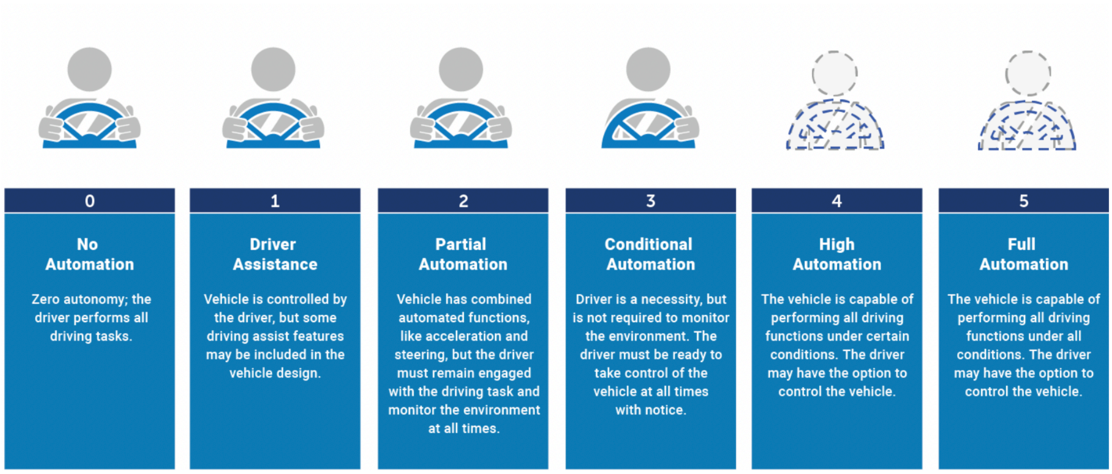
Is there a fundamental difference between self-driving cars and a "slaughter army" of killer drones?

Self-driving cars are a common talking point of modern technological ethics, with their constantly predicted take-over yielding divided opinions. A self-driving car is defined as 'capable of sensing its environment and moving safely with little or no human input[1]', with ambiguity clarified using the SAE's automation levels (Figure 1[2]).



Although unspecified, this question can be assumed to refer to fully autonomous vehicles (AVs) of Level 4/5, so these technologies will be examined further. Similarly, whilst a 'slaughter army of killer drones' seems a nightmarish sci-fi proposition, similar devices exist as Lethal Autonomous Weapon Systems (LAWS)[3], which can be used a benchmark for comparison. Whilst methods for investigation are primarily hypothetical since Level 5 vehicles are yet to exist, I will examine the proposed 'fundamental' differences of whether each will increase or decrease fatalities, alongside whether deaths caused by each are morally comparable. I aim to argue that despite some similarities, there are fundamental differences between self-driving cars and 'killer drones' in these areas.

Predicted reasons for increased or decreased fatalities from the state of each technology's current climate

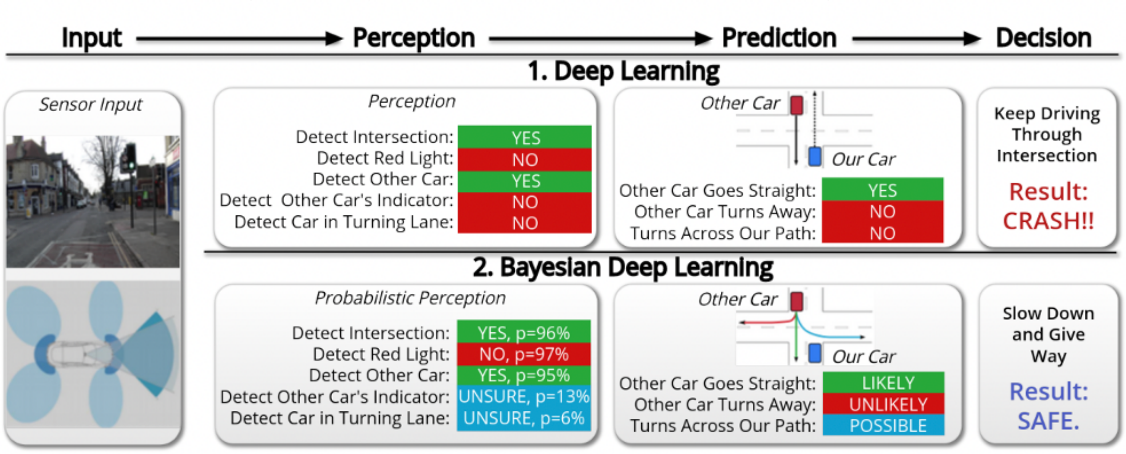
This question assumes AVs can cause excessive harm, increasing road accidents to a level of 'slaughter'. Whilst many are skeptical about the inevitable errors, fuelled by biased media coverage and fear of the unknown, the possibility of AVs becoming dangerous to this extent is unlikely. Arguably, the primary goal of autonomous driving has always been to reduce accidents, with estimation that full AV deployment could do so by 90%[4]. Conversely, LAWS are indisputably designed to favour killing, and are predicted to worsen the number of military casualties.

