends. *Knowledge and Information Systems,* 12(2), pp. 147-168.

Hegner, S., Beldad, A. & Brunswick, G., 2019. In Automatic We Trust: Investigating the Impact of Trust, Control, Personality Characteristics, and Extrinsic and Intrinsic Motivations on the Acceptance of Autonomous Vehicles. *International Journal of Human-Computer Interaction.*

Hevelke, A. & Nida-Rümelin, J., 2015. Responsibility for Crashes of Autonomous Vehicles; An Ethical Analysis. *Sci Eng Ethics,* Volume 21, pp. 619-630.

Hörl, S., Ciari, F. & Axhausen, K. W., 2016. *Recent perspectives on the impact of autonomous vehicles.* Zurich: Swiss Federal Institute of Technology Zurich.

JafariNiami, N., 2018. Our Bodies in the Trolley's Path, or Why Self-driving Cars Must *Not* Be Programmed to Kill. *Science, Technology, & Human Values,* 43(2), pp. 302-323.

Jha, S., Raman, V. & Sadigh, D., 2018. Safe Autonomy Under Perception Uncertainty Using Chance-Constrained Temporal Logic. *J Autom Reasoning,* Volume 60, pp. 43-62.

Johnson, D. & Verdicchio, M., 2017. AI, agency and responsibility: the VW fraud case and beyond. *AI & Society.*

Johnson, D. & Verdicchio, M., 2017. Reframing AI Discourse. *Minds & Machines,* Volume 27, pp. 575-590.

Karnouskos, S., 2018. Self-Driving Car Acceptance and the Role of Ethics. *IEEE Transactions on Engineering Management.*.

Kumfer, W., Levulis, S., Olson, M. & Burgess, R., 2016. *A Human Factors Perspective on Ethical Concerns of Vehicle Automation.* Washington, HFES.

Likert, R., 1932. A technique for the measurement of attitudes. *Archives of Psychology,* 22(140).

Liliequist, E., 2018. *Artificial Intelligence - Are there any social obstacles?.* Stockholm: KTH.

Li, S. et al., 2018. Ethical and Legal Dilemma of Autonomous Vehicles: Study on Driving Decision-Making Model under the Emergency Situations of Red Light-Running Behaviors. *Electronics,* 7(10), pp. 264-282.

London, A. & Danks, D., 2018. *Regulating Autonomous Vehicles: A Policy Proposal.* New Orleans, AAAI.

Manikonda, L. & Kambhampati, S., 2017. *Tweeting AI: Perceptions of Lay vs Expert Twitterati.* Arizona: Arizona State University.

Obozintsev, L., 2018. *From Skynet to Siri: An Exploration of the Nature and Effects of Media Coverage of Artificial Intelligence.* Delaware: University of Delaware.

Pohl, J., 2015. *Artificial Intelligence: Extinction or Nirvana?.* Baden-Baden, InterSymp.

Prediger, L., 2017. On the Importance of Monitoring and Directing Progress in AI. *AI Matters,* 3(3), pp. 30-38.

Sadigh, D., 2017. *Safe and Interactive Autonomy: Control, Learning, and Verification.* Berkeley: University of California.

Salkind, N. & Rasmussen, K., 2007. Descriptive Research. In: *Encyclopedia of Measurement and Statistics.* Thousand Oaks: SAGE Publications, pp. 251-254.

Schmidt, W., 1997. World-Wide Web survey research: Benefits, potential problems, and solutions. *Behvaiour Research Methods, Instruments, & Computers,* 29(2), pp. 274-279.

Shuttleworth, M., 2008. *Descriptive Research Design.* [Online] Available at: https://explorable.com/descriptive-research-design [Accessed 28 April 2019].

Stokvis, E.-J., 2018. *How can we control the ethical implications of artificial (super)intelligence?.* [Online] Available at: https://www.linkedin.com/pulse/how-can-we-control-ethical-implications-artificial-ernst-jan-stokvis [Accessed 24 April 2019].

Taeihagh, A. & Lim, H. S. M., 2018. Governing autonomous vehicles: emerging responses for safety, liability, privacy, cybersecurity, and industry risks. *Transport Reviews,* 39(1), pp. 103-128.

Thomopoulos, N. & Givoni, M., 2015. The autonomous car—a blessing or a curse for the future of low carbon mobility? An exploration of likely vs. desirable outcomes. *European Journal of Futures Research,* 3(14).

Tussyadiah, I. & Miller, G., 2019. Perceived Impacts of Artificial Intelligence and Responses to Positive Behaviour Change Intervention. In: *Information and Communication Technologies in Tourism.* s.l.:s.n., pp. 359-370.

Tussyadiah, I., Zach, F. & Wang, J., 2017. Attitudes toward Autonomous on Demand Mobility System: The Case of Self-Driving Taxi. In: *Information & Communication Technologies in Tourism.* s.l.:Springer International.

Tussyadiah, I., Zach, F. & Wang, J., 2018. *Do Traverlers Trust Intelligent Service Robots?*, s.l.: s.n.

