Designing Driving Automation for Human Autonomy

Self-determination, the Good Life, and Social Deliberation

Santoni de Sio, Filippo; Fossa, Fabio

Citation (APA)
Santoni de Sio, F., & Fossa, F. (2023). Designing Driving Automation for Human Autonomy: Self-determination, the Good Life, and Social Deliberation. In *Studies in Applied Philosophy, Epistemology and Rational Ethics* (pp. 19-36). (Studies in Applied Philosophy, Epistemology and Rational Ethics; Vol. 67). Springer. https://doi.org/10.1007/978-3-031-39991-6_2

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Designing Driving Automation for Human Autonomy: Self-determination, the Good Life, and Social Deliberation

Filippo Santoni de Sio and **Fabio Fossa**

Abstract The present chapter analyses the complex ways in which driving automation affects human autonomy with the aim of raising awareness on the design and policy challenges that must be faced to effectively align future transportation systems to this ethical value. Building on the European report *Ethics of Connected and Automated Vehicles*, we consider three dimensions of the relation between human autonomy and driving automation: autonomy as self-determination of driving decisions; autonomy as freedom to pursue a good life through mobility; and, finally, autonomy as the capacity and opportunity to influence social deliberation concerning transportation policies and planning. In doing so, the chapter shows that delegating driving tasks to CAVs might both infringe and support user autonomy, thus calling for a reconsideration of widespread frameworks concerning the role of humans and technological systems in this domain. Moreover, it stresses the importance of promoting inclusive and participated decision-making processes on transportation policies and planning, so to avoid situations where the development and adoption of transport innovations are led by agents willing to respond only to a limited set of stakeholders' needs.

Keywords Driving automation · Human autonomy · Self-determination · Transport policy and planning · Inclusive social deliberation

F. Santoni de Sio
Section Ethics/Philosophy of Technology, Delft University of Technology, TBM Building, 2628BX Delft, The Netherlands
e-mail: f.santonidesio@tudelft.nl

F. Fossa (✉)
Department of Mechanical Engineering, Politecnico di Milano, Via Privata Giuseppe La Masa, 1, 20156 Milan, Italy
e-mail: fabio.fossa@polimi.it

1 Introduction

The purpose of the present chapter is to analyse the complex ways in which driving automation affects the ethical value of human autonomy. In doing so, it aims at raising awareness on the design and policy challenges that must be faced to effectively align future transportation systems to relevant claims grounded on this value.

The intersection of human autonomy and technological automation is notoriously controversial, with technical automation concurrently opening possibilities for supporting and restricting the enjoyment of autonomy. As such, it lies at the heart of several ethical quandaries across various AI-based applications [30]—e.g., Autonomous Weapon Systems [54] and Recommender Systems [64].

Connected and Automated Vehicles (CAVs) make no exception [6]. In the context of driving automation, threats to and opportunities for human autonomy are so numerous and deeply entangled with each other that much philosophical work is needed to clarify how this value is to be effectively pursued and promoted.

Building on the analysis proposed in the European report *Ethics of Connected and Automated Vehicles* [26], the chapter considers three dimensions of the relation between human autonomy and driving automation. First, we analyse how CAVs impact the self-determination of driving decisions and actions such as swerving, speeding, or choosing routes. Secondly, we examine how driving automation impinges on human autonomy as freedom to purse a good life through mobility. Finally, we assume a wider perspective and investigate how current narratives surrounding driving automation (see also Chap. 10) affect human autonomy intended as free and open social deliberation and policy-making.

For each case, we discuss the challenges of aligning driving automation to the demands of human autonomy. Our analysis, then, is based on the presupposition that the promotion of human autonomy as an ethical value is not to be understood only as a desired side-effect of CAV adoption. Rather, we intend to explore how driving automation can be *a means to* support this value in the practical domain of transportation. Therefore, the scope of human autonomy is neither defined by reference to the SAE levels of automation nor by taking the widespread adoption of CAVs for granted. Both levels of automation and modes of social adoption are instead assessed on the basis of their compatibility with relevant claims grounded on human autonomy.

To be clear, we do not assume that autonomy is the only or most important value to be promoted by CAVs. In line with the European report, we acknowledge that the design, development, and use of CAVs should comply with or directly promote a number of other values such as safety, justice, responsibility, and so on. In this chapter, however, our focus is exclusively on human autonomy as one important value in itself and as an example of the way in which human values should be analysed in relation to technological development.

The remainder of this chapter is structured as follows. Section 2 introduces some general remarks on the ethics of technology, thus sketching the theoretical background of our analysis. Section 3 discusses how the ethical value of human autonomy

has been brought to bear on driving automation in the context of the European report *Ethics of Connected and Automated Vehicles* [26]. Section 4 draws attention to the composite nature of the definition of autonomy there proposed. Three dimensions of human autonomy as affected by driving automation are thus specified: the self-determination of driving tasks, the freedom to pursue a good life through mobility, and the openness of social deliberation on transportation planning.

The subsequent sections discuss the alignment of driving automation to these three dimensions of human autonomy. Section 5 shows how upholding the self-determination of driving tasks suggests supporting conditional forms of driving automation. However, Sect. 6 argues that important benefits in terms of freedom to pursue a good life through mobility can only be reaped through full automation. This sets the stage for future design challenges aimed at moving beyond the SAE levels of automation framework to strike new balances between such seemingly contradictory claims. Finally, Sect. 7 tackles issues at the intersection of driving automation and transportation policy-making—the very context from which the European Commission Report and other analogous documents have originated. Finally, Sect. 8 summarises the results of the analysis and presents some conclusive remarks.

