Delft University of Technology

Values and Design

van de Poel, I.R.

Publication date
2021

Document Version
Final published version

Published in
The Routledge Handbook of the Philosophy of Engineering

Citation (APA)
van de Poel, I. R. (2021). Values and Design. In D. P. Michelfelder, & N. Doorn (Eds.), *The Routledge Handbook of the Philosophy of Engineering* (pp. 300-314). Routledge - Taylor & Francis Group.

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

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' - Taverne project

https://www.openaccess.nl/en/you-share-we-take-care

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.

22
VALUES AND DESIGN

Ibo Van de Poel

22.1. Introduction

It is increasingly recognized that technological design expresses certain values and that it is therefore desirable to explicitly address values during the design process of new technology. One approach to do so is value sensitive design (VSD). The VSD approach was originally developed by Batya Friedman, and it has been further developed by her and her colleagues during the last decades (Friedman and Kahn 2003; Friedman et al. 2006; Davis and Nathan 2015). Also a range of similar approaches has been articulated and developed, like Values in Design (VID), Values at Play, and Design for Values (DfV) (e.g. van den Hoven et al. 2015; Flanagan et al. 2008). In addition, there are a range of approaches that focus more specifically on one value (or a limited number of values); for example, design for sustainability, inclusive design, and privacy-by-design (Keates and Clarkson 2003; Bhamra and Lofthouse 2007; Birkeland 2002).

The aim of this contribution is to discuss a number of general and more fundamental issues that should be addressed if one wants to systematically address values in the design of new technology. The focus is therefore not on one specific approach (like VSD) but on a number of more general conceptual and methodological issues that are crucial for systematically integrating values into technical design. These issues are: (1) what values are, and how to identify the values that should be integrated into a particular design; (2) the conceptualization, operationalization and specification of values which is required to make values operational in the design process; (3) issues of conflicting values and possible ways to address them and (4) the possibility of value change and options to address it during design. Before discussing these four issues, I start with a brief historical overview.

22.2. Historical Overview

It has long been recognized in the philosophy of technology that technology is value-laden (e.g. Heidegger 1962; Ellul 1964). The first to give a systematic account of how specific artefacts (rather than technology as such) and design choices may be value-laden was probably Langdon Winner (1980). This value-ladenness of technological artefacts is now more or less generally accepted (see also Chapter 53, 'Engineering and Contemporary Continental Philosophy of Technology' by Diane P. Michelfelder, this volume), although there is not yet agreement on how it can best be understood, as different philosophical accounts have been proposed (Kroes and Verbeek 2014).

Initial work on VSD did not originate from the philosophy of technology, but from the area of information systems in combination with social science. Since the early 1990s, Batya Friedman and colleagues have been working on proposing and further developing the VSD approach (Davis and Nathan 2015). Although this approach was initially developed for the design of information systems, it is more widely applicable.

VSD takes what Friedman et al. (2006) call an interactional stance with respect to how values can get embedded in technology. With this, they mean that which values are realized will depend on both how a technology has been designed and how it is implemented and operated; both technical and social factors thus play a role, and it is the interplay between them that makes technology value-laden.

A further core element of the VSD approach is a tripartite methodology consisting of an iterative combination of empirical, conceptual and technical investigations (Friedman and Kahn 2003). Empirical investigations, among others, involve empirically investigating relevant actors and their values; conceptual investigations involve a further conceptualization of these values and possible value trade-offs; technical investigations involve implementing values in technical design choices but also identifying ethical and value issues on basis of existing designs and technologies.

In addition to the tripartite methodology, Flanagan et al. (2008) propose three phases or activities for integrating values into design:

1. *Discovery*, resulting in a list of relevant values for the design project at hand.
2. *Translation*, i.e. "the activity of embodying or expressing . . . values in system design" (Flanagan et al. 2008: 338).
3. *Verification*, the assessment—e.g. through simulation, tests or user questionnaires—of whether the design indeed has implemented the values that were aimed at.