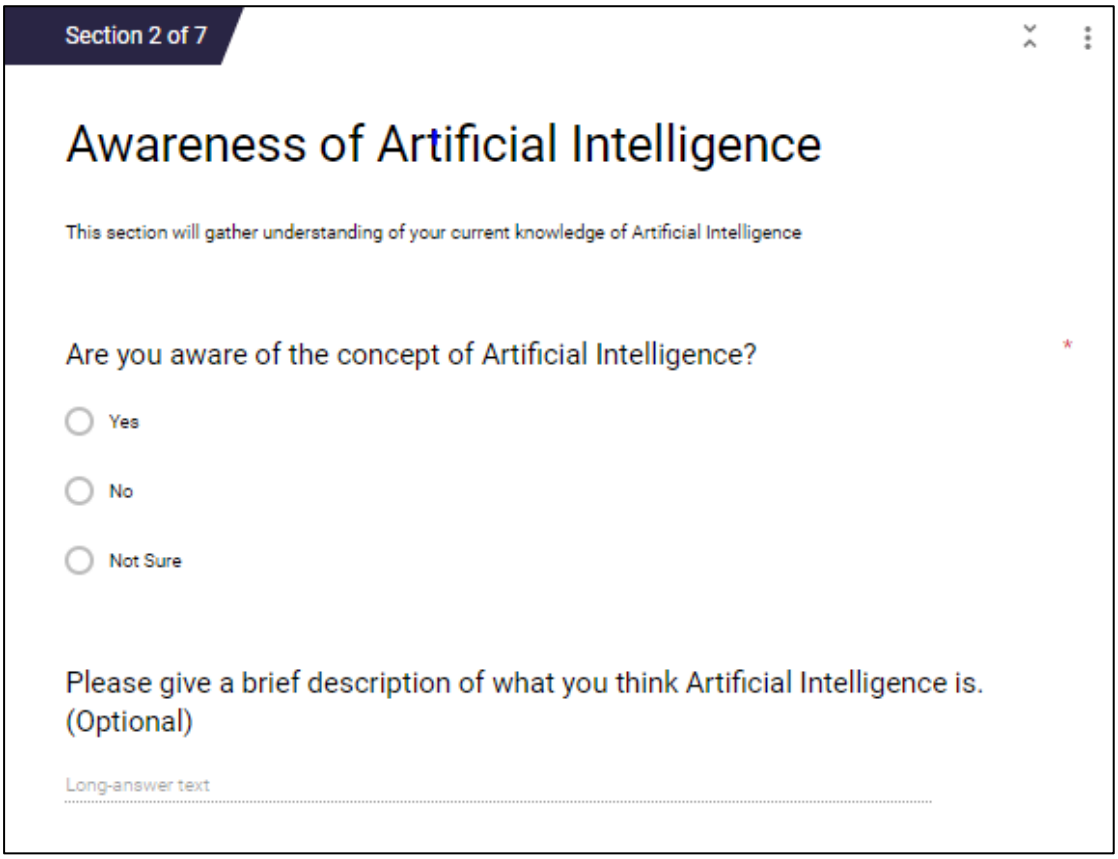
Williams, C., 2007. Research Methods. *Jounrnal of Business & Economic Research,* 5(3), pp. 65-72.

Wirtz, B., Weyerer, J. & Geyer, C., 2019. Artificial Intelligence and the Public Sector - Applications and Challenges. *International Journal of Public Administration,* 42(7), pp. 596-615.

8. Appendix

8.1. Appendix 1 – Online Survey Facebook Advert

8.2. Appendix 2 – Online Survey



Section 3 of 7

Radar-Guided Cruise Control

(Also known as 'Adaptive Cruise Control'), is similar to standard cruise control, however the system automatically adjusts the vehicles speed to maintain a safe distance from the vehicles ahead.

Are you aware of this system? *

- ○ Yes
- ○ No
- ○ Not Sure

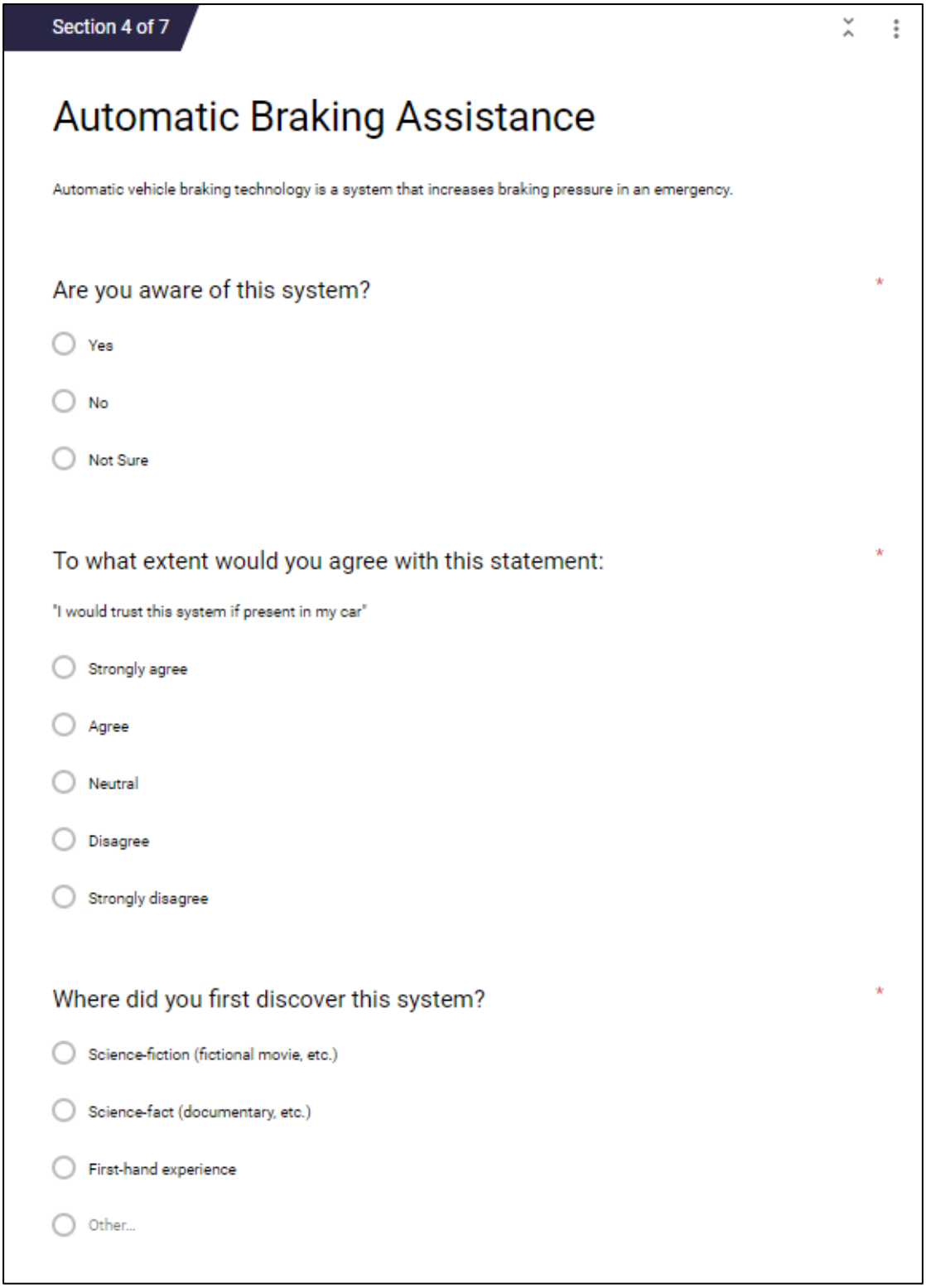
To what extent would you agree with this statement: *

"I would trust this system if present in my car"

- ○ Strongly agree
- ○ Agree
- ○ Neutral
- ○ Disagree
- ○ Strongly disagree

Where did you first discover this system? *

- ○ Science-fiction (fictional movie, etc.)
- ○ Science-fact (documentary, etc.)
- ○ First-hand experience
- ○ Other...