2   Ethics of Technology and Driving Automation

Ethical inquiries surrounding CAVs evidently belong to the wider field of the ethics of technology [22, 43]. As an instance of applied ethics [1], the ethics of technology is characterised by the exposure to a highly multifaceted set of challenges [17]. On a conceptual level, ethical values and notions relevant to the case at hand must be identified, determined, and discussed [2, 57]. On a practical level, the various needs, interests, expectations, perspectives, and aims of different stakeholders must be considered and agreement must be sought out of inclusive social deliberation processes [45, 47, 66]. On a technical level, both theoretical and practical insights must be operationalised so to be understandable and actionable for engineers, designers, and other professionals involved in the development of technological products, which come with their own share of complexity [29, 61]. Moreover, each level iteratively affects the others, contributing to an open (and necessarily fuzzy) process of critical appraisal, refinement, and specification [16]. At the same time, however, adherence to the criteria of applicability, effectiveness, and usability pulls for the provision of concrete recommendations, methodologies, and best practices. Such a composite blend of theoretical, practical, ethical, social, political, legal, policy, and engineering ingredients is distinctive of applied ethics efforts and lies at the basis of their intricacy.

Faced with the compound composition of socio-technical systems, ethical efforts endeavour to align the technical development, social deployment, and use of technological products to the relevant ethical values. In more concrete terms, these efforts aim at minimising harm caused by technology and putting it at the service of the social good [23].

Clearly, similar objectives can only be pursued through interdisciplinary collaboration [63]. Finding a common ground on which to build shared languages, methodologies, and practices is critical to the mission that the ethics of technology is called to carry out. Concurrently, a truly informed debate can only arise from the sharing of thorough disciplinary perspectives. Up to a point, then, disciplinary analysis must be pursued also separately, by recurring to the epistemological resources and tools of the relative research fields. For what concerns ethics, the identification, definition, and practical application of relevant values constitute perhaps the most remarkable task to attend to—and one that must also be executed with philosophical means of inquiry.

Building on these considerations, the present chapter intends to focus on human autonomy as an ethical value that requires to be adequately integrated to the driving automation domain. Driving automation promises to reorganise the way in which many transport decisions and actions are taken, shifting the boundaries of human autonomy on different levels. These reconfigurations of the balance between autonomy and automation in road transport call for further analysis and clarification. The importance of such clarification should not be underestimated. Critically examining how widespread conceptions of human autonomy apply to the case of CAVs is key to realise relevant ethical opportunities and risks. Effective design choices and policy decisions considerably depend on it.

3   Human Autonomy and Driving Automation

From an ethical perspective, the importance of aligning the design, deployment, and use of CAVs to the value of human autonomy could hardly be belittled. Surely, the philosophical status of human autonomy is controversial and intricately intertwined with thorny notions such as free will, freedom, self-determination, agential causation, subjectivity, and so on [28]. As a key ethical value, however, it enjoys widespread socio-political recognition (see, e.g., the UN Universal Declaration of Human Rights). In this sense, the value of autonomy characterises human beings as self-determining entities who deserve respect and protection. Moreover, at least within Western philosophy and culture, it buttresses fundamental components of the moral life, such as the exercise of responsible behaviour and the full enjoyment of human dignity [7, 11]. Given its relevance, human autonomy evidently qualifies as a value to be pursued through technological innovation as well.

The European approach to the ethics of driving automation confirms the latter statement. In 2020, an interdisciplinary group of fourteen experts appointed by the European Commission authored the report *Ethics of Connected and Automated Vehicles. Recommendations on Road Safety, Privacy, Fairness, Explainability and Responsibility* [26]. The document establishes an ethical framework for CAVs and offers concrete recommendations aimed at guiding stakeholders in the effort of aligning driving automation to relevant ethical values.

In close connection with the European approach to trustworthy Artificial Intelligence [25], the report starts by identifying and describing the basic normative cornerstones of the framework. Acknowledging its relevance, the authors indicate human autonomy as one of the eight overarching ethical principles for driving automation along with non-maleficence, beneficence, dignity, responsibility, justice, solidarity, and inclusive deliberation [51].

According to the report, the principle of autonomy states that human beings are to be conceived as "free moral agents" [26: 22] whose right to self-determination ought to be respected. In relation to driving automation, the principle of autonomy demands that CAVs are designed so to "protect and promote human beings' capacity to decide about their movements and, more generally, to set their own standards and ends for accommodating a variety of conceptions of a 'good life'" [26: 22]. As such, autonomy plays a crucial role in several recommendations, ranging from the protection of privacy rights and the promotion of user choice to reducing opacity and enhancing explainability (see Chap. 3).

The insistence on protecting and promoting human autonomy in the context of driving automation rests on solid grounds. Evidence from the ethics of technology and innovation clearly stresses the importance of upholding the value of human autonomy throughout the technological domain. Bypassing individual decision-making through technical means might risk leading to situations where personal decisions are taken by actors (e.g., designers, engineers, manufacturers, policy-makers) who, however, have no right nor particular competence to do so [38]. On a more social level, unduly constraining public deliberation processes might lead to harmful or short-sighted decisions [15], and in any case to limiting the people's contribution to determining their future. These states of affair are evidently incompatible with the individual and social right to self-determination and should be carefully avoided when designing, deploying, and using CAVs.

The protection of autonomy in driving automation is also critical to support other important ethical values. Consider, for instance, the case of responsibility (see Chap. 1). Qualifying human beings as free moral agents by principle means, at the same time, holding them to be responsible agents as well to the extent that they can exercise such freedom. This is a necessary presupposition to establishing who is responsible, and why, when harmful consequences follow from the design, deployment, and use of CAVs [42].

The value of human autonomy, then, is vital to the ethics of driving automation for many reasons. On the one hand, CAVs designed, deployed, and used in ways that promote autonomy will meet claims grounded on the protection of human dignity, thus supporting social acceptance and trust. On the other hand, upholding human autonomy is key to distributing responsibility in a clear and fair way while, at the same time, encouraging responsible behaviour. But how is driving automation to be concretely aligned to the demands of human autonomy?

