Introduction

Jeroen van den Hoven

The last two decades of the twentieth century and the first decade of the twenty-first century have seen a significant increase in attention to ethical problems in all areas of public life. The UN Sustainable Development Goals provide a detailed list of some of the biggest challenges the world is confronted with in the near future: climate change, renewable energy, safety and security, affordable health care for all, safe food production and access to clean drinking water. These problems are hard to solve for a number of reasons. Their formulation is contested, proposed solutions are controversial, the indicators and metrics of success are under constant review and discussion, the subject matter is complex, the parts of the world they are concerned with are dynamic and in constant flux, the technology involved is often new, and the institutional context is populated by a great number of stakeholders with different views and values. Financial crises are examples of these 'wicked problems'. They are exacerbated by systemic failures of different sorts: cascading effects, political instability triggered by fake news on social media, natural disasters aggravated by corruption. Our own responses to problems often also make things worse in unpredictable ways. Our world has become a highly interconnected place and it is perhaps more accurate to say for example that the banks in the financial crisis were 'too connected to fail', rather than 'too big to fail'. In a world with these characteristics it is very difficult to figure out where responsibilities lie, which regulations, governance schemes and institutional frameworks are adequate, which mix of economic incentives, education policies and innovation strategies will help to bring about the desired collective behaviour. Those who attempt to solve the problems soon feel overburdened by a range of conflicting value commitments and multiple constraints.

Applied Turn in Ethics

Moral philosophers have started to study these problems more systematically around the mid-twentieth century with the aim to have impact on the real world of politics and public policy and to improve the quality of decision making and public policy regarding the Grand Challenges of our time. It seems that

moral philosophy has thereby entered a phase of recovery in the sense that John Dewey (1998) referred to when he remarked: 'Philosophy recovers itself when it ceases to be a device for dealing with the problems of philosophers and becomes a method, cultivated by philosophers, for dealing with the problems of man' (p. 8).

After a linguistic turn in the first half of the twentieth century, analytical philosophy witnessed a distinctively 'applied turn' in the early sixties. Professional philosophers in the USA and in the UK started to contribute to public debates about the moral aspects of societal problems: the war in Vietnam, civil rights, animal rights, abortion and euthanasia, environmental protection and the use of nuclear power and nuclear weapons. Philosophical papers on these subjects found their ways into the philosophy curriculum and anthologies in practical ethics. The renewed attention to practical ethics was occasioned by prominent societal debates, appreciation of social criticism and political progressive forces, but it was also facilitated by the retreat of non-cognitivism and emotivism in meta-ethics. Morality was no longer seen – as it had been before – as concerned with emotions. As Passmore (1985) observed in his supplement to *A Hundred Years of Philosophy*, 'The Applied Turn' (see DeMarco and Fox, 1986) was marked out by the fact that philosophers suddenly 'were no longer prepared to grant that to participate in fundamental social controversies was to step outside philosophy' (p. 3). This had been the position for example of influential philosophers such as Bertrand Russell and Alfred Ayer. Russell did not see his contributions to public and political discussions about war, peace and international politics as contributions he made *as a philosopher*. The programmatic statement of the influential journal *Philosophy and Public Affairs* – founded in 1971 – clearly marks a paradigm shift toward an emphasis on the practical relevance of moral philosophy: 'PAPA is founded in the belief that a philosophical examination of these issues can contribute to their clarification and to their resolution'. In the second half of the twentieth century and in the first decades of the twenty-first century the prominence of practical ethics and applied philosophy has only expanded. Bioethics, medical ethics, business ethics, police ethics, computer ethics, sports ethics, academic ethics, to name just a few, have become highly institutionalized and specialized journals and dedicated conferences have become associated with them.

The Design Turn in Applied Ethics

Applied Philosophers no longer think that prescriptions deduced from normative ethical theories are particularly helpful when it comes to solving the problems that the world is confronted with in the twenty-first century. Too often practical syllogism-style contributions have remained without impact in the real

world. It is one thing to say that the categorical imperative or the principle of utility has certain practical implications in a particular case, it is quite another thing to realize those conclusions in a meaningful way. Reiterating the Kantian view on lying does not bring honesty and integrity in the world of business and finance one step closer.