Section 4 of 7

Automatic Braking Assistance

Automatic vehicle braking technology is a system that increases braking pressure in an emergency.

Are you aware of this system? *

- ○ Yes
- ○ No
- ○ Not Sure

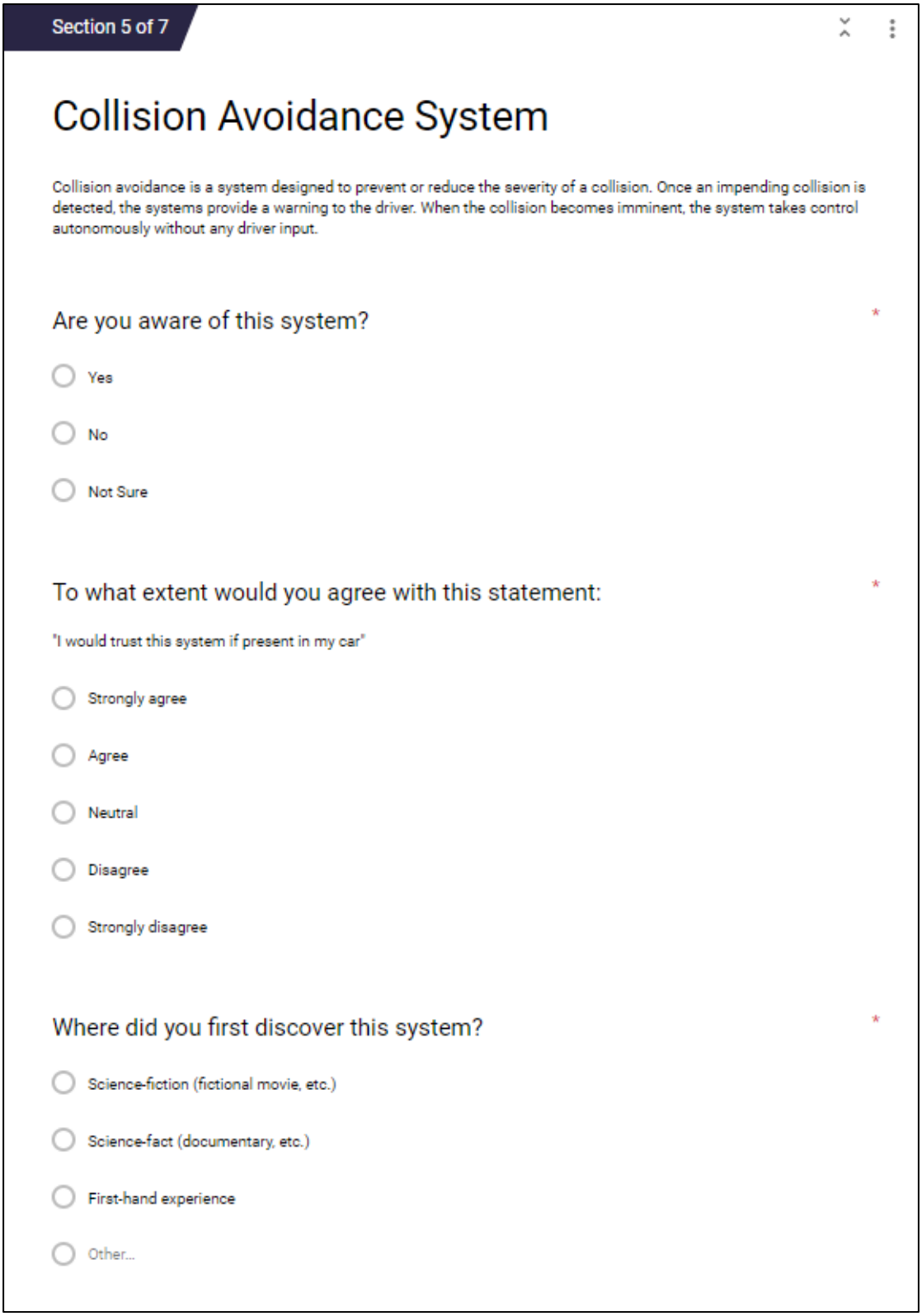
To what extent would you agree with this statement: *

"I would trust this system if present in my car"

- ○ Strongly agree
- ○ Agree
- ○ Neutral
- ○ Disagree
- ○ Strongly disagree

Where did you first discover this system? *

- ○ Science-fiction (fictional movie, etc.)
- ○ Science-fact (documentary, etc.)
- ○ First-hand experience
- ○ Other...



Section 5 of 7

Collision Avoidance System

Collision avoidance is a system designed to prevent or reduce the severity of a collision. Once an impending collision is detected, the systems provide a warning to the driver. When the collision becomes imminent, the system takes control autonomously without any driver input.