Decreased fatalities due to AVs

Computers, especially AI based, are better suited to multichannel information processing than the human brain[5], meaning AV technology may surpass human drivers in handling decisions regarding large sets of input data, precisely and simultaneously. In the average road scenario, 'human drivers still generally maintain an advantage in terms of reasoning[5]'. However, when additional stimuli appear at once, the driver will inevitably be shocked, potentially risking an unsafe reaction, especially on long journeys. In an AV, exteroceptive sensors such as LiDAR, Radar or 2D cameras, alongside proprioceptive sensors, provide a comprehensive stream of perception data, which decision-making AI software constantly analyses[6]. The vehicle is always aware of all possible hazard types, rather than shifting its focus level like a human. Considering most road accidents are attributed to complex simultaneous factors[7], the ability of AVs to better deal with many simultaneous stimuli, would mean they better handle scenarios likely to cause an accident without becoming 'overwhelmed', reducing the risk of choosing an inadequate action, causing a collision.

In addition, legal driving requirements such as 'assured clear distance ahead' (maintaining a distance between a vehicle and obstacle for a given speed that allows stopping without a collision[8]), can have their mathematical formula programmed into an AV. Through systems such as Mobileye's RSS model, recently confirmed to have increased safety performance over humans[9], the 'safe distance for each velocity' won't be violated, even when other stimuli normally take priority in for humans. Autonomous decision-making extends to, for example, splitting surroundings into an 'occupancy grid', and finding a route between two positions under multiple constraints[6], relying on mathematical techniques which provide continuous accuracy that humans cannot achieve without fluctuations.

Of the three broad subtasks of perception, planning, and control in AV software, perception contributes the 'main source of vulnerabilities'[6]. Perception relies heavily on machine learning to form a 3D map of the environment, increasingly through deep neural networks[6]. Subsequently, it carries the inherent limitations of machine learning. 'If sensors "see" what they have seldom seen before, they don't know how to handle those situations'[10], evident in most past fatal AV crashes. Two major crashes involving Tesla's Model S Autopilot, are attributed to AVs' struggle in detecting objects perpendicular to the vehicle's direction of travel, due to this situation rarely arising in training data[11]. Elaine Herzberg's death in 2018 was speculated to be due to her bicycle's metal parts preventing sensors classifying her correctly until 1.2 seconds before impact[12]. Whilst initially these cases are due to unreliable perception, the former were caused by Level 2 AVs, and the latter by trial Level 3 AV, with a safety driver[13]. The extent of harm was ultimately the drivers' responsibility, since SAE states Level 2/3 vehicles require human presence, able to 'remain engaged' and 'ready to take control' 'at all times[2]', echoed by Elon Musk's claims that Autopilot is never responsible for any crash[14]. Whilst some studies suggest 'severe injuries can happen if' the automated vehicle 'is the major responsible party for the crash[15]', these conclusions are based on test vehicles which will be highly modified before Level 4/5 release. Novel approaches promise improved perception technology, such as Bayesian probability methods which counteract input uncertainties that propagate through AV components from the model's ignorance about the world (Figure 2)[16].



This is especially effective in 'end-to-end' deep learning systems, to 'focus its learning on aspects of the scene most useful for the end task[16]', an implementation increasing in popularity[6]. This cutting edge AI research, alongside other strands, such as the 'resurgence of structure' in neural networks by reconsidering 'hard-coded' elements[17], will continue to be implemented into progressing AV technology, and alongside models continually being trained to recognise more flexible scenarios, future Level 5 self-driving cars should include computer vision superior to both current systems, and human drivers. In short, 'the technology is improving, while humans are not[18]'.