4　One Concept, Three Dimensions

Whilst the ethical relevance of human autonomy to driving automation is evident, it is difficult to specify how the value is to be operationalised into specific socio-technical systems. Driving automation, after all, consists in the delegation of driving tasks from human agents to technical systems. Arguably, this amounts to a reduction of human (driving) autonomy. What needs to be further clarified, then, is how to automate driving functions without impacting too negatively on human autonomy. This raises thorny practical questions. What aspects of human autonomy are relevant to driving automation? Which of them should be prioritised? What model of driving automation should be promoted through design and policy decisions? What guidelines should be offered to practitioners in this sense?

To answer such questions, it is first necessary to identify what aspects of driving are relevant for moral autonomy. These aspects, in turn, would serve as tangible constraints to driving automation: CAVs should be developed and deployed in ways that allow for their exercise. In sum, defining the boundaries of human autonomy in driving automation is a necessary step towards providing effective guidelines to its operationalisation [14].

In this sense, we propose an approach that differs from the way in which the relation between human autonomy and driving automation has been commonly debated. Building on SAE levels of automation [48], human autonomy has been primarily conceived in terms of what remains to be handled once a particular driving system is deployed (see Chap. 9). In other words, human autonomy is defined by reference to which driving tasks are delegated to the driving system and which remain competence of the human driver. In what follows, we overturn the order of the terms. Instead of inducing the scope of human autonomy by moving from the scope of driving automation, we ask how different forms of driving automation can serve different ethically valuable aspects of human autonomy.

Accordingly, the ethical worthiness of CAV development, deployment, and use will be assessed on the basis of their potential in enabling human autonomy within future transportation systems. By doing so, we wish to stress the priority of the ethical value of human autonomy over choices concerning driving automation and its levels. Integrating ethics and technology, we contend, does not only require ensuring that new products comply with relevant ethical values. It also (and, perhaps, most importantly) requires determining how products should be designed, deployed, and used as means to proactively support fundamental ethical values.

Therefore, our ethical analysis of the different dimensions of autonomy cuts across and does not coincide with the engineering definition of the levels of driving automation. And it is not necessarily the case that higher levels of automation will necessarily cause lower levels of human autonomy (even though this may sometimes be the case). As it will become clearer in the remaining of the chapter, the relationship between human autonomy and driving automation is more complex than this.

The discussion of human autonomy provided in the European report represents a good starting point. Three dimensions of the relation between human autonomy and driving automation can be found through the report:

(a) autonomy as self-determination of driving decisions;
(b) autonomy as freedom to pursue a good life through mobility; and
(c) autonomy as the capacity and opportunity to influence social deliberation concerning transportation policies and planning.

Let us take a closer look to these three dimensions of human autonomy as they are affected by driving automation.

(a) Autonomy as self-determination of driving decisions concerns the exercise of individual control by human drivers/operators over decisions that pertain to driving behaviour—i.e., to the ways in which the vehicle reaches its destination from its starting point. To use traffic psychologist John Michon's control framework [34, 35; see 33], the scope of (a) includes tactical and operational levels of control over driving systems. At the tactical level, drivers exercise control over the undertaking of traffic manoeuvres such as overtaking, turning, stopping, and speeding. The operational level pertains to the physical activation of driving controls that is necessary to execute traffic manoeuvres: pushing pedals, turning the steering wheel, and so on. In the European report, this component is referred to when authors recommend to design, deploy, and use CAVs so to "protect and promote human beings' capacity to decide about their movements" [26: 22]. In this sense, respecting human autonomy would mean to let users exercise some sort of control over those system operations, if any, that impact on their moral sphere.

(b) Autonomy as freedom to pursue a good life through mobility exhibits a wider scope. It refers to the freedom of pursuing happiness and to mobility as a noteworthy enabler of what makes life worth living. In Michon's terms [34: 5–6], this aspect of human autonomy pertains to the strategic level of control, which comprises "the planning stage of a trip, incorporating the determination of trip, goal, route and vehicle choice, and evaluation of the costs and risks involved". In this sense, driving automation is not considered only as supporting a given experience—a journey between two points in space –, but rather as one important element for a good life. As stated in the European report, upholding human autonomy also means to "protect and promote human beings' capacity to (…) set their own standards and ends for accommodating a variety of conceptions of a 'good life'" [26: 22]. From this perspective, aligning driving automation to the principle of human autonomy would mean to envision CAVs as means to support the individual pursuit of personal flourishing and well-being—i.e., as means to realise significant transportation needs grounded on the value of autonomy.

(c) This third dimension of autonomy acknowledges the relevance of policy decisions concerning transport solutions and their massive impacts on people's wellbeing—including the operationalisation of (a) and (b). Indeed, policy-makers

(along with users and designers) figure as one of three stakeholder classes to which the report is mainly addressed. Moreover, inclusive deliberation—the obligation to promote a widely participated debate on the design, deployment, and use of CAVs—is endowed with the status of fundamental principle. By insisting on the necessity of inclusive deliberation processes and participation, the authors of the report stress the importance of promoting open, well-informed, unbiased, and independent policy-making when it comes to transportation planning. Shedding light on the ways in which widespread visions of driving automation as a resource to improve the ethical conditions of current transportation systems affect social decision-makers' autonomy is yet another aspect of our problem that requires to be carefully examined.

In what follows, the ethical challenges related to these three dimensions of human autonomy in driving automation are outlined. The results of the analysis can be used to ground more concrete design and policy initiatives to promote autonomy in the context of driving automation.

5 Driving Automation and the Self-determination of Driving Tasks

Let us start by considering how human autonomy as in (a) could be promoted through driving automation. In this sense, human autonomy partakes in driving automation mostly as a threatened individual value that requires to be adequately safeguarded. Particular care is required since the exercise of human autonomy at the operational and strategic levels of control is variously constrained by driving automation [67]. However, the experience of manual driving is a complex one, composed by a myriad of decisions. Some of these decisions might have a considerable impact on the moral sphere. Taking ethically relevant driving decisions out of users' hands might lead to infringements of their autonomy.