Are you aware of this system? *

- Yes
- No
- Not Sure

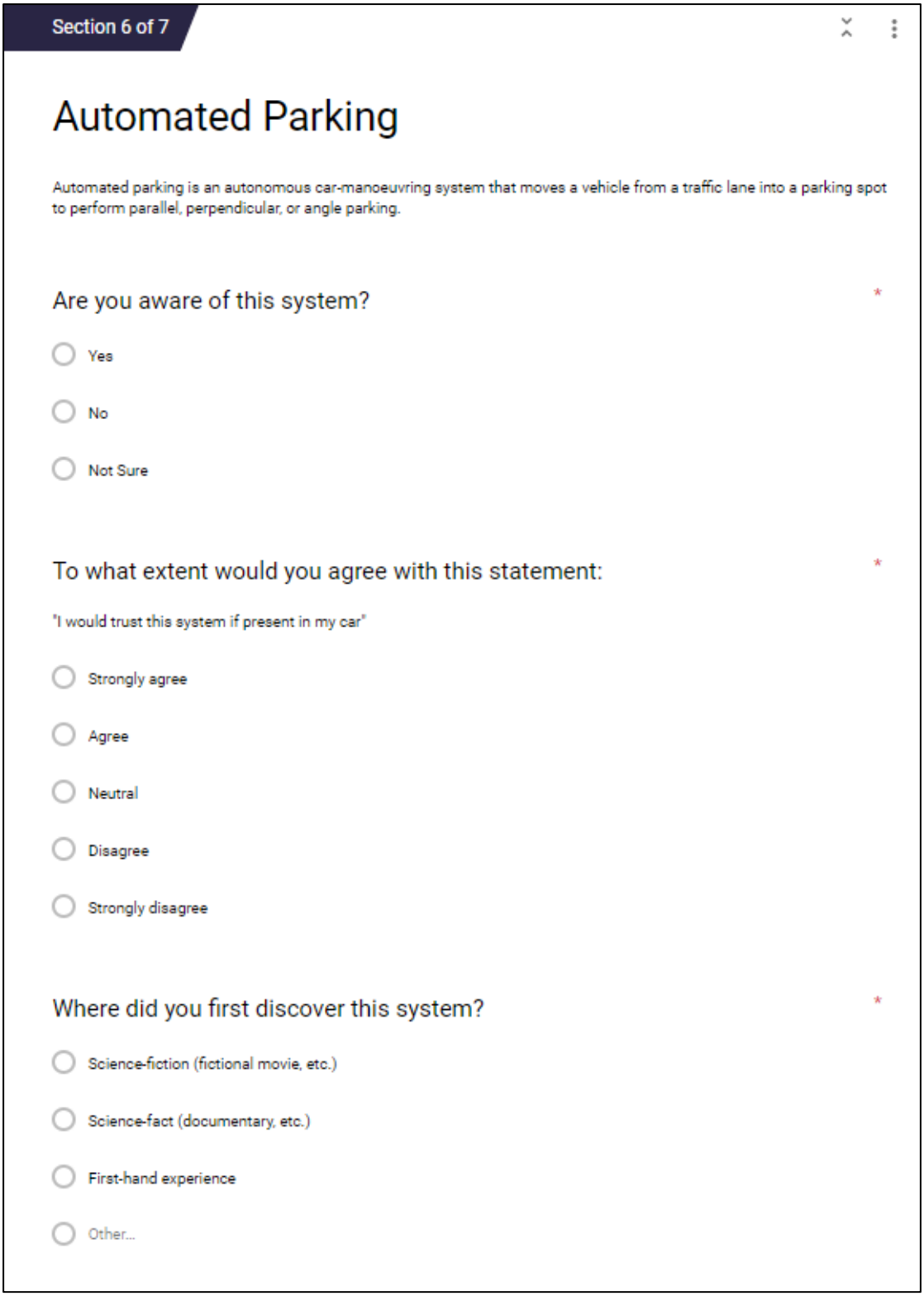
To what extent would you agree with this statement: *

"I would trust this system if present in my car"

- Strongly agree
- Agree
- Neutral
- Disagree
- Strongly disagree

Where did you first discover this system? *

- Science-fiction (fictional movie, etc.)
- Science-fact (documentary, etc.)
- First-hand experience
- Other...



Section 6 of 7

Automated Parking

Automated parking is an autonomous car-manoeuvring system that moves a vehicle from a traffic lane into a parking spot to perform parallel, perpendicular, or angle parking.

Are you aware of this system? *

- Yes
- No
- Not Sure

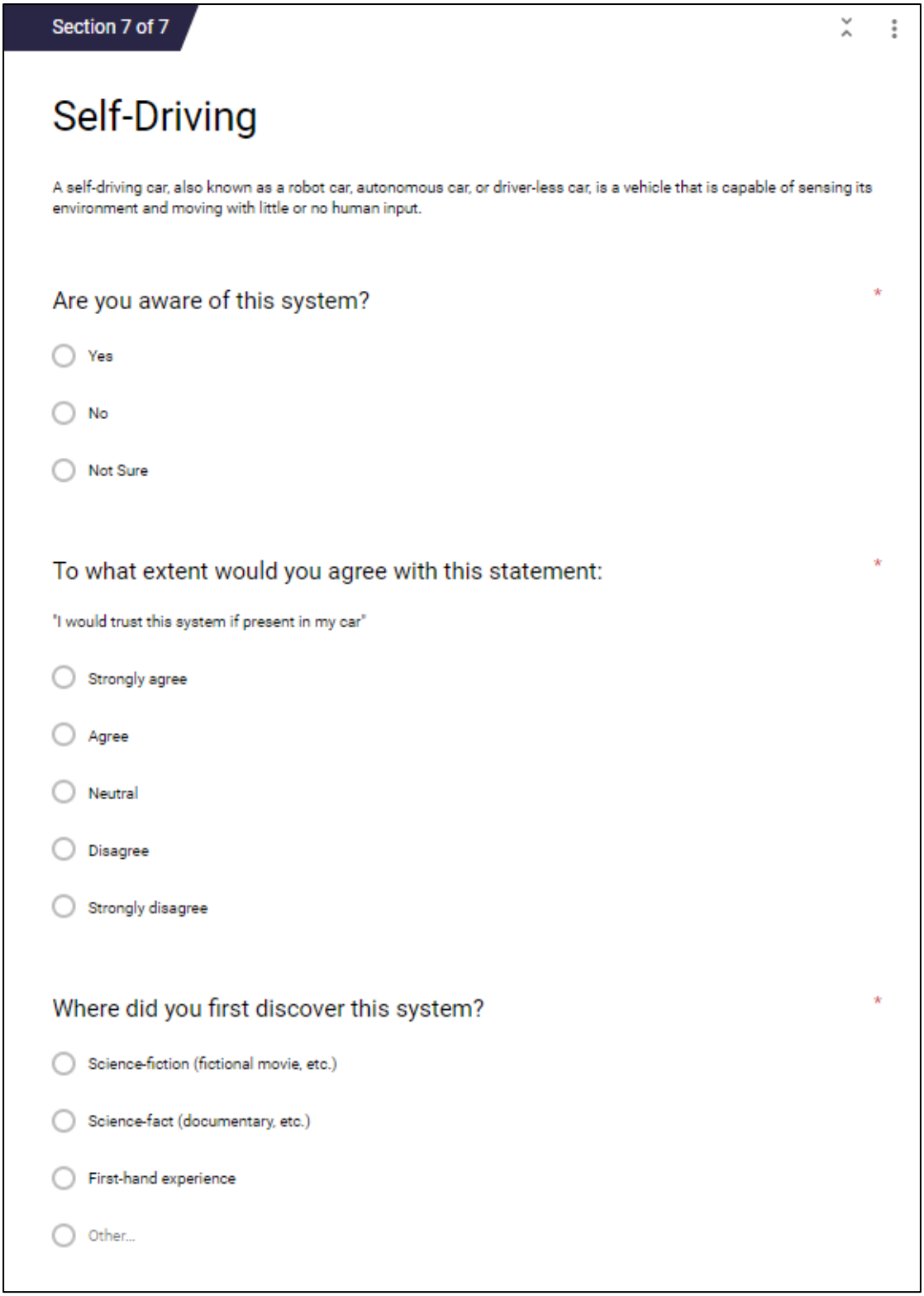
To what extent would you agree with this statement: *

"I would trust this system if present in my car"

- Strongly agree
- Agree
- Neutral
- Disagree
- Strongly disagree

Where did you first discover this system? *

- Science-fiction (fictional movie, etc.)
- Science-fact (documentary, etc.)
- First-hand experience
- Other...