AV systems cannot succumb to 'all-too-human fits of anger, sadism, cruelty[19]', nor disobey safe driving rules. 'Driving too fast for the road or driving conditions is still one of the most common contributory factors in collisions[20]', implying humans are what often jeopardise road safety. This question however, suggests humans are the reliable element of driving, and their replacement with autonomy would be dangerous, questionable after examining AI's nature. Humans learn the nuances of driving from hours of trial-and-error in a car, and therefore their driving style mimics the conventional behaviour of other drivers, primarily their instructor. In contrast, the 'planning' module of AV software is either traditionally implemented to recognise 'hand-crafted' features determined by ideal road conduct, or increasingly, trained 'through deep reinforcement learning to perform safe and efficient driving' without pre-set parameters[6]. Either way, the system learns to exhibit cautious behaviour, with predictable, consistent control, and without conventional human complacency. Inevitably, it's easier to control the safety of AV technology and associated infrastructure, than actions of individual drivers such as fluctuating attention or impairment, meaning that if regular examinations and clear regulations are facilitated, the risks AVs may pose are more rectifiable than human mistakes which endanger current roads the most.

Whilst the question discusses the general danger of self-driving cars, when full Level 5 vehicles are deployed, this will likely be done as a collective system; a network of 'connected vehicles that can "talk" to other traffic[5]'. On analysis of Californian traffic accidents involving AVs and conventional vehicles (CVs) from 2015-17, the most statistically significant difference of AVs being more 'dangerous', was the number of 'rear-end' collisions; 64.2% in AV accidents, and 28.3% with only CVs[4]. The assumption for this, as echoed in other research including simulations, is that 'drivers of the CVs are not accustomed to the driving style of AVs' complying more with regulations[4]. Proposed coherence between 'high level driving strategies' of individual AVs will maximise awareness between AVs in their observations and actions[21], and provide a safer collective future road environment. Whether through real-time wireless communication with surrounding infrastructures/vehicles[22], or establishing characteristics in driving strategy which ensure the 'compatibility and coexistence' of different manufacturers' AVs[21], treating them as one system rather than single entities, is 'potentially one of the most transformational technologies' in improving road safety[23], avoiding any human miscommunication through angry gestures and eye contact.

95% of current US road accidents are due to driver behaviour[15], and so whilst rare situations are still unfamiliar to AVs, as they progress to Level 5, overall road accident fatalities should reduce as a result of the advantages of automation explored above.

<u>Increased fatalities due to LAWS</u>

Conversely, little evidence is needed to suggest that future LAWS will increase destruction in warfare, despite claims they could reduce fatalities. Some delegations at the CCW 2015 meeting claimed wars between robots could 'avoid human casualties altogether', or at least 'reduce civilian casualties' from 'more precise targeting[3]' .

However, most argue that easy mass production of these weapons, alongside the 'physical and emotional distance' of deploying them compared to human soldiers will 'change how humans exercise control over the use of force[24]', resulting in their excessive use, potentially outside of clear-cut military settings[3]. Additionally, whilst malfunctions are not a disadvantage of LAWS inherently, operational flaws in this setting would likely default to causing more damage than good, due to their destructive nature[3]. If LAWS were to become the future norm, levels of fatality would continually increase in line with the technology's progression, becoming less predictable and more dangerous as the systems advance[24].

Ultimately, these grave predictions outweigh any other advantages of LAWS[3]. This is well recognised, with tens of thousands of lead researchers signing the 2015 petition to the UN to ban the proliferation of fully autonomous weapons, clearly believing that a future with LAWS would kill more people than wars involving conventional weapons[3].

Whilst further investigation is required to solidify all predictions raised in the section above, the contrast in how each technology will supposedly affect future safety, poses one fundamental difference between the technologies.

Philosophical differences between the morality of intent and damage in self-driving cars and autonomous weaponry

Whilst AVs and 'killer drones' (LAWS) are each able to cause damage, the question is raised if a mass of deaths in each case is morally comparable, and whether these differences in intent are 'fundamental'. The Cambridge dictionary defines 'slaughter' as 'to cruelly and unfairly kill a lot of people', with the definition of 'cruel' concluding with 'intentionally[25]'. Whilst these technologies are physically similar, consisting of a mobile hardware system, environmental sensing capabilities, and tactical decision-making frameworks[3], they have been created with different intentions.