Indeed, much of what is valued in driving stems from control over our vehicles and, through it, our movements. Steering, accelerating, and braking according to personal traits, attitudes, and needs offer a tangible medium to the expression of one's own inner self [49, 60]. More importantly, driving is a moral experience. Different driving styles convey various moral values such as safety, respect for others, care for vulnerable road users, environmental friendliness, and so on. Recklessness, aggressivity, carelessness, and negligence are signs of unethical driving attitudes that all stem from how—in the Michon's terms presented above—tactical and operational control is exercised. The relation between human driving and ethical values is significantly determined by the element of control that drivers exercise on their vehicles. The delegation of driving tasks to automated systems poses the risk of bypassing human judgment in this ethically-laden domain, thus restricting the scope of human autonomy.

Of course, not every instance of driving task delegation is to be understood as a threat to autonomy. However, in some occasions driving automation might bypass users' moral judgment on matters that could be construed as lying within the purview of user autonomy. At low levels of automation, for instance, a speed control system that could not be overridden by human intervention even in case of emergency might be considered as problematic with reference to human autonomy [44, 53, 56]. At the opposite extreme, suppose that fully autonomous vehicles will be able to distribute harm during unavoidable collisions according to given ethical values (see Chaps. 7 and 8). In this case, it might be problematic in terms of human autonomy if said values were set not by passengers themselves, but rather by other stakeholders [9, 36, 37]— which would also create a potentially problematic shift of responsibility away from the drivers. Considering less futuristic scenarios involving high levels of automation, automated features concerning ethical driving behaviour—e.g., regarding the safety distance to be accorded to vulnerable road users or traffic etiquette at pedestrian crossings—might qualify as constrains to the exercise of human autonomy. Finally, relying on CAVs would restrict the possibility of taking timely decisions concerning routes, which could variously impact the execution of self-determined intentions— e.g., staying away from given roads to protect one's privacy [4].

In light of the above, it seems reasonable to conclude that the rush towards full automation—be that in the name of road safety, efficiency, sustainability—should not obfuscate the value of drivers' control over driving tasks, at least when this would serve the legitimate expression of human autonomy.

As suggested, for instance, by the Meaningful Human Control approach [52], if there are solid ethical reasons for driving decisions to be left to users, then CAVs should be designed, developed, and regulated to allow for their autonomy to be expressed [33].[1] Suppose, for instance, that users could legitimately lay claims on ethical decisions regarding CAV behaviour during unavoidable collisions, as Millar suggests [36].[2] As a consequence, they would have to be put in the condition of deciding how to shape CAV behaviour during these traffic situations. Admittedly, in a context of full automation user control over operational and tactical levels could only be accomplished indirectly—e.g., through the setting of user preferences. It is at least uncertain, however, whether this form of indirect control over system operations would satisfy the demands of the principle of human autonomy. More likely, (a) seems to encourage the development of automated features that leave enough space for the exercise of user autonomy—as happens in conditional automation, where control over driving tasks is shared with the system rather than fully delegated to it.

---

[1] The authors do not endorse the claim that drivers should remain in control, but rather that some human actors should. Their framework however clearly explains how ethical choices about the protection of the moral agency and responsibility of different actors should be reflected in design and policy choices.

[2] For a different opinion, see e.g. [20].

6 Driving Automation and the Good Life

The claim according to which human autonomy would be better served by conditional automation—or other design choices that would protect the drivers' control over driving operations—may be challenged, however, if human autonomy is intended as in (b): the freedom to pursue a good life. Higher levels of driving automation can arguably have beneficial impacts on human autonomy as the freedom to pursue a good life. At least two opportunities stand out: inclusive transportation and the improvement of travellers' overall well-being. Both importantly enable the possibility to fulfil personal needs and desires, thus increasing the chances to live a good life.

First, CAVs could massively enhance the autonomy of social categories that are currently excluded from manual driving because of physical and cognitive impairments. Independent access to transportation is critical for pursuing personal well-being and leading a satisfying social life [13, 46, 55]. Private and independent mobility, moreover, is commonly considered as the best solution in cases where various impairments stemming from conditions such as disability and old age make it harder to rely on public transport [5, 10]. However, current technical and societal limitations exclude many individuals from getting a driver license or being able to drive and, thus, from manual driving. As a result, those who need assistance to enjoy the freedom of road transport experience an unjustly limited and expensive access to it. In turn, this reduces opportunities for social interaction, political involvement, and employment, while hindering the recourse to important services such as education and medical care. Hence, failures in providing inclusive transport options heavily affect the personal well-being of already vulnerable social groups (e.g., [3]). If designed, funded, introduced, and regulated with this goal in mind, driving automation might play a considerable role in changing this situation. Since driving tasks would be automated, physical and cognitive impairments would no longer constitute an insurmountable barrier to the autonomous use of road vehicles [31].

Second, driving automation could support the self-determined pursuit of a good life by creating the conditions for a better travel experience. First of all, CAVs would allow users to reclaim travel time. Freed from the burden of driving themselves, CAV users would be able to employ travel time as they prefer. In addition, robust driving assistance might substantially reduce negative externalities associated with the driving experience such as stress, fatigue, and road rage, thus enhancing psychological well-being and contributing to the enjoyment of a good life [55]. Finally, autonomous decision-making on matters that importantly impact on individual well-being would also be supported. For instance, decisions about where to live would be less constrained by work locations and other circumstantial factors [24].

In both cases above, human autonomy benefits entirely depend on higher levels of automation. As a matter of fact, individuals excluded from manual driving would be poor candidates for shared control as well [19]. Similarly, full delegation is necessary for CAV users to freely engage in other, more satisfying activities. As the accidents

involving Tesla and Uber automated vehicles demonstrate [40, 41], current shared control design does not allow for user distraction—even when they include tools to prompt users to remain focused on supervision tasks [21]. In order to support autonomy as freedom to pursue one's own conception of a good life, then, it seems as if human intervention and supervision should be increasingly automated away.