The work of John Rawls (1999) gave an impulse to further developments in applied ethics by foregrounding a design stance in ethics and political theory. The principles of justice in his proposed theory of justice can be construed as requirements for social and institutional design. Construing ethical considerations as requirements for design helps to fill the gap between moral theory and change in the real world. Rawls's student Thomas Pogge was even more explicit in his attention to realization by design of moral theory. In his study of Rawls's theory of justice entitled *Realizing Rawls*, one of his central concerns was how institutions and social mechanisms can be designed in greater detail in order to implement abstract principles of justice. Thinking along similar lines, Russell Hardin (1996) has argued that 'morality must be built in' (p. 128). As the present collection of contributions illustrates, an increasing number of authors working in normative ethical theory are starting to think in terms of design in order to make applied ethics more useful and relevant to solving our real-world problems. We refer to this notable shift in attention in applied ethics as *the design turn in applied ethics*.

Russell Hardin (1996), Dennis Thompson (2005), Richard Thaler and Cass Sunstein (2008), Robert Goodin (1996), Lawrence Lessig (1999) and Thomas Pogge (see this volume) all discuss and suggest designs of economic, institutional, organizational or technical conditions required for the realization of normative ethical views. The work of these authors is primarily focused on *institutional* design, but *The Design Turn* also brings into view the design of *socio-technical systems*, *technological artefacts* and *infrastructures*. Design-oriented applied ethics asks how we can design the ensemble of systems, institutions, incentive structures, economic mechanisms, infrastructures and IT applications in the world of health care, finance, business, energy, safety so as to enable users to do what they ought to do.

Design products, both institutional and technological, are all around us, they constitute our life-world. They include highway systems, energy systems, social media, rail transportation systems, cars, tax systems, laws and regulatory frameworks, schools, universities, hospitals, corporations, business processes, clinical protocols, computers and smart phones, household appliances, streets and squares, public buildings and smart houses. More and more people live in these extensively designed and engineered environments. Design is a ubiquitous feature of our contemporary world. These institutional and technological contrivances empower us in a myriad of ways. They bring buyers and

sellers together (markets), they educate us (schools), they provide us with transport (motorcars), allow us to communicate over long distances instantaneously (mobile phones and Internet), and manage smart infrastructures.

However, these very same humanly designed institutions and technologies also take away options and constrain our freedom to act. Motorways enable us to travel speedily by car from A to B, but motorways also require us to drive on one side of the road and along a predetermined pathway between A and B. The institution of the police enables us to conduct our daily business safely and without unwarranted interference; but equally the police constrain our activities in the public domain to do as we see fit. Computer algorithms and computers allow us to process large volumes of data, but they also determine how we perceive the world when we run our computer models. Social media allow us to connect to friends, but they also structure our interactions in particular ways. Smart infrastructures allow us to manage public utilities efficiently and sustainably, but they also require personal data of users.

So design products are ubiquitous and are in part constitutive of our human environments. They empower us *and* constrain us in particular ways. Indeed, quite frequently institutions and technologies have dual uses, ethically speaking: they can be used for good, but they can also be used for evil. Consider techniques for genetic engineering or the modern corporation. As such, design products have an important *ethical* and *normative* dimension.

Sometimes a moral end or moral feature is *designed into* an institution or technology; sometimes a morally desirable outcome is the fortuitous, but unintended, consequence of an institutional arrangement or technological invention. A paradigm case of a morally objectionable end being *designed into an artefact* was Robert Moses's use of low hanging overpasses in New York, which were intended to prevent the buses from the poor black neighbourhoods being routed to the beaches near New York – a favorited destination of white middle class families (Winner, 1980). This notion of *designing in ethics* is complex and elusive. In what sense, for example, are the ethical ends or moral aspects of hospitals or pacemakers a constitutive feature of these entities? And in what sense is there a designer of them? They are complex systems that have come about as a result of the contributions of many agents with different expertise over a course of time. Institutions, for example, are typically the result of the actions of multiple agents interacting over generations (Miller, 2003). This also applies to infrastructures and large socio-technical systems, like grids or smart cities. A further issue pertains to the conceptual, practical and moral limits on design. Here it is especially important to note the moral limits on notions of designing or redesigning human beings *per se*. Some recent discussions on human enhancement, for example, offer a brave new world of future 'super-humans' with greatly enhanced sensory apparatus and possessed of prodigious memories and calculative capacities (Agar, 2004; Savulescu and Bostrom, 2009). The

underlying conception of a human being informing some of the more glib pronouncements made in this connection, seems that of a computerized robot – a conception that is, to say the least, contestable (Sandel, 2007). Artificial Intelligence and machine learning opens up a completely new field of design issues. How can systems with these AI components be designed in a transparent and accountable way? Not only should moral problems thus be studied in light of possible design solutions to them, but existing designs and proposed design solutions should also be evaluated from a moral point of view.