A wide range of more specific methods for VSD is now available. Some of these have been specifically developed for VSD, other are more generally applicable social science methods that are also useful for VSD. Friedman et al. (2017) list the following 14 methods: (1) direct and indirect stakeholder analysis; (2) value source analysis; (3) co-evolution of technology and social structure; (4) value scenarios; (5) value sketches; (6) value-oriented semi-structured interview; (7) scalable information dimensions; (8) value-oriented coding manual; (9) value-oriented mock-ups, prototypes, and field deployments; (10) ethnographically informed inquiry regarding values and technology; (11) model for informed consent online; (12) value dams and flows; (13) value sensitive action-reflection model and (14) envisioning cards.

A good overview of the developments and discussions in VSD is presented in Davis and Nathan (2015), while van den Hoven et al. (2015) provide an overview of the broader field of design and values. As said in the Introduction, my aim for the remainder of this chapter is to elaborate a number of general themes that are relevant to any attempt to systematically include values into design.

22.3. What Are Values?

In the literature on VSD, values are usually defined as referring "to what a person or a group of people consider important in life" (Friedman et al. 2006: 349).[1] Although this is helpful as a first-order characterization of values, it also raises questions. For example, is everything that a person or group concerns important in life also a value? What if people consider it important to treat people of a different race or ethnicity differently? It seems, then, that there are things people find important but which do not amount to values. For such reasons, it has been argued that values should be distinguished from mere preferences or interests of people, and that we should distinguish between what people find important in life and what they have good *reasons* for to find important in life.

In the psychological literature, somewhat more precise definitions of values can be found. Rokeach (1973), for example, defines values as "enduring beliefs that a specific mode of conduct is personally or socially preferable to an opposite or converse mode of conduct or end-state of existence" (Rokeach 1973: 5). According to Schwartz and Bilsky (1987) "[v]alues are (a) concepts or beliefs, (b) about desirable end states or behaviors, (c) that transcend specific situations, (d) guide selection or evaluation of behavior and events, and (e) are ordered by relative importance" (Schwartz and Bilsky 1987: 551).

There are no clear-cut definitions in the philosophical literature, but usually values are associated with what is *good* (Dancy 1993; Raz 1999). The notion of value can then refer to what is good (metaphysics), or what we believe (epistemology) or express (semantics) to be good (Hirose and Olson 2015). Different values may be understood as different varieties of goodness, i.e. as different ways in which a state of affairs (or an object) can be good (von Wright 1963). General as this philosophical characterization of values is, it nevertheless helps to distinguish values from a number of related concepts:

1. From *attitudes*. The difference is that values are more stable and enduring and are also more general and abstract than attitudes (Rokeach 1973; Hechter 1993);
2. From *interests and preferences*. While some authors have associated values with interests (Schwartz and Bilsky 1987), values are not just beliefs or expressions about what is in the interest of an agent, but rather expressions or beliefs about what is *good* either for an agent or more generally.
3. From *norms*. Norms are more specific than values and contain prescriptions (including recommendations, obligations, prohibitions, restrictions) for action often based on sanctions (Hechter 1993). They belong to the deontic domain of normativity, while values belong to the evaluative domain (Dancy 1993; Raz 1999). Values help to evaluate certain states of affairs in terms of goodness.
4. From *goals and aims*. Goals and aims are more specific and concrete than values. Values help to evaluate states of affairs and may thus suggest certain goals or aims to strive for, but they are not themselves goals.

Even if we have a general idea what values are, the question for VSD remains: how are we to determine what values are relevant and important for the design of a specific technology? Friedman et al. (2006) suggest a list of values—like human welfare, privacy, accountability, freedom from bias, trust and environmental sustainability—that are often implicated in system design, but they stress that this list is not exhaustive. Others have questioned the usefulness of such lists, as designers should be open to other relevant values and values should always be elicited in the context of specific design projects and require the input from users and stakeholders (Borning and Muller 2012; Dantec et al. 2009).