The conclusion just reached could, however, be challenged both by remaining within the scope of (b) and by comparing it to claims based on (a). For what concerns the first case, it has been argued that full driving automation would hinder access to significant sources of well-being for those who attach value to manual driving. Indeed, manual driving is valued by many as a source of pleasure and a powerful medium to enjoy meaningful aspects of human autonomy such as freedom and self-determination [39, 58]. Moreover, manual driving might be valued by those who distrust technological innovation in general or the peculiar configuration of CAVs—which, e.g., might expose not just CAV users, but all road users to privacy risks they might be unwilling to face [18, 27].

On the other hand, it is evident that (b) and (a) lead to conflicting conclusions. Benefits related to inclusivity and user well-being are strongly dependent on high or full levels of automation. However, these forms of driving automation would leave little space to the exercise of users' autonomy at tactical and operational levels of control. If considerations on (a) and (b) are taken together, support for both partial and full automation can be advocated. As a result, compliance with the ethical principle of human autonomy would steer in directions that are difficult to harmonise. Within the framework of SAE levels of automation, then, it is hard to realise how protecting the exercise of user self-determination over driving decisions can go hand in hand with protecting the right to a self-determined good life pursued through mobility.[3]

This ambiguity, that stems from the complexity of the notion of autonomy and competing claims on driving automation, poses a challenge to designers and engineers. Perhaps, moving beyond the SAE framework to explore new forms of cooperation between human beings and driving systems might help open other avenues to design for autonomy in this field [59]. For now, the previous analysis has shown that design for autonomy requires a clear understanding of the nuanced ways in which different dimensions of this value can be served by driving automation. Moreover, it requires a thorough study of the tensions that might arise when design solutions aimed at supporting these nuances are introduced.

Fine-grained knowledge on the intersection between user autonomy and driving automation is necessary to strike acceptable trade-offs in this sense. In the spirit of [63], it may be even said that real innovative solutions in driving automation should strive to loosen the existing tension between these different interpretations of the value of autonomy, by designing new socio-technical systems that allow to promote (more of) both of them. Moreover, such knowledge would provide a solid ground for political and policy decision-making, which arguably represents the most

---

[3] As discussed in [26], Chap. 2, it is also important to consider the extent to which data-based transportation systems may create new forms of discrimination or domination, as is already the case with many other digital, data-based services.

adequate context where choices concerning similar trade-offs are to be made. Here lies the relevance to our analysis of the third dimension of autonomy, (c) autonomy as the capacity and opportunity of people to influence social deliberation concerning transportation policies and planning. The next session is dedicated to it.

7 Driving Automation and Independent Policy-Making

Consider again the so-called moral dilemmas with self-driving cars—exceptional hypothetical situations in which a fully automated vehicle faces an unavoidable crash where the only options open are (seriously) harming one group of agents or another. In Sect. 5 we have presented the claim put forward by some ethicists and policy-makers according to which such choices should be left—directly or indirectly—to the vehicles' drivers, to protect their moral autonomy. However, one could also take a different ethical approach. These choices, it may be argued, will crucially affect the distribution of risks and harms in the public space, and should therefore be regulated according to some principles or norms of social justice, rather than just being left to the free interaction of the individual choices of drivers [26, Chap. 1]. It would arguably not be fair to have certain categories of people—say, elderly persons, cyclists, people from minority ethnic groups etc.—being systematically penalised in the crashes' harm allocation. This may happen due to the result of the aggregation of preferences of drivers[4] (see Chap. 8), and/or given the impossibility of other road users to influence the decision-making process and to have their interests and rights protected [32].

At a more procedural level, this seems to suggest that the ultimate authority to decide on such cases should be given not as much to individual users but rather to some collective, typically parliamentary or governmental, agency entrusted with the legitimate power to represent the interests of all the people [50]. This shows the importance of the third dimension of autonomy (c): the possibility for, ideally, all people to (indirectly) influence the decision-making process that will determine the design, development, and regulation of CAVs, and to make sure that their interests and rights are sufficiently reflected in the process. It is, as it were, autonomy as democratic freedom or power.

This form of autonomy is very important beyond the relatively marginal issue of decision-making in dilemmatic crash-avoidance scenarios. Consider, for instance, the issue of inclusivity discussed in the previous section. Driving automation has the potential to give more people the possibility to independently use motor vehicles, by providing them with specific forms of driving assistance, or even full automated driving capabilities. This may dramatically enhance their capacity and opportunity to freely pursue a good and meaningful life, that is autonomy in our second sense (b).

---

[4] By the way, some of these preferences would even be openly discriminatory and their implementation in the driving system therefore just illegal in many jurisdictions (imagine, e.g., a vehicle programmed to systematically hit women or people of colour).

However, to realise this potential it is crucial that the development and introduction of automated driving happen with these stakeholders and their interests and values in mind. To the extent to which the development of driving automation technology is guided, for instance, by big (luxury) car manufacturers and tech companies, with the goal of embedding these technological features in vehicles designed for wealthy, able-bodied, neurotypical drivers, then the promise of inclusiveness and autonomy-enhancing is not likely to be realised. If, on the contrary, the interests and values of a broader range of stakeholders are seriously, not only in words, considered and embedded in the development process of the technology, since its early stages, then more well-being and autonomy for more people can be expected.

This leads to an even more general point. In his 1980 book *The Social Control of Technology* [8], the social scientist David Collingridge reflected on the ways in which technological development can help achieve broad societal goals, as opposed to just deliver new technical functionalities. Collingridge was quite sceptical that this could be achieved by leaving technological development only in the hand of scientists and engineers. He wrote:

> Ask technologists to build gadgets which explode with enormous power or to get men to the moon, and success can be expected, given sufficient resources, enthusiasm and organization. But ask them to get food for the poor; to develop transport systems for the journeys which people want; to provide machines which will work efficiently without alienating the men who work them; to provide security from war, liberation from mental stress, or anything else where the technological hardware can fulfil its function only through interaction with people and their society, and success is far from guaranteed. [8: 15].