In military settings, LAWS make decisions similarly to the AI systems of AVs; identifying their opposing targets using computer vision, under conditions set by developers, and executing their kill as an AV would a manoeuvre, with the help of weaponry such as the 1kg explosives in Turkey's Kargu drone[3]. When considering LAWS in a hypothetical setting more comparable to AVs, a 'slaughter army' can be assumed to aim for as many deaths as possible, using AI for vision, navigation, and control. Either implementation of a 'slaughter army' shows it has been intentionally instructed to kill, whether through directly defined targets, or indirectly through the software's underlying purpose.

To examine this difference's significance, we can look to English Law, and the conditions of murder and manslaughter. When comparing LAWS directly to AVs, disregarding the altered moral conditions for warfare, and quoting the designers of the respective systems as the conviction focus, there remain differences in how killings by each technology are viewed. The primary distinctions between murder and manslaughter are a 'partial defence' despite murderous intent (voluntary manslaughter), or accidental killing, due to gross negligence, or an 'unlawful and dangerous act' (involuntary manslaughter)[26]. The basis on which these convictions differ is clearly **intent**, echoed in the legal slogan 'actus non facit reum nisi mens sit rea: an act does not make a person guilty unless her mind is guilty too[27]'.

As explored previously, deaths caused by future AVs should be the result of a technical malfunction (unless corrupted), and therefore unintentional by developers, since a main reason for the creation of AVs is improved safety. Deaths caused by AVs would therefore likely fall under gross negligence manslaughter. Deaths caused by LAWS (outside military settings) would likely be classed as murder, with lawyers echoing there's 'no new gap here between criminal intent and criminal negligence[3]'. Although 'slaughter' isn't synonymous with murder, relating deaths from killer drones to murder, and AVs to manslaughter, provides evidence for intent being a fundamental difference between these technologies.

Similarly, as proposed by Aristotle, an agent holds 'moral responsibility' for something 'if and only if the action and/or disposition is voluntary', where 'it must be up to the agent whether to perform that action' and 'the agent must be aware of what it is she is doing[28]'. This provides two philosophical conditions; control and awareness. In this scenario, the latter can be somewhat overlooked, with the technologies as an extension of their creators' awareness in programming decisions, and disregarding conscious machines. Conscious machines are not an impossibility, however such proposed systems are likely to 'require a particular architecture' not indicative of AVs[29], and are unlikely to be allowed to completely overwrite human-set objectives anyway, due to increasingly tight regulation[30]. Both technologies' creators therefore satisfy Aristotle's second condition for awareness.

However, these technologies differ in first condition. Although the ML of AVs is not explicitly programmed, the decision-making framework is heavily tested to avoid fatal mistakes[6]. Whether algorithms are based around a consequentialist or deontological ethical approach[31], ultimately, developers have not controlled the system to cause harm, therefore any deaths from malfunction are involuntary, and the creators hold no negative 'moral responsibility'. On the contrary, creators of LAWS including 'killer drones', do have control over intentionally killing, as explained above, therefore possessing a negative 'moral responsibility', since this action is voluntary.

Modern philosophy also considers that 'our responsibility practices are inherently social', and over-intellectualising morality is irrelevant without considering life's 'interpersonal nature' and individual attitudes[28]. In accidental situations, 'people clearly are reluctant to say that x intentionally brought about y, no matter what they suppose about x's responsibility for y[32]'; the difference in intent between deaths, is as 'fundamental' as society perceives it to be, regardless of classical standpoints. Clearly, the vast majority of society does differentiate between fatalities by AVs and LAWS, with BAE chairman Roger Carr stating 'delegating kill decisions to machines is fundamentally wrong[3]'.

Whilst law's treatment of deaths caused by AVs is yet to be known, in short, LAWS are being restricted, whilst alongside regulations, autonomous vehicles are being encouraged, evidence there is a fundamental philosophical difference between the technologies' moral standings in society.

<u>**Conclusion**</u>

Whilst there are evidenced differences between self-driving cars and 'killer drones' in both their predicted effects, and philosophical purposes, whether this constitutes a 'fundamental' difference, is subjective. Their architecture and operation is similar; subsequently, the explored differences in how these technologies, and the deaths they may cause, are perceived, are inherently socio-psychological. Despite my belief they can be differentiated, the extent of this ‘fundamental’ difference in future societies, will depend on how their development is controlled by humans.