Section 7 of 7

Self-Driving

A self-driving car, also known as a robot car, autonomous car, or driver-less car, is a vehicle that is capable of sensing its environment and moving with little or no human input.

Are you aware of this system? *

- ○ Yes
- ○ No
- ○ Not Sure

To what extent would you agree with this statement: *

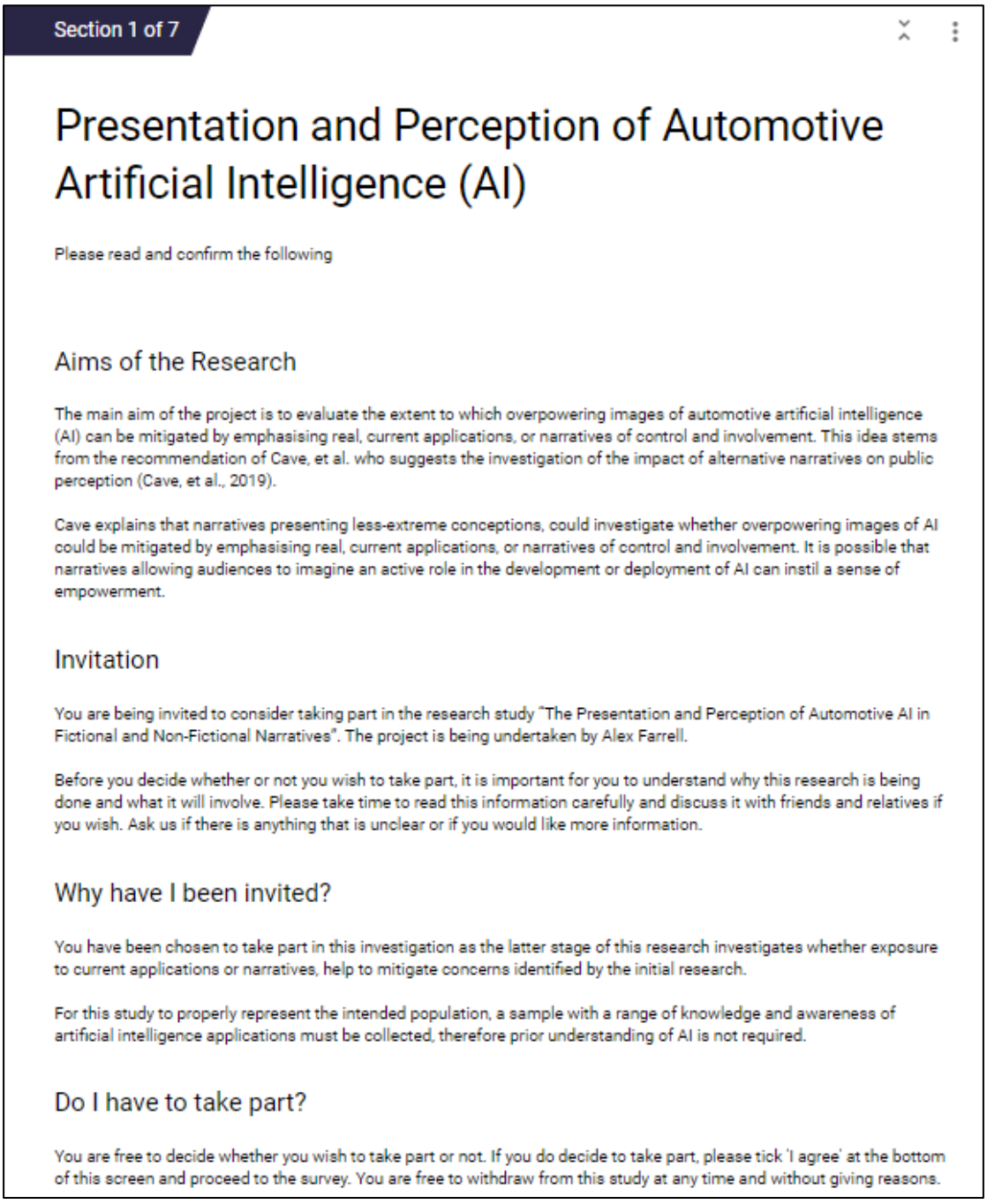
"I would trust this system if present in my car"

- ○ Strongly agree
- ○ Agree
- ○ Neutral
- ○ Disagree
- ○ Strongly disagree

Where did you first discover this system? *

- ○ Science-fiction (fictional movie, etc.)
- ○ Science-fact (documentary, etc.)
- ○ First-hand experience
- ○ Other...

8.3. Appendix 3 – Online Survey Consent Form



Section 1 of 7

Presentation and Perception of Automotive Artificial Intelligence (AI)

Please read and confirm the following

Aims of the Research

The main aim of the project is to evaluate the extent to which overpowering images of automotive artificial intelligence (AI) can be mitigated by emphasising real, current applications, or narratives of control and involvement. This idea stems from the recommendation of Cave, et al. who suggests the investigation of the impact of alternative narratives on public perception (Cave, et al., 2019).

Cave explains that narratives presenting less-extreme conceptions, could investigate whether overpowering images of AI could be mitigated by emphasising real, current applications, or narratives of control and involvement. It is possible that narratives allowing audiences to imagine an active role in the development or deployment of AI can instil a sense of empowerment.

Invitation

You are being invited to consider taking part in the research study "The Presentation and Perception of Automotive AI in Fictional and Non-Fictional Narratives". The project is being undertaken by Alex Farrell.

Before you decide whether or not you wish to take part, it is important for you to understand why this research is being done and what it will involve. Please take time to read this information carefully and discuss it with friends and relatives if you wish. Ask us if there is anything that is unclear or if you would like more information.

Why have I been invited?