After Collingridge, a full thread of academic and policy studies has emerged, under the name of *Responsible Innovation*. According to this approach, to ensure that the technological process is societally beneficial, at least three forms of "responsiveness" should be pursued: (a) between innovators and stakeholders [45, 65]; (b) between the innovation process, the changing information environment, and changing values (adaptivity; cfr. [12, 45, 62]); (c) among stakeholders [65]. Different aspects are emphasised by different authors in this tradition. But the general point is that for technology to be responsive to a broader range of interests and needs of a broader range of stakeholders, stakeholders' interests and values should somehow be firmly embedded at the different levels of the technological process.

This is needed both for epistemological and political reasons. On an epistemological level, it is to be noted that the knowledge required to develop a societally beneficial technology is distributed across society. "Experts", be them engineers or policy-makers, cannot be expected to possess all the relevant knowledge. Their decisions should therefore be supported and guided by the knowledge of other stakeholders as well. From a political viewpoint, it is important to stress that bearers of the relevant knowledge and interests should have sufficient power to counteract the dominant approaches enforced by more powerful actors in the technology and policy game. We do not want to fall back in "the notorious example" of road building, where citizens were asked what the best route for the new road would be, but not whether the road was needed at all [8: 191]. The same goes with driving automation: we should ensure that we give different stakeholders the possibility and power to contribute to

the question, not only of how much automation they want, but also which automation, for whom, to achieve which values and goals—and possibly, sometimes, if, considering all possible other technological and policy options to promote their interests and needs, they need any driving automation at all.

8 Conclusions

Human autonomy and technological automation intertwine in complex and often multifaceted ways. If the ethical values of freedom and self-determination are to be preserved and supported at both individual and societal levels, it is necessary to examine the ways in which our interactions with technological systems impact the scope and possibilities of our agency.

The present chapter has proposed an analysis of this problem as it applies to the domain of driving automation. It has argued that delegating driving tasks to CAVs might both infringe and support user autonomy, thus calling for a reconsideration of widespread frameworks concerning the roles of humans and driving systems. Moreover, it has highlighted the importance of designing CAVs by keeping in mind their impacts on human autonomy—along with other relevant ethical values as well. Finally, it has focused the attention on decision-making processes concerning transportation policy and identified possible limitations of democratic freedom, such as in cases where the development and adoption of transport innovations are led by agents who respond only to a limited set of stakeholders' values, needs, and interests. In light of what has been discussed, we conclude that further clarifications of how human autonomy is constrained, served, and transformed by driving automation are essential to guide the development of the technology towards ethically acceptable directions and, in so doing, contribute to improving the conditions of future transportation systems.

References

1. Beauchamp, T.: The nature of applied ethics. In: Frey, R.G., Wellman, C.H. (eds.) A Companion to Applied Ethics, pp. 1–16. Blackwell, Malden (2015)
2. Bednar, K., Spiekermann, S.: Eliciting values for technology design with moral philosophy: an empirical exploration of effects and shortcomings. Sci. Technol. Hum. Values (2022). https://doi.org/10.1177/01622439221122595
3. Bennett, R., Vijaygopal, R., Kottasz, R.: Willingness of people who are blind to accept autonomous vehicles: an empirical investigation. Transp. Res. F: Traffic Psychol. Behav. **69**, 13–27 (2019). https://doi.org/10.1016/j.trf.2019.12.012
4. Boeglin, J.A.: The costs of self-driving cars: reconciling freedom and privacy with tort liability in autonomous vehicle regulation. Yale J. Law Technol. **17**(4), 171–203 (2015)
5. Bradshaw-Martin, H., Easton, C.: Autonomous or 'driverless' cars and disability: a legal and ethical analysis. Euro. J. Curr. Legal Issues **20**(3), 1–17 (2014). http://webjcli.org/index.php/webjcli/rt/printerFriendly/344/471