The question of how to determine the values that are to be taken into account in the design of technology has two aspects. One is what might be called the discovery of values. Discovery is aimed at finding the potentially relevant values for the design of a technology. A large range of methods and approaches is available for value discovery, including checklists, interviews, surveys, focus groups, (value) scenarios, discourse analysis, hermeneutics, ethnography, participatory research, living labs and Q-methodology. The other aspect is how to decide which of the values resulting from value discovery *should* be taken into account in a specific design process. This is a normative question that extends beyond listing the (potentially) relevant values. This may require a substantive ethical theory (Manders-Huits 2011) or at least a criterion to distinguish 'mere' values from 'real' values or, more important, from less important values. With respect to the latter, philosophers sometimes distinguish between intrinsic (or final) values, that are good for their own sake, and instrumental (or extrinsic) value, that are good insofar as they contribute to other intrinsically good values (Zimmerman 2004; Moore 1922), although such distinctions have also been criticized (Dewey 1922).

The normative question of what values to take into account also has a political dimension: Who gets to decide what values are to be included? Most of the VSD literature (and other literature on values and design) advocates an approach in which not just the designers, but also direct stakeholders (users) and indirect stakeholders (others affected by the technology) have a say in answering this question. However, this may itself raise further normative (and political) questions. For example, in the development of new encryption software, criminals are, strictly speaking, also an indirect stakeholder; but should their values also be taken in account?

22.4. The Conceptualization, Operationalization and Specification of Values

Identifying relevant values and selecting those on which the design process should focus is a first step in any systematic attempt to take values into account in the design of new technology. However, for these values to impact the actual design process, they need to be made more specific and operational so that they can guide design choices. I propose here to distinguish between three main activities that are important in order to make values bear on the actual design process of new technologies:

- *Conceptualization of values* is "the providing of a definition, analysis or description of a value that clarifies its meaning and often its applicability" (Van de Poel 2013: 261). Conceptualizations typically provide an explanation or reasons why a value is good or positive, as well as an interpretation of how to understand the value. Conceptualization of values is largely a philosophical activity that is independent from the specific context where the value is applied, although some conceptualizations may be more appropriate or adequate for certain contexts than for others. For example, privacy may be conceptualized in terms of the "right to be left alone" (Warren and Brandeis 1890), but also in terms of control over what information about one self is shared with others (cf. Koops et al. 2017). Although both conceptualizations are general, in the design of a specific information system, the second one may be more relevant than the first.
- *Operationalization of values* refers to the process of making values measurable, so that it becomes possible to measure to what degree a state of affairs or a certain design realizes (or meets) a certain value. Value measurement has similarities but also distinct differences with the measurement of physical concepts like temperature (Kroes and Van de Poel 2015). Unlike physical concepts, values are normative (and not descriptive) in nature. This makes it more difficult to measure values objectively, but it does not rule out the operationalization of values, nor does it make value measurements arbitrary, as I will discuss in greater detail.
- *Specification of values* refers to the translation of values into more specific norms and design requirements that can guide the design process of new technology. In specification, contextual information is added, which makes it more specific what it means to strive for (or respect or meet) a certain value in a certain context. For example, in the context of designing a chemical plant, safety may be specified in terms of minimizing explosion risks, as well as in terms of providing containment, so that in case of an accident, it is less likely that hazardous materials leak into the environment of the plant.

Next, I will further discuss and illustrate these three activities with the help of a number of examples.

22.4.1 Conceptualization

Concerning conceptualization, it is often the case that different philosophical conceptions of a value are available in the literature. Take for example well-being. In the philosophical literature, three main theories of well-being have been articulated (Crisp 2013):

- Hedonism, which understands human well-being in terms of pleasurable experiences.
- Desire satisfaction accounts which conceptualize well-being in terms of the fulfilment of people's desires.

- Objective list accounts which understand well-being in terms of a list of general prudential values (such as accomplishment, autonomy, liberty, friendship, understanding and enjoyment; cf. Griffin 1986).