Bibliography

[1] Wikipedia (2022) ‘Self-driving car’ [Online] Available from: https://en.wikipedia.org/wiki/Self-driving_car [accessed 6 March 2022]
[2] NHTSA (2020) ’Automated Vehicles for Safety’ [Online] Available from: https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety [accessed 4 February 2022]
[3] BBC Sounds (2021) ‘The Reith Lectures: AI in Warfare’ [Online podcast] Available from: https://www.bbc.co.uk/programmes/m00127t9 [accessed 5 March 2022]
[4] Transportation Research Procedia Volume 45 p.161-168 (2020) ‘Traffic Accidents with Autonomous Vehicles: Type of Collisions, Manoeuvres and Errors of Conventional Vehicles’ Drivers’ [Online] Available from: https://www.sciencedirect.com/science/article/pii/S2352146520301654 [accessed 27 February 2022]
[5] B. Schoettle (2017) ‘Sensor Fusion: A Comparison of Sensing Capabilities of Human Drivers and Highly Automated Vehicles’ [Online] Available from: http://websites.umich.edu/~umtriswt/PDF/SWT-2017-12_Abstract_English.pdf [accessed 23 February 2022]
[6] ENISA-JRC (2021) ‘Cybersecurity Challenges in the Uptake of Artificial Intelligence in Autonomous Driving’ [Online] Available from: https://www.enisa.europa.eu/publications/enisa-jrc-cybersecurity-challenges-in-the-uptake-of-artificial-intelligence-in-autonomous-driving/ [accessed 5 February 2022]
[7] Accident Analysis & Prevention Volume 115 p.11-24 (2018) ‘What are the factors that contribute to road accidents? An assessment of law enforcement views, ordinary drivers’ opinions, and road accident records’ [Online] Available from: https://www.sciencedirect.com/science/article/pii/S0001457518300873 [accessed 20 February 2022]
[8] Columbus Criminal Attorney (2022) ‘Columbus ACDA Ticket Lawyer’ [Online] Available from: https://www.columbuscriminalattorney.com/traffic-lawyer/assured-clear-distance-ahead/ [accessed 16 February 2022]
[9] Sensors, Volume 2, Article 6733 (2021) ‘On the Development of Autonomous Vehicle Safety Distance by an RSS Model Based on a Variable Focus Function Camera’ [Online] Available from: https://www.mdpi.com/1424-8220/21/20/6733 [accessed 27 February 2022]
[10] A. Hawkins (2019) ‘Tesla didn’t fix an Autopilot problem for three years, and now another person is dead’ [Online] Available from: https://www.theverge.com/2019/5/17/18629214/tesla-autopilot-crash-death-josh-brown-jeremy-banner [accessed 29 January 2022]
[11] A. Hawkins (2019) 'Tesla’s Autopilot lulled driver into a state of ‘inattention’ in 2018 freeway crash’ [Online] Available from: https://www.theverge.com/2019/9/4/20849499/tesla-autopilot-crash-culver-city-2018-ntsb-report [accessed 10 February 2022]
[12] R. Gonzales (2019) ‘Feds Say Self-Driving Uber SUV Did Not Recognize Jaywalking Pedestrian In Fatal Crash’ [Online] Available from: https://www.npr.org/2019/11/07/777438412/feds-say-self-driving-uber-suv-did-not-recognize-jaywalking-pedestrian-in-fatal-?t=1644772228740 [accessed 11 February 2022]
[13] C. Coppola, B. Frank (2018) 'Report: Uber driver was watching 'The Voice' moments before fatal Tempe crash’ [Online] Available from: https://eu.azcentral.com/story/news/local/tempe-breaking/2018/06/21/uber-self-driving-car-crash-tempe-police-elaine-herzberg/724344002/ [accessed 12 February 2022]
[14] S. O’Kane (2018) ‘Tesla will regularly release data about the safety of Autopilot, Elon Musk says’ [Online] Available from: https://www.theverge.com/2018/5/2/17313324/tesla-autopilot-safety-statistics-elon-musk-q1-earnings [accessed 11 February 2022]
[15] International Review of Law and Economics, Volume 68, Article 106022 (2021) 'Car accidents in the age of robots’ [Online] Available from: https://www.sciencedirect.com/science/article/pii/S0144818821000466 [accessed 19 February 2022]