6. Chiodo, S.: Human autonomy, technological automation (and reverse). AI Soc. **37**, 39–48 (2022). https://doi.org/10.1007/s00146-021-01149-5
7. Christman, J.: Autonomy in moral and political philosophy. In: Zalta, E. (ed.) The Stanford Encyclopedia of Philosophy (Fall 2020 Edition). https://plato.stanford.edu/archives/fall2020/entries/autonomy-moral/
8. Collingridge, D.: The Social Control of Technology. St. Martin's Press, New York (1980)
9. Contissa, G., Lagioia, F., Sartor, G.: The Ethical Knob: ethically-customisable automated vehicles and the law. Artif. Intell. Law **25**, 365–378 (2017). https://doi.org/10.1007/s10506-017-9211-z
10. Crayton, T.J., Meier, B.M.: Autonomous vehicles: developing a public health research agenda to frame the future of transportation policy. J. Transp. Health **6**, 245–252 (2017). https://doi.org/10.1016/j.jth.2017.04.004
11. Darwall, S.: The value of autonomy and autonomy of the will. Ethics **116**(2), 263–284 (2006)
12. de Saille, S.: Innovating innovation policy: the emergence of 'Responsible Research and Innovation'. J. Respons. Innov. **2**(2), 152–168 (2015). https://doi.org/10.1080/23299460.2015.1045280
13. Epting, S.: Automated vehicles and transportation justice. Philos. Technol. **32**, 389–403 (2019). https://doi.org/10.1007/s13347-018-0307-5
14. Fossa, F., et al.: Operationalizing the ethics of connected and automated vehicles: an engineering perspective. Int. J. Technoethics **13**(1), 1–20 (2022). https://doi.org/10.4018/IJT.291553
15. Foxon, T.J.: Technological lock-in and the role of innovation. In: Atkinson, G., Dietz, S., Neumayer, E., Agarwala, M. (eds.) Handbook of Sustainable Development. Edward Elgar Publishing, Cheltenham (2014). https://doi.org/10.4337/9781782544708.00031
16. Friedman, B., Kahn, P.H.: Human values, ethics, and design. In: Sears, A., Jacko, J.A. (eds.) The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, II, pp. 1177–1201. L. Erlbaum Associates Inc., New York (2008)
17. Friedman, B., Kahn, P.H., Borning, A.: Value sensitive design and information systems. In: Himma, K.E., Tavani, H.T. (eds.) The Handbook of Information and Computer Ethics, pp. 69–101. Wiley, Hoboken (2008)
18. Glancy, D.J.: Privacy in autonomous vehicles. Santa Clara Law Rev. **52**(4), 3, 1171–1239 (2012). https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?article=2728&context=lawreview&httpsredir=1&referer=
19. Goggin, G.: Disability, connected cars, and communication. Int. J. Commun. **13**, 2748–2773 (2019). https://ijoc.org/index.php/ijoc/article/view/9021
20. Gogoll, J., Müller, J.F.: Autonomous cars: in favor of a mandatory ethics setting. Sci. Eng. Ethics **23**, 681–700 (2017). https://doi.org/10.1007/s11948-016-9806-x
21. Hancock, P.A.: Some pitfalls in the promises of automated and autonomous vehicles. Ergonomics **62**(4), 479–495 (2019). https://doi.org/10.1080/00140139.2018.1498136
22. Hansson, S.O. (ed.): The Ethics of Technology. Methods and Approaches. Rowman and Littlefield, London (2017)
23. Harris, C.E.: Engineering ethics: from preventive ethics to aspirational ethics. In: Michelfelder, D., McCarthy, N., Goldberg, D. (eds.) Philosophy and Engineering: Reflections on Practice, Principles and Process. Philosophy of Engineering and Technology, vol. 15. Springer, Dordrecht (2013). https://doi.org/10.1007/978-94-007-7762-0_14
24. Heinrichs, D.: Autonomous driving and urban land use. In: Maurer, M., Gerdes, J., Lenz, B., Winner, H. (eds.) Autonomous Driving, pp. 213–231. Springer, Berlin-Heidelberg (2016). https://doi.org/10.1007/978-3-662-48847-8_11
25. High-Level Expert Group on Artificial Intelligence: Ethics Guidelines for Trustworthy AI (2019). https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai. Accessed 3 May 2022
26. Horizon Commission expert group to advise on specific ethical issues raised by driverless mobility (E03659). Ethics of Connected and Automated Vehicles: recommendations on road safety, privacy, fairness, explainability and responsibility (2020). https://op.europa.eu/en/publication-detail/-/publication/89624e2c-f98c-11ea-b44f-01aa75ed71a1/language-en. Accessed 3 May 2022

27. Jannusch, T., David-Spickermann, F., Shannon, D., Ressel, J., Völler, M., Murphy, F., Furxhi, I., Cunneen, M., Mullins, M.: Surveillance and privacy—Beyond the panopticon. An exploration of 720-degree observation in level 3 and 4 vehicle automation. Technol. Soc. **66**, 101667 (2021). https://doi.org/10.1016/j.techsoc.2021.101667

28. Kane, R. (ed.): The Oxford Handbook of Free Will, II Oxford University Press, Oxford (2011)

29. Kroes, P., van de Poel, I.: Design for Values and the definition, specification, and operationalization of values. In: van den Hoven, J., Vermaas, P., van de Poel, I. (eds.) Handbook of Ethics, Values, and Technological Design. Springer, Dordrecht (2015). https://doi.org/10.1007/978-94-007-6970-0_11

30. Laitinen, A., Sahlgren, O.: AI systems and respect for human autonomy. Front. Artif. Intell. **4**, 705164, 1–14 (2021). https://doi.org/10.3389/frai.2021.705164

31. Lim, H.S.M., Taeihagh, A.: Autonomous vehicles for smart and sustainable cities: an in-depth exploration of privacy and cybersecurity implications. Energies **11**(5), 1062 (2018). https://doi.org/10.3390/en11051062

32. Liu, H.Y.: Irresponsibilities, inequalities and injustice for autonomous vehicles. Ethics Inf. Technol. **19**, 193–207 (2017). https://doi.org/10.1007/s10676-017-9436-2

33. Mecacci, G., Santoni de Sio, F.: Meaningful human control as reason-responsiveness: the case of dual-mode vehicles. Ethics Inf. Technol. **22**, 103–115 (2020). https://doi.org/10.1007/s10676-019-09519-w

34. Michon, J.A: Dealing with danger. Technical Report nr. VK 79-01. Traffic Research Centre, University of Groningen (1979). https://jamichon.nl/jam_writings/1979_dealing_with_danger.pdf

35. Michon, J.A.: Human Behavior and Traffic Safety. Springer, Boston (1985). https://doi.org/10.1007/978-1-4613-2173-6.

36. Millar, J.: An ethics evaluation tool for automating ethical decision-making in robots and self-driving cars. Appl. Artif. Intell. **30**(8), 787–809 (2016). https://doi.org/10.1080/08839514.2016.1229919

37. Millar, J.: Ethics settings for autonomous vehicles. In: Lin, P., Abney, K., Jenkins, R. (eds.), Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence, pp. 20–34. Oxford University Press, Oxford (2017). https://doi.org/10.1093/oso/9780190652951.003.0002

38. Mitcham, C.: Justifying public participation in technical decision making. IEEE Technol. Soc. Mag. **16**(1), 40–46 (1997). https://doi.org/10.1109/44.584650

39. Müller, J.F., Gogoll, J.: Should manual driving be (eventually) outlawed? Sci. Eng. Ethics **26**, 1549–1567 (2020). https://doi.org/10.1007/s11948-020-00190-9

40. National Highway Traffic Safety Administration: Special Crash Investigations: On-Site Automated Driver Assistance System Crash Investigation of the 2015 Tesla Model S 70D. DOT HS812 481. National Highway Traffic Safety Administration, Washington, DC (2018)

41. National Transportation Safety Board: Collision Between Vehicle Controlled by Developmental Automated Driving System and Pedestrian, Tempe, Arizona, March 18, 2018. Accident Report NTSB/HAR-19/03PB2019-101402 (2019). https://www.ntsb.gov/investigations/accidentreports/reports/har1903.pdf