Each of these conceptualizations suggests another approach to how to design for well-being. On a hedonistic account of well-being, design should be aimed at creating products that give users and other stakeholders pleasurable experiences. A number of design approaches are aimed at just that (e.g. Koskinen et al. 2003; Desmet and Hekkert 2007). On a desire-satisfaction account, people's (actual) desires should be the starting point of design, and a method like Quality Function Deployment (QFD) (e.g. Akao 1990; Hauser and Clausing 1988) that starts from user demands and translates these into engineering characteristics of the product-to-be-designed could be used. For objective list accounts, there are not many systematic methods, but a design approach is suggested in Van de Poel (2012).

The fact that there are various, often competing, philosophical conceptions of many values (like well-being, privacy, sustainability, justice) raises the question which conception practitioners who want to design for a certain value should adopt. There is no straightforward answer to this question, but it is important to be aware that two types of considerations are relevant in answering this question in a specific case. One type consists of more general philosophical considerations about the adequacy of certain conceptualizations. The mentioned philosophical theories of well-being have each been criticized, which in response has resulted in more sophisticated accounts that try to meet some of the raised criticism, and one might have good philosophical reasons to prefer one account over the others. In addition to such philosophical considerations, there are also more practical considerations. Not every technology will (potentially) affect the same dimensions of well-being, and depending on how a technology may or may not impact humans, there may be good reasons to focus on a specific notion of well-being in the design of a specific technology. For example, if a technology mainly affects human experiences and to a lesser extent other aspects of well-being (according to the other two accounts), it may be justified to focus on a more hedonistic notion of well-being, while in other cases such a notion may be too narrow, for example because a technology also affects friendship or personal relations.

22.4.2 Operationalization

Operationalization of values is the process of making values measurable. In some cases, it might be possible to directly measure values in a subjective way. We can, for example, ask users (or other stakeholders) to rank design options in terms of a value like safety; the result would be a subjective measurement of the value of safety.

In many cases, however, it is hard to measure values directly in a reliable way. In such cases, it may be more appropriate to first further operationalize the value at stake. Operationalization of values, in many cases, can be modelled as a two-stage process, namely: first, a translation of the value into a number of evaluation criteria that are important for judging whether the value is met (in the specific context considered); and second, the association of attributes, that can be more directly and objectively measured, with these criteria (Kroes and Van de Poel 2015). The attributes can, then, be seen as proxies for the attainment of the criteria. Figure 22.1 gives an impression of what the operationalization of values might look like in a concrete case (adopted from Kroes and Van de Poel 2015).

The case at hand is the design of alternative coolants for refrigerators. Values that are relevant in this case include environmental sustainability, safety and health. These values can be associated with certain evaluation criteria and attributes. For the attributes mentioned in the figure, there are measurement procedures and measurement scales that are, for example, laid down in relevant technical codes and standards.