You have been chosen to take part in this investigation as the latter stage of this research investigates whether exposure to current applications or narratives, help to mitigate concerns identified by the initial research.

For this study to properly represent the intended population, a sample with a range of knowledge and awareness of artificial intelligence applications must be collected, therefore prior understanding of AI is not required.

Do I have to take part?

You are free to decide whether you wish to take part or not. If you do decide to take part, please tick 'I agree' at the bottom of this screen and proceed to the survey. You are free to withdraw from this study at any time and without giving reasons.

What will happen if I take part?

If you decide to take part, you will first be asked your current understanding of the concept of Artificial Intelligence, followed by questions on various intelligent systems in cars. The survey will only take a few minutes to complete.

What are the risks of taking part?

Due to the nature of the research, the concerns expressed could relate to such things as car accidents, therefore if you have been involved in or recall any negative experiences that could cause emotional stress, there is no obligation to take part.

How will information about me be used?

This survey is completely anonymous, therefore no personal information about you will be collected or stored.

What if there is a problem?

If you have a concern about any aspect of this study, you may wish to speak to the researcher who will do their best to answer your questions. You should contact Alex Farrell on w4s22@students.keele.ac.uk. Alternatively, if you do not wish to contact the researcher you may contact Theocharis Kyriacou at t.kyriacou@keele.ac.uk.

If you remain unhappy about the research and wish to raise a complaint about any aspect of the way that you have been approached or treated during the course of the study please write to Nicola Leighton who is the University's contact for complaints regarding research at the following address: -

Nicola Leighton
Research Governance Officer
Research & Enterprise Services
Dorothy Hodgkin Building
Keele University
ST5 5BG

Email: n.leighton@uso.keele.ac.uk
Tel: 01782 733306

Please check if you agree to participate *

☐ I agree

8.4. Appendix 4 – Interview Script

Radar-Guided Cruise Control

1. What is your understanding of Rader-Guided Cruise Control?
2. *(Interviewer prompt: Give brief definition if subject is unsure or gives incomplete definition)*
3. If this system were present in a vehicle you were travelling in, would you trust it (1 – 7)?
4. *(Interviewer prompt: Show video from URL: https://www.youtube.com/watch?v=F949Lms4Uvg)*
5. Having watched the video, would you trust this system in a vehicle you were travelling in (1 – 7)?

Automatic Braking Assistance

1. What is your understanding of Rader-Guided Cruise Control?
2. *(Interviewer prompt: Give brief definition if subject is unsure or gives incomplete definition)*
3. If this system were present in a vehicle you were travelling in, would you trust it (1 – 7)?
4. *(Interviewer prompt: Show video from URL: https://www.youtube.com/watch?v=LivMzWoHAuQ)*
5. Having watched the video, would you trust this system in a vehicle you were travelling in (1 – 7)?

Collision Avoidance System

1. What is your understanding of Rader-Guided Cruise Control?
2. *(Interviewer prompt: Give brief definition if subject is unsure or gives incomplete definition)*
3. If this system were present in a vehicle you were travelling in, would you trust it (1 – 7)?
4. *(Interviewer prompt: Show video from URL: https://www.youtube.com/watch?v=rYckJqp4XTc)*
5. Having watched the video, would you trust this system in a vehicle you were travelling in (1 – 7)?

Automated Parking

1. What is your understanding of Rader-Guided Cruise Control?
2. *(Interviewer prompt: Give brief definition if subject is unsure or gives incomplete definition)*
3. If this system were present in a vehicle you were travelling in, would you trust it (1 – 7)?
4. *(Interviewer prompt: Show video from URL: https://www.youtube.com/watch?v=0kRArbNfOGA)*
5. Having watched the video, would you trust this system in a vehicle you were travelling in (1 – 7)?

Self-Driving

1. What is your understanding of Rader-Guided Cruise Control?
2. *(Interviewer prompt: Give brief definition if subject is unsure or gives incomplete definition)*
3. If this system were present in a vehicle you were travelling in, would you trust it (1 – 7)?
4. *(Interviewer prompt: Show video from URL: https://www.youtube.com/watch?v=VG68SKoG7vE)*
5. Having watched the video, would you trust this system in a vehicle you were travelling in (1 – 7)?

8.5. Appendix 5 – Interview Information Sheet

Keele UNIVERSITY

Information Sheet

Study Title

The Presentation and Perception of Automotive AI in Fictional and Non-Fictional Narratives

Aims of Research

The main aim of the project is to evaluate the extent to which overpowering images of automotive artificial intelligence (AI) can be mitigated by emphasising real, current applications, or narratives of control and involvement. This idea stems from the recommendation of Cave, et al. who suggests the investigation of the impact of alternative narratives on public perception (Cave, et al., 2019).

Cave explains that narratives presenting less-extreme conceptions, could investigate whether overpowering images of AI could be mitigated by emphasising real, current applications, or narratives of control and involvement. It is possible that narratives allowing audiences to imagine an active role in the development or deployment of AI can instil a sense of empowerment.

Invitation

You are being invited to consider taking part in the research study "The Presentation and Perception of Automotive AI in Fictional and Non-Fictional Narratives". The project is being undertaken by Alex Farrell.

Before you decide whether or not you wish to take part, it is important for you to understand why this research is being done and what it will involve. Please take time to read this information carefully and discuss it with friends and relatives if you wish. Ask us if there is anything that is unclear or if you would like more information.

Why have I been invited?

You have been chosen to take part in this investigation as the latter stage of this research investigates whether exposure to current applications or narratives, help to mitigate concerns identified by the initial research.

For this study to properly represent the intended population, a sample with a range of knowledge and awareness of artificial intelligence applications must be collected, therefore prior understanding of AI is not required.

Do I have to take part?

You are free to decide whether you wish to take part or not. If you do decide to take part you will be asked to sign two consent forms, one is for you to keep and the other is for our records. You are free to withdraw from this study at any time and without giving reasons.

What will happen if I take part?

If you decide to take part, you will be presented with an explanation of a particular concern and then shown a real-world application or narrative to judge whether exposure can decrease pessimism towards the concern.

Version No: 1
Date: 01/04/2019
1 for participant, 1 for researcher Page 1 of 3 ERP27314

What are the risks of taking part?

Due to the nature of the research, the concerns expressed could relate to such things as car accidents, therefore if you have been involved in or recall any negative experiences that could cause emotional stress, there is no obligation to take part.

How will information about me be used?