42. Nyholm, S.: The ethics of crashes with self-driving cars: a roadmap II. Philos. Compass **13**(7), e12506 (2018). https://doi.org/10.1111/phc3.12506

43. Nyholm, S.: This is Technology Ethics: an Introduction. Wiley Blackwell, New York (2022)

44. Nyholm, S., Smids, J.: Automated cars meet human drivers: responsible human-robot coordination and the ethics of mixed traffic. Ethics Inf. Technol. **22**, 335–344 (2020). https://doi.org/10.1007/s10676-018-9445-9

45. Owen, R., Stilgoe, J., Macnaghten, P., Gorman, M., Fisher, E., Guston, D.: A framework for responsible innovation. In: Owen, R., Bessant, J., Heintz, M. (eds.) Responsible Innovation: managing the Responsible Emergence of Science and Innovation in Society, pp. 27–50. Wiley, Chichester (2013). https://doi.org/10.1002/9781118551424.ch2

46. Parviainen, J.: Kinetic values, mobility (in)equalities, and ageing in smart urban environments. Ethical Theory Moral. Pract. **24**, 1139–1153 (2021). https://doi.org/10.1007/s10677-021-10249-6

47. Pinch, T.J., Bijker, W.E.: The social construction of facts and artefacts: or how the sociology of science and the sociology of technology might benefit each other. Soc. Stud. Sci. **14**(3), 399–441 (1984)
48. SAE International: J3016. (R) Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. Superseding J3016 JUN2018 (2021)
49. Sagberg, F., Selpi, G., Bianchi Piccinini, F., Engström, J.: A review of research on driving styles and road safety. Hum. Factors **57**(7), 1248–1275 (2015). https://doi.org/10.1177/001872 0815591313
50. Santoni de Sio, F.: Killing by autonomous vehicles and the legal doctrine of necessity. Ethical Theory Moral Pract. **20**, 411–429 (2017). https://doi.org/10.1007/s10677-017-9780-7
51. Santoni de Sio, F.: The European commission report on ethics of connected and automated vehicles and the future of ethics of transportation. Ethics Inf. Technol. **23**, 713–726 (2021). https://doi.org/10.1007/s10676-021-09609-8
52. Santoni De Sio, F., van den Hoven, J.: Meaningful human control over autonomous systems: a philosophical account. Front. Robot. AI **5**(15), 1–14 (2018). https://doi.org/10.3389/frobt. 2018.00015
53. Schoonmaker, J.: Proactive privacy for a driverless age. Inf. Commun. Technol. Law **25**, 1–33 (2016). https://doi.org/10.1080/13600834.2016.1184456
54. Sharkey, A.: Autonomous weapons systems, killer robots and human dignity. Ethics Inf. Technol. **21**, 75–87 (2019). https://doi.org/10.1007/s10676-018-9494-0
55. Singleton, P.A., De Vos, J., Heinen, E., Pudāne, B.: Chapter Seven—Potential health and well-being implications of autonomous vehicles. In: Milakis, D., Thomopoulos, N., van Wee, B. (eds.) Advances in Transport Policy and Planning, vol. 5, pp. 163–190. Academic Press, Cambridge (2020). https://doi.org/10.1016/bs.atpp.2020.02.002
56. Smids, J.: The moral case for intelligent speed adaptation. J. Appl. Philos. **35**, 205–221 (2018). https://doi.org/10.1111/japp.12168
57. Smits, M., Ludden, G.D.S., Peters, R., Bredie, S.J.H., van Goor, H., Verbeek, P.P.: Values that matter: a new method to design and assess moral mediation of technology. Des. Issues **38**, 39–54 (2022)
58. Sparrow, R., Howard, M.: When human beings are like drunk robots: driverless vehicles, ethics, and the future of transport. Transp. Res. Part C: Emerg. Technol. **80**, 206–215 (2017). https:// doi.org/10.1016/j.trc.2017.04.014
59. Stayton, E., Stilgoe, J.: It's time to rethink levels of automation for self-driving vehicles. IEEE Technol. Soc. Mag. **39**(3), 13–19 (2020). https://doi.org/10.1109/MTS.2020.3012315
60. Taubman Ben Ari, O., Yehiel, D.: Driving styles and their associations with personality and motivation. Accid. Anal. Prevent. **45**, 416–422 (2012). https://doi.org/10.1016/j.aap.2011. 08.007
61. van de Poel, I.: Translating values into design requirements. In: Michelfelder, D., McCarthy, N., Goldberg, D. (eds.) Philosophy and Engineering: reflections on Practice, Principles and Process. Philosophy of Engineering and Technology, vol. 15. Springer, Dordrecht (2013). https://doi. org/10.1007/978-94-007-7762-0_20
62. van den Hoven, J.: Value sensitive design and responsible innovation. In: Owen, R., Bessant, J., Heintz, M. (eds.) Responsible Innovation: managing the Responsible Emergence of Science and Innovation in Society, pp. 75–83. Wiley, Chichester (2013). https://doi.org/10.1002/978 1118551424.ch4
63. van den Hoven, J., Vermaas, P., van de Poel, I. (eds.).: Handbook of Ethics, Values, and Technological Design. Springer, Dordrecht (2015). https://link.springer.com/referencework/https:// doi.org/10.1007/978-94-007-6970-0
64. Varshney, L.R.: Respect for Human Autonomy in Recommender Systems (2020). https://arxiv. org/abs/2009.02603v1

65. von Schomberg, R.: Towards responsible research and innovation in the information and communication technologies and security technologies fields (2011). https://ssrn.com/abstract=2436399
66. Winner, L.: Do artefacts have politics? Daedalus **109**(1), 121–136 (1980)
67. Xu, W.: From automation to autonomy and autonomous vehicles. Interactions, 49–53 (2021). http://dl.acm.org/ft_gateway.cfm?id=3434580&type=pdf&dwn=1