Hitherto, the study of ethics has tended to focus on the analysis of pre-existing concepts and the adjudication between predetermined options – such adjudication often being understood in terms of a process of weighing up the moral considerations regarding these given options.

Notwithstanding the applicability of this general approach in a variety of situations, we want to point in this volume to a somewhat different approach to the study of contemporary ethical issues – an approach which highlights the possibility of design for values, value sensitive design or 'designing in ethics' (van den Hoven, 1997; Whitbeck, 1996, 2011; Friedman et al., 2006; van den Hoven et al., 2015). This approach emphasizes the need to create or expose additional design options and, thereby, reconfigure or re-design the choice set.

This *design turn in applied ethics* can be seen as the third and the most recent phase in the development of contemporary ethics. After an almost exclusive focus on meta-ethics in the beginning of the twentieth century, there was an applied turn in the second half of the twentieth century. However, this applied turn largely consisted in the application of existing theory to given practical problems, and involved primarily a static process of analysis and adjudication in light of pre-existing options.

By contrast, we are insisting on the need to identify and articulate a rather different category of questions in applied ethics. It is not simply a matter of the familiar type of question: '*Given* this situation and *given* the fact that there are two options A and B open to the individual chooser (e.g. the chooser in a prisoner's dilemma scenario and the chooser in a trolley scenario) what should she do?' For we also need to focus on the following type of question: 'How can we re-design the option set such that there is a morally preferable option C – an option that is additional to the currently existing options A and B, and meets all our ethical requirements, so that we do not need to choose between A and B?'

Consider a situation in which a large number of moral agents each performs an action of a given type, e.g. each dumps their waste into a nearby river that provides fresh water for drinking, washing and so on. Each action considered on its own does no harm. Cumulatively, however, the actions cause great harm – specifically, the river is seriously polluted. On a static conception of the matter

it is not obvious whether, and on what basis, to ascribe moral responsibility and to punish. On the one hand, given the relatively benign intention and tiny causal contribution of each agent, it is not obvious that anyone is morally responsible; so presumably no one ought to be held liable and punished. On the other hand, there is the fact of great harm having been caused by these agents, indeed foreseeably caused.

However, there is a third option, C, that could be designed and implemented – indeed, has been designed and implemented. This third option consists in ascribing, i.e. creating, an institutional responsibility on the part of each agent not to pollute, and doing so by means of an enforceable regulation or law to that effect. This is an institutional responsibility that is designed to serve an important moral purpose; moreover, it is an institutional duty that is now also a moral duty. Thus, a citizen or other institutional role occupant who fails in his or her institutional duty by making what is only a tiny causal contribution to some morally untoward outcome, nevertheless, might be held *fully institutionally* responsible for the outcome and, as a result, *fully morally* responsible for the outcome; or at least he might only enjoy a minor diminution in moral responsibility, and certainly not a moral diminution commensurate with his very minor share of (so to speak) natural responsibility for the outcome. These alternatives by design could also consist of technical innovations such as affordable waste water treatment supported by institutional design solutions. A similar type of solutions has been brought about in the context of the choice between (national) security and privacy. Here so-called privacy enhancing (see discussion in van den Hoven et al., 2014) and 'privacy respecting' technologies allow us to benefit from information technology without having to suffer violations of privacy. Clean technology and renewable energy technology allow us reap the socio-economic benefits of economic growth, without suffering the detrimental impact of growth on the environment and climate. These solutions do not come about simply as a result of moral reasoning and analysis in the context of traditional concepts and pre-existing options of on the one hand limiting growth or on the other hand accepting climate change and pollution.

Morally motivated design solutions to real-world problems are discussed throughout this volume and a number of straightforward examples of technological innovation are provided in the first chapter. Design thinking in ethics does not make moral analysis superfluous or obsolete, it amplifies the scope and efficacy of practical ethics.

Globalization, Institutions and Collective Responsibility

If designing-in ethics is an important new methodological orientation for applied ethics, globalization is providing many of the practical ethical problems upon which to deploy this methodology.