[16] Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence p.4745-4753 (2017) ’Concrete Problems for Autonomous Vehicle Safety: Advantages of Bayesian Deep Learning’ [Online] Available from: https://www.repository.cam.ac.uk/bitstream/handle/1810/266683/McAllister.pdf?sequence=5&isAllowed=y [accessed 26 February 2022]
[17] P. Veličković (2019) ‘The resurgence of structure in deep neural networks’ [Online] Available from: https://www.repository.cam.ac.uk/handle/1810/292230 [accessed 1 March 2022]
[18] D. Newcomb (2017) ‘Humans Are Still Better Drivers Than Machines—for Now’ [Online] Available from: https://uk.pcmag.com/old-news/90692/humans-are-still-better-drivers-than-machines-for-now [accessed 15 February 2022]
[19] F. Pasquale (2020) “Machines set loose to slaughter’: the dangerous rise of military AI’ [Online] Available from: https://www.theguardian.com/news/2020/oct/15/dangerous-rise-of-military-ai-drone-swarm-autonomous-weapons [accessed 1 March 2022]
[20] AA (2020) ‘Driven to collision: common causes of road accidents’ [Online] Available from: https://www.theaa.com/car-insurance/advice/road-accident-car-damage [accessed 14 February 2022]
[21] Accident Analysis & Prevention, Volume 150, Article 105937 (2021) 'A comparative study of state-of-the-art driving strategies for autonomous vehicles’ [Online] Available from: https://www.sciencedirect.com/science/article/pii/S0001457520317577 [accessed 19 February 2022]
[22] Sensors, Volume 19, Article 3165 (2019) 'A Blockchain Framework for Securing Connected and Autonomous Vehicles’ [Online] Available from: https://www.mdpi.com/1424-8220/19/14/3165 [accessed 22 February 2022]
[23] IEEE Access, Volume 4, p.9413-9420 (2016) ’Decentralized Cooperative Lane-Changing Decision-Making for Connected Autonomous Vehicles’ [Online] Available from: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7811295 [accessed 22 February 2022]
[24] R. Surber (2018) ‘Artificial Intelligence: Autonomous Technology (AT), Lethal Autonomous Weapons Systems (LAWS) and Peace Time Threats’ [Online] Available from: https://ict4peace.org/wp-content/uploads/2018/02/2018_RSurber_AI-AT-LAWS-Peace-Time-Threats_final.pdf [accessed 6 March 2022]
[25] Cambridge Dictionary [Online] Available from: https://dictionary.cambridge.org/dictionary/english/slaughter and https://dictionary.cambridge.org/dictionary/english/cruel [accessed 15 February 2022]
[26] The Crown Prosecution Service (2021) 'Homicide: Murder and Manslaughter’ [Online] Available from: https://www.cps.gov.uk/legal-guidance/homicide-murder-and-manslaughter [accessed 5 March 2022]
[27] Philosophy Compass, Volume 4, p.978-986 (2009) ‘Legal and Moral Responsibility’ [Online] Available from: https://dspace.stir.ac.uk/bitstream/1893/2472/1/Duff%20Legal%20and%20Moral%20Responsibility.pdf
[28] A. Eshleman (2014) ‘Moral responsibility’ [Online] Available from: https://pilotscholars.up.edu/cgi/viewcontent.cgi?article=1001&context=phl_facpubs [accessed 5 March 2022]
[29] A. Kadin (2021) ‘Conscious Machines for Autonomous Agents and Cybersecurity’ [Online] Available from: https://www.techrxiv.org/articles/preprint/Conscious_Machines_for_Autonomous_Agents_and_Cybersecurity/17308406 [accessed 1 March 2022]
[30] S. Shead (2021) ‘A.I. researchers urge regulators not to slam the brakes on its development’ [Online] Available from: https://www.cnbc.com/2021/03/29/ai-researchers-urge-regulators-not-to-slam-brakes-on-development.html [accessed 2 March 2022]
[31] H. Bradshaw-Martin (2020) ‘Could your self-driving car choose to kill you?’ [Online] Available from: https://www.sciencefocus.com/future-technology/could-your-self-driving-car-choose-to-kill-you/ [accessed 27 February 2022]
[32] Philosophical Studies, Volume 82, p265-287 (1996) ‘Intention, Intentional Action, and Moral Responsibility’ [Online] Available from: https://www.jstor.org/stable/4320677 [accessed 5 March 2022]