The researcher will be taking an audio recording of the 1 on 1 discussion so that he can analyse your feedback. This recording will only be accessible by the researcher for the purpose of this study. Once he has extracted your feedback from this recording, the audio file will be deleted.

Who will have access to information about me?

As mentioned, audio recordings of your responses in the discussion will be collected and only accessible to the researcher. These responses will be analysed, and key points extracted. Any points will be referenced anonymously in his project report. Om completion of the study, your information will be kept in a locked filing cabinet on the Keele University Campus.

What if there is a problem?

If you have a concern about any aspect of this study, you may wish to speak to the researcher who will do their best to answer your questions. You should contact Alex Farrell on w4s22@students.keele.ac.uk. Alternatively, if you do not wish to contact the researcher you may contact Theocharis Kyriacou at t.kyriacou@keele.ac.uk.

If you remain unhappy about the research and wish to raise a complaint about any aspect of the way that you have been approached or treated during the course of the study please write to Nicola Leighton who is the University's contact for complaints regarding research at the following address: -

Nicola Leighton

Research Governance Officer

Research & Enterprise Services

Dorothy Hodgkin Building

Keele University

ST5 5BG

Email: n.leighton@uso.keele.ac.uk

Tel: 01782 733306

Version No: 1
Date: 01/04/2019
1 for participant, 1 for researcher Page 2 of 3 ERP27314

8.6. Appendix 6 – Interview Consent Form

Keele UNIVERSITY

Consent Form

Title of Project

The Presentation and Perception of Automotive AI in Fictional and Non-Fictional Narratives

Name and contact details of Principle Investigator

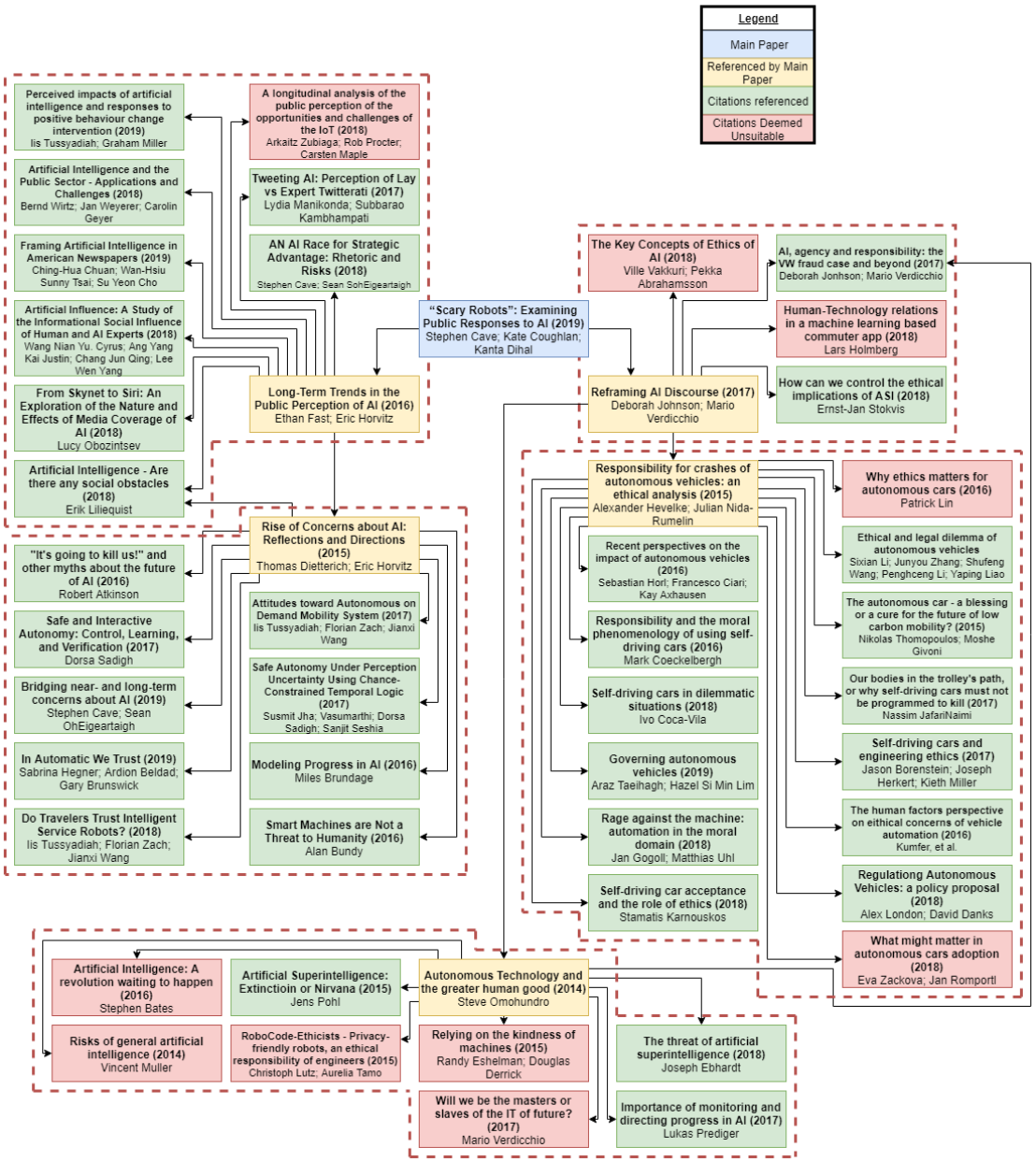
Alex Farrell, w4s22@students.keele.ac.uk

		Please tick box if you agree with the statement
1.	I confirm that I have read and understood the information sheet dated 01/04/2019 (version no. 1) for the above study and have had the opportunity to ask questions	☐
2.	I understand that my participation is voluntary and that I am free to withdraw at any time	☐
3.	I agree to take part in this study	☐
4.	I understand that data collected about me during this study will be anonymised before it is submitted for publication	☐
5.	I agree to the interview being audio recorded	☐
6.	I agree to allow the dataset collected to be used for future research projects	☐

Name of Participant	Date	Signature
Researcher	Date	Signature

Version No: 1
Date: 01/04/2019
1 for participant, 1 for researcher Page 3 of 3 ERP27314

8.7. Appendix 7 – Literature Review Structure (Hierarchy)



Project Plan

Project Overview and Description

Student Name: Alexander Farrell
Student Username: w4s22
Student Number: 15005594
Module (*delete as appropriate)*: CSC-40047
Supervisor Name: Dr Theocharis Kyriacou
Project Title: The Presentation and Perception of Automotive AI in Fictional and Non-Fictional Narratives

Please provide a brief Project Description:

It is commonly misconceived that the automobile industry can be lethargic in its adaptation to new technologies, including artificial intelligence, however in reality, there are a number of new technologies that are deployed through vehicles that are then translated to other industries (Gusikhin, et al., 2007). This realisation correlates to the concerns that people may have regarding new technology and their subsequent acceptance.