On the one hand, global communication, international financial transactions and transnational movement of people, give rise to new ethical problems, or new versions of old problems. On the other hand, the solutions to these problems require transnational and transcultural cooperation. Now more than ever there is a need to shoulder *collective* moral responsibility for ethical problems (Scheffler, 1995; Miller, 2006). Corruption, genocide, economic and financial crises, terrorism, climate change and environmental degradation, pandemics require collective action on the part of governments and other organizations. The global nature of the problems, the high degrees of communication and information exchange, connectivity, interdependence, international mobility, means that ethics, including issues of social justice, can no longer be seen as relativized to the individual nation state (Pogge, 2002). What is happening to people in other parts of the world can no longer be seen to be exclusively their business, but not ours. Self-evidently, ethics now has an important global focus. Furthermore, moral problems should be studied in light of possible design solutions to them and those designs should also be evaluated from a moral point of view.

The contributions in this volume all exemplify one or more features of the 'design turn in applied ethics' as discussed above. Van den Hoven argues that applied ethics needs to be reconstructed in order to adequately address the problems of the twenty-first century. It needs to become more focused on real-world problems in all of their complexities and it needs to readjust some of the basic conceptual frameworks with which it addresses these problems. More specifically he argues that attention should be paid to *design agents* and *design histories*, while pointing out that our conception of agency should be expanded so as to comprise the *design contexts* in which our own agency and actions of others is structured and shaped.

Dennis Thompson elucidates how two general approaches (the individualist and collectivist) are both inadequate with regard to assigning responsibility when confronted with such problems of many hands. He therefore proposes to 'use the results of the investigations into responsibility for past outcomes as a guide to changing an organization to clarify individual responsibility for future outcomes and future oversight. Specifying the responsibility for monitoring the reforms in the structures of responsibility – design responsibility – is often neglected but is no less important than the responsibility for outcomes'.

Ibo van de Poel argues in his chapter that dealing with moral dilemmas requires combining both analytic and synthetic strategies. In morally problematic situations in which we become confronted with conflicting values or principles, we could adopt analytical strategies such as revision of our moral obligations to deal with such a dilemma. However, it could be simultaneously possible to develop new options for actions that avoid the conflict at hand. The latter is a synthetic strategy of innovation. Design is an activity that combines

both synthetic and analytic strategies and is therefore, as argued by van de Poel, especially apt as a strategy to deal with moral dilemmas. He illustrates his argumentation further by focusing on the formulation of design requirements, which play a role both in the synthesis of new options and in choice.

Situations where individuals have to contribute to joint efforts or share scare resources are ubiquitous. In the absence of proper mechanisms of cooperation, pursuits to maximize individual success could thereby tend to create a tragedy of the commons (such as over-fishing or the destruction of our environment).

Peter-Paul Verbeek argues that technology is morally charged and we should embrace this condition to actively moralize technology. The notion that technology could influence and guide human behaviour requires the integration of ethical analysis in the design processes. In his chapter, Verbeek reflects on two problems in the design of 'moral technologies', which are identified as the designer fallacy and the democratic challenge of behaviour-guiding design.

Thomas Pogge in his contribution discusses the fact that current rules governing pharmaceutical innovation cause advanced medicine to be priced beyond the reach of the poor and steer medical research away from diseases concentrated among them. He envisions how, through the establishment of the Health Impact Fund (HIF), financial support will be provided to innovators in order to increase worldwide access to affordable medicine. Furthermore, the HIF will direct pharmaceutical innovation toward the most serious diseases, including those concentrated among the poor. And it incentivizes innovators to promote the optimal use of their HIF-registered medicines. Magnifying one another's effects, these advances would engender large global health gains.

In his contribution, Tom Sorell analyses how micro-lending could serve as an instrument for undoing financial exclusion and reducing global poverty. On the basis of the constraints such micro-lending (microfinance) institutions face, he tries to identify the preferred design for these same institutions in order for lenders to realize sustainable and effective 'financial inclusion' of the poorest to the largest extent. Seumas Miller's chapter on Australia's superannuation system discusses the ethical issues that arise in relation to the designing of an institutional arrangement for the financial needs of retirees. In the course of doing so, he outlines his normative theoretical account of socio-economic institutions and applies it to this system.

In his chapter, Andrew Alexandra questions whether the institution of war itself is the most effective means of achieving the goals which are the reasons for its existence. The design of pacifist institutions are brought forth as a potential alternative to the institution of war in achieving desired goals without the use of violence.