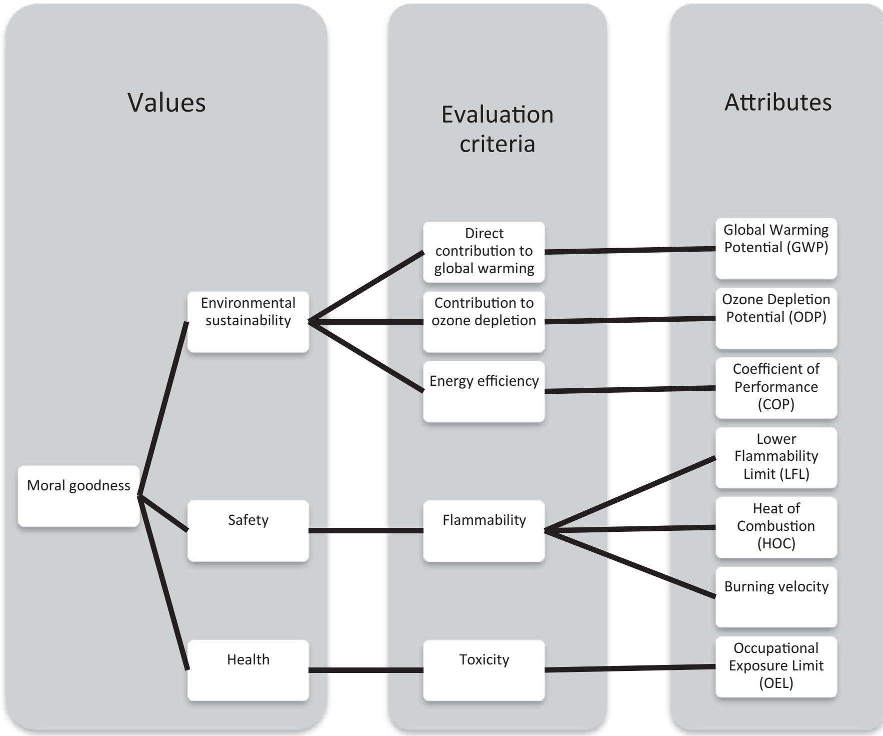

Figure 22.1 The operationalization of the values environmental sustainability, safety and health for the case of refrigerant coolants

Source: From Kroes and Van de Poel (2015).

A number of observations are in place with respect to this process of operationalization of values. First, it is context-specific. The operationalizations shown in Figure 22.1 are sensible for the case of coolants, but they would probably make no sense in the context of the design of another technology (like, for example, cars, where the same values would be relevant). Second, it involves so-called second-order value judgements, i.e. value judgements not about the technology designed (first-order value judgements), but about how to interpret and operationalize the relevant values in the specific context. Third, often more than one attribute will be associated with one value (as suggested in the figure). This means that these attributes are not only just proxies, but also that there is the question how to aggregate the different attributes into one measure for the relevant value. Obviously, any attempt to answer this aggregation question would involve a range of further value judgements (see also Chapter 7, 'Engineering Design and the Quest for Optimality' by Maarten Franssen, this volume).

Although the operationalization of values involves value judgements, and often different operationalizations are justifiable, it does not follow that operationalization is arbitrary. Not any operationalization will do, and in practice there are often operationalizations available for various technological domains, for which there is a degree of social consensus; for example, those laid down in technical codes and standards (see also Chapter 42, 'Standards in Engineering' by Paul B. Thompson, this volume). Of course, this does not necessarily imply that these operationalizations are also philosophically and morally justified, or are beyond debate. There may sometimes be good reasons to change an operationalization of a value in the light of new technological or social developments.

22.4.3 Specification

A third relevant activity to make values bear on design processes is specification. This process involves translating values into more concrete norms and design requirements that can guide the design of new technology. A useful tool here might the so-called values hierarchy (Van de Poel 2013). Figure 22.2 shows an example of a values hierarchy.

A values hierarchy consists of three main layers, i.e. values, norms and design requirements; each layer may in turn have a number of sublayers. A values hierarchy is held together by two relations. Top-down, this is the relation of specification, i.e. lower-level elements in the values hierarchy are specifications of higher-level elements. Specification is a non-deductive reasoning process, in which contextual information is taken into account in specifying what a higher level element (like a value) implies in a specific context. Bottom-up, the relation between the elements can be characterized as "for the sake of", i.e. the lower-level elements are strived for the sake of higher-level elements. For example in Figure 22.2, the design requirement that biofuels should (preferably) be based on non-edible crops is desirable for the sake of avoiding an increase in food process through competition, which, in turn, is desirable for the sake of intergenerational justice.

As suggested by these two relations, a values hierarchy can be construed top-down as well as bottom-up. In the first case, we start with values that are then specified in terms of norms and design requirements. In the second case, we can start, for example, with already formulated design requirements for a design task, and ask the question for the sake of what are these design requirements strived for, and so reconstruct the underlying values. In practice, the construction of a values hierarchy will usually be an iterative process, consisting of moving bottom-up as well as top-down. For example, if values are first reconstructed bottom-up (on basis of a set of given design requirements), we can then ask the question whether the set of design requirements is indeed the most appropriate specification of these values, or should perhaps be adapted.