Public acceptance can be influence by a number of factors, one of which is the presentation and perception of technology in both fictional and non-fictional narratives. A study conducted by Cave, et al. identified 8 factors categorized as 'hopes and fears' that are commonly represented in narratives.
Background

What are the aims and objectives of the Project?

The main aim of the project is to evaluate the extent to which overpowering images of automotive AI can be mitigated by emphasising real, current applications, or narratives of control and involvement. This idea stems from the recommendation of Cave, et al. who suggests the investigation of the impact of alternative narratives on public perception (2019).

Please provide a brief overview of the key literature related to the Project:

(Cave, et al., 2019) – This will be the main paper that the research is based on, looking at the perception of AI in the media. This research will be adapted to fit Ai systems in the automotive industry.
(Fast & Horvitz, 2017)
(Johnson & Verdicchio, 2017)

Project Process and Method

Please provide a brief overview of the Methodology to be used in the Project (inc. an overview of best practice within the Methodology):

- Conduct a systematic literature review to uncover concerns that the public have towards automotive AI
- Conduct a survey to gather information on current perception of AI and the narratives those perceptions originate from
- Conduct interviews to test whether exposure to real-world examples that depict concerns, can mitigate these concerns and improve trust.

Which Data Collection Methods will be employed (e.g. card sorts, questionnaires, simulations, …)?

- Systematic Literature Review
- Online Survey
- Semi-structured Interview

Time and Resource Planning

Will Standard Departmental Hardware be used? YES

If NO please outline the Hardware/Materials to be used:

N/A

Will Software which is already available in department be used? YES

If NO please outline the Software to be used including how any necessary licences will be obtained:

N/A

Will the project require any Programming? NO

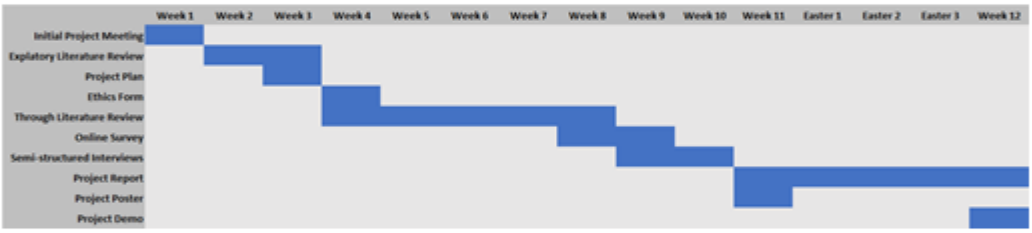
If YES please list the (potential) Programming Languages to be used (including any IDEs and Libraries you may make use of):

N/A

Table of Risks (*if non Standard Hardware and/or Software to be used please include backup options/ contingency plans here*):

Risk	Likelihood	Avoidance	Actions on occurrence
Running out of time	Possible	Aim to stick to schedule laid out on Gantt Chart	
No concerns identified	Unlikely	Search as wide-a-range of systems and literature	Choose relatively new concerns that may only exist in science-fiction
Lack of survey answers	Unlikely	Publish survey to substantial audience to maximise return	Use survey answers collected as proof of concept
Lack of interviewees	Possible	Offer some sort of incentive, either financial or material	Use interview results collected as proof of concept

Gantt Chart/ Pert Chart (must include milestones and deliverables):



References and Administration

Please include a list of References used in this Plan:

Cave, S., Coughlan, K. & Dihal, K., 2019. *"Scary Robots": Examining Public Responses to AI.* Honolulu, Association for the Advancement of Artificial Intelligence.

Dietterich, T. & Horvitz, E., 2015. Rise of Concerns about AI: Relfections and Directions. *Communications of the ASM,* 58(10), pp. 38-40.

Fast, E. & Horvitz, E., 2017. *Long-Term Trends in the Public Perception of Artificial Intelligence.* San Francisco, AAAI.

Gusikhin, O., Rychtyckyj, N. & Filev, D., 2007. Intelligent systems in the automotive industry: applications and trends. *Knowledge and Information Systems,* 12(2), pp. 147-168.

Hevelke, A. & Nida-Rümelin, J., 2015. Responsibility for Crashes of Autonomous Vehicles; An Ethical Analysis. *Sci Eng Ethics,* Volume 21, pp. 619-630.

Johnson, D. & Verdicchio, M., 2017. Reframing AI Discourse. *Minds & Machines,* Volume 27, pp. 575-590.

Ethics Form

School of Computing and Mathematics: Student Project Ethics Committee Application form (U/G and PGT Students)

Please print off a hard copy of this form and submit it to the School Office.
Please remember to sign it, date it, and to get your supervisor's signature on it.

Student name:	**Alex Farrell**
Student number:	15005594
Course:	Computer Science MComp
Date:	28/03/2019

Part A: (all students)

Does the topic of the project involve any of the following?	**YES**	**NO**
Recall of personal or sensitive memories		X
Reporting or discussion of personal or sensitive topics		X
Tasks which could be harmful or distressing		X
A significant risk of participants later regretting taking part		X
Procedures which are likely to provoke inter-personal or inter-group conflict?		X

If you answer "Yes" to any questions on Part A, then please also complete the University Ethics form (on the KLE) and seek guidance from the School Research Governance officer. The School Research Governance officer (mentioned in the university form) is the projects co-ordinator, Gordon Rugg.

Part B: (if you are doing a software design and/or software build or gathering data from human participants)

Requirements gathering and evaluation: use of unusual techniques	YES	NO
Will the techniques you are using for requirements gathering and software evaluation be unusual in a way which could cause ethical problems?		X
Will any of the participants be from a vulnerable group (e.g. under 18, or with learning difficulties, or under pressure to help you)?		X

If you answer "Yes" to any questions on Part B, then please also complete the University Ethics form (on the KLE) and seek guidance from the School Research Governance officer. The School Research Governance officer (mentioned in the university form) is the projects co-ordinator, Gordon Rugg.

Student's signature and date:	
Supervisor's signature and date:	