Seumas Miller in his essay on designing an anti-corruption system for police organizations, outlines the nature and causes of police corruption prior to describing the main features of an anti-corruption system. He completes the

essay with a discussion of two specific challenges faced by designers of anti-corruption systems for police, namely, police culture (specifically, the so-called blue wall of silence) and the use and misuse of integrity testing police officers (so-called sting operations).

In their chapter, Toonen and Doorn develop a tentative set of recommendations for the good governance of Common Pool Resources (CPR). On the basis of a case study on the governance of the Dutch Wadden Sea, they 'show that good governance for the Commons can be conceived as a design problem with one overarching criterion, viz. legitimacy'.

Christian Illies and Nick Ray argue in their chapter that aesthetic considerations should constitute the primary aim of architecture. Rather than just being focused on functionality, architectural design should be directed at the realization of accessible beauty in the construction of buildings. Thus, when it comes to reconciling conflicting demands, beauty is the key. Illies and Ray describe that this over-riding duty obliges architects to 'aim toward a satisfaction of needs that entails the moral imagination of a better world'.

REFERENCES

Agar, N. (2004). *Liberal Eugenics: In Defence of Human Enhancement*. Oxford: Blackwell Publishing.

DeMarco, J. P., and Fox, R. M. (1986). *New Directions in Ethics: The Challenge of Applied Ethics*. New York: Routledge.

Dewey, J. (1998). The need for a recovery of philosophy. (1917). *The Essential Dewey*, 1, 46–70.

Friedman, B., Kahn, P. H., Jr., and Borning, A. (2006). Value sensitive design and information systems. In P. Zhang and D. Galletta (eds), *Human-Computer Interaction in Management Information Systems: Foundations*, 348–372. London: M. E. Sharpe.

Goodin, R. E. (1996). Institutions and their design. In R. E. Goodin (ed.), *The Theory of Institutional Design*, 1–53. Cambridge: Cambridge University Press.

Hardin, R. (1996). Institutional morality. In R. E. Goodin (ed.), *The Theory of Institutional Design*, 126–153. Cambridge: Cambridge University Press.

Lessig, L. (1999). *Code and Other Laws of Cyberspace*. New York: Basic Books.

Miller, S. (2003). Individual autonomy and sociality. In F. Schmitt (ed.), *Socialising Metaphysics: Nature of Social Reality*, 269–300. Lanham, MD: Rowman and Littlefield.

(2006). Collective moral responsibility: An individualist account. *Midwest Studies in Philosophy*, 30(1), 176–193.

Passmore, J. A. (1985). *Recent Philosophers*. London: Duckworth.

Pogge, T. (2002). *World Poverty and Human Rights*. Cambridge: Polity Press.

Rawls, J. (1999). *A Theory of Justice*. Rev. ed. Cambridge, MA: Harvard University Press.

Sandel, M. (2007). *The Case against Perfection: Ethics in the Age of Genetic Engineering*. Cambridge, MA: Harvard University Press.

Savulescu, J., and Bostrom, N., eds. (2009). *Human Enhancement*. Oxford: Oxford University Press.

Scheffler, S. (1995). Individual responsibility in a global age. *Social Philosophy and Policy*, 12(1), 219–236.

Thaler, R. H., and Sunstein, C. R. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.

Thompson, D. (2005). *Restoring Responsibility: Ethics in Government, Business, and Healthcare*. Cambridge: Cambridge University Press.

van den Hoven, M. J. (1997). Computer ethics and moral methodology. *Metaphilosophy*, 28(3), 234–249.

van den Hoven, M. J., Blaauw, M., Pieters, W., and Warnier, M. (2014). Privacy and information technology. In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Winter 2014 ed. http://plato.stanford.edu/archives/win2014/entries/it-privacy/.

van den Hoven, M. J., Vermaas, P. E., and van de Poel, I. (2015). *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains*. Houten, Netherlands: Springer.

Verbeek, P. P. (2005). *What Things Do: Philosophical Reflections on Technology, Agency and Design*. Philadelphia: Pennsylvania State University Press.

Whitbeck, C. (1996). Ethics as design: Doing justice to moral problems. *The Hastings Center Report*, 26(3), 9–16.

(2011). *Ethics in Engineering Practice and Research*. Cambridge: Cambridge University Press.

Winner, L. (1980). Do artifacts have politics? *Daedalus*, 109(1), 121–136.