Since specification is non-deductive, the translation of values into design requirements involves second-order value judgements (like in the case of value operationalization), and often different specifications of a set of values may be justifiable. It should be noted that a values hierarchy as such does not answer the question how to make such value judgements, or which specification to choose from a set of justifiable but competing specifications. Nevertheless, the values hierarchy might still be useful to locate where value judgements are to be made, and to track possible disagreements about the specification of values. As such, the values hierarchy is perhaps better seen as a deliberative tool, rather than a straightforward method to specify values.

As in the case of operationalization, not any value specification will do. In fact there are a number of criteria to judge the adequacy of a specification (Van de Poel 2013). For example, one crucial issue is whether meeting the design requirements would count as an instance of meeting the underlying value.

22.4.4 Value Conflicts

In most design processes, not just one but a range of values is relevant. Oftentimes, these values will conflict. Value conflicts can take many forms, but we can distinguish two main varieties:

1. Conflicts between values, i.e. situations in which it is not possible to fully realize the various relevant values at the same time in the design of a technology. In other words, the realization of one value comes at the cost of another value. For example, safety and sustainability may be conflicting in car design, because lightweight cars usually consume less energy and are therefore more sustainable. However, in order to make a really lightweight car, it may be necessary not to include certain safety features (that often add quite a lot of weight), so that sustainability is achieved at the cost of safety. We will call this types of conflicts multiple-value conflicts.

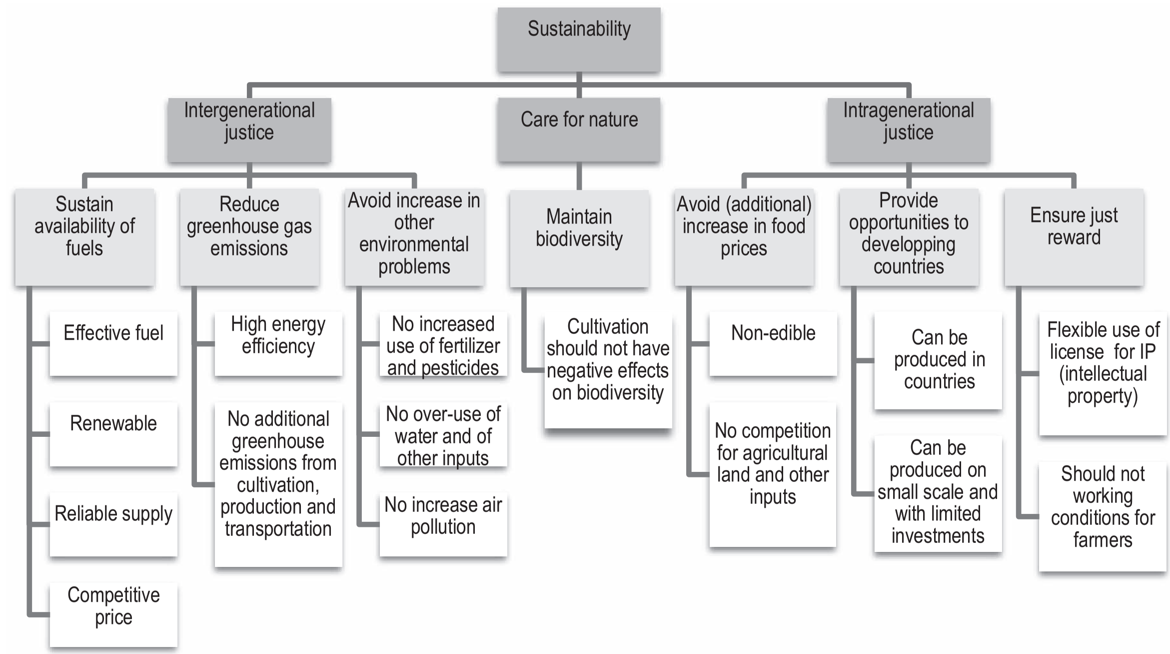

Figure 22.2 Possible values hierarchy for biofuels

Note: Values are dark gray; norms are light gray and design requirements are white

Source: From Van de Poel (2017b).

2. Conflicts between agents, i.e. situations in which agents disagree about the values. One form this may take is that agents attach different priorities (or weights) to the relevant values. In such cases, different agents may prefer differ solutions to multiple-value conflicts. However, agents may as well disagree about what values are relevant for a design, or about the conceptualization, operationalization and specification of these values. We will call this type of conflicts multiple-agent conflicts.

In reality, value conflicts will often be multiple-value as well as multiple-agent. Nevertheless, in some cases, it may be justified to treat them primarily as either multiple-value or multiple-agent conflicts. For example, if we consider the situation that one designer is designing for a multiplicity of values, it may be appropriate to treat this primarily as a multiple-value conflict. Similarly, if we take the perspective of the government, which has the responsibility to safeguard certain public values in the design of technology, it may also be proper to treat value conflicts primarily as multiple-value conflicts.[2] In other situations, it may be more appropriate to treat the conflict primarily as a multiple-agent conflict; for example, the design of large infrastructures that impact a large number of stakeholders, and for which the support of a wide diversity of agents is required for the success of the project.

There are many approaches to deal with value conflicts; overviews are presented in Van de Poel (2009, 2015). Here I will restrict myself to three main categories of strategies that can be applied to multiple-value as well as multiple-agent conflicts. These three strategies are: (1) creating a win/win option, (2) compromise and (3) finding an integrated solution.

22.4.5 Win/Win

In case of multiple-value conflicts, a win/win option is an option that improves the current situation in terms of all relevant values, and, similarly, in case of a multiple-agent issue, it is an improvement for all agents involved.[3] At the start of the design process, such win/win options may not exist, or at least not be evident. However, since design is a creative and innovative process that may result in *new options*, it may be seen as the challenge for the designer(s) to create such win/win options, in order to solve the value conflict. Van den Hoven et al. (2012) argue that designers have a moral obligation to look for such win/win options in particular if a multiple-value conflict amounts to a moral dilemma.

It should be noted that the creation of a win/win option for multiple-agent conflicts does not necessarily require that all the agents agree on the relevant values (or on the priority of these values, and their conceptualization, operationalization and specification). It 'only' requires that there is (at least) one option that, for all relevant agents, is an improvement compared to the current situation. The values, on basis of which the individual agents decide whether this is the case, may differ from agent to agent. Moreover, for each individual agent, the win/win option may still be a compromise, in the sense that it does not fully meet the values of that agent but still is considered an improvement (with respect to the current situation) for that agent. For that reason, a win/win option in terms of agents may not be a win/win option in terms of values.

Since win/win options for multiple-value conflicts do not coincide with win/win options for multiple-agent conflicts, designers who aim at creating a win/win option should first answer the question whether the value conflict at hand is best approached as a multiple-agent or a multiple-value conflict (or perhaps as both). Both approaches can be sensible, but each comes with its own pitfalls of which the designers should be aware. Generally speaking, the main pitfall of treating a value conflict solely as a multiple-value conflict is that, perhaps, an ethically acceptable solution (in terms of the relevant values) may be found, but that this solution is not accepted by the relevant agents, and therefore not executed. Conversely, if the conflict is solely treated as a multiple-agent issue, the resulting win/win solution may be accepted by all parties, but it may be ethically unacceptable, or it may neglect the legitimate interest of agents that cannot be involved (like future generations). Ideally a win/win solution, then, both implies an improvement in terms of the relevant values, as well as an improvement for all relevant agents, but obviously such ideal solutions may often be hard, if not impossible, to find even for the most creative designers.

22.4.6 Compromise

A compromise is a solution to a value conflict, in which either one (or more) of the relevant values is compromised (in case of a multiple-value conflict) or the perspective (or interest) of one (or more) of the agents is compromised (in case of a multiple-agent conflict). Because of this characteristic, compromises are often seen as less than ideal solutions to value conflicts. But, generally speaking, compromises are better than solutions, in which just one value or one agent dictates the solution, or in which a solution is reached by means of violence. So, even if compromises are less than ideal, they may the best that can be achieved.

There are many approaches to reach a compromise in a multiple-value or multiple-agent conflict. For multiple-value conflicts, compromising may often involve trading off some value(s) against one another, as is done, for example, in multiple criteria decision-making (MCDM) but also in many more specific engineering methods like Quality Function Deployment (QFD) or Pugh charts (e.g. Pugh 1991; Bogetoft and Pruzan 1991; Hauser and Clausing 1988). There are, however, also balancing methods for values that do not involve (direct) trade-offs between values (Van de Poel 2015). For multiple-agent conflicts, achieving a compromise will often involve some form of negotiation between the relevant agents.

Although compromises are sometimes unavoidable (or to be preferred to dominated or violent solutions), in general, compromises raise two further issues. One is that a specific compromise may be ethically unacceptable. In particular, negotiation in a multiple-agent conflict may result in solutions that are accepted by the most powerful agents but that either neglect important values or are achieved at the cost of less powerful agents. Another issue is that a compromise may be an unstable solution. The reason for this is that in case of a compromise, there is always at least one value or actor perspective compromised; however, the reasons for that value or agent perspective are typically not annulled (Chan and Protzen 2018). In other words, from the viewpoint of the value or agent that is compromised, there remain good reasons to reject the compromise, once there is the opportunity to do so.

22.4.7 Integrated Solution

An integrated solution will here be understood as a solution to a value conflict that does justice to all relevant values (in case of a multiple-value conflict) or to all relevant agents (in case of a multiple-agent conflict). So, contrary to a compromise, there is not a value or agent perspective compromised, and hence there is no remaining reason to undo the solution, as in the case of a compromise. At the same time, an integrated solution does not necessarily create a win/win situation, at least not in terms of the original value or agent perspectives that defined the original value conflict. The reason for this is that achieving an integrated solution involves *deliberation* about the values and the value conflict and may lead to a reformulation of the original value conflict, for example by reinterpreting or respecifying the relevant values, or because agents revise their perspective. This reformulation or reframing of the original value conflict may open the way to solutions that are acceptable in the light of all relevant values (in case of a multiple-value conflict) or to all agents (in case of a multiple-agent conflict), thus leading to a consensus solution rather than a compromise.

One possibility for the case of multiple-value conflicts is that deliberation results in a *respecification* of the relevant values in terms of design requirements. In multiple-value conflicts, often the values are not directly conflicting but, rather, the specifications of these values (in terms of design requirements) are conflicting. Because often different specifications of a value are rationally and morally defensible, it may be possible to respecify the values in such a way that the resulting design requirements are no longer conflicting, while the values are still all respected.

In case of multiple-agent conflicts, agents may need to adopt a solution-oriented attitude rather than a competing, or interest-maximizing, perspective in order to achieve an integrated solution (Chan and Protzen 2018; Carens 1979). It may particularly be helpful if agents come to a shared understanding of the value conflict, for example through reframing the problem, which may require agents to reflect on, and be willing to adapt, their own belief and value systems (Van de Poel and Zwart 2010). An integral solution, however, does not require that all agents adopt the same perspective. The result may also be what John Rawls has called an "overlapping consensus" (Rawls 2001). That is consensus (in this case about a solution to a value conflict) that each agent can morally justify in terms of his or her moral belief system, without all agents agreeing on the underlying moral belief system.

Achieving an integral solution may of course not always be possible. Nevertheless, the process of looking for one may trigger deliberation and lead to reflection on the relevant values and agent perspectives, which in itself may not only make it easier to achieve a compromise but also lead to better compromises (in terms of meeting a number of values and agent perspectives). At the same time, too much emphasis on consensus may result in an unjustified watering down of values, or may create too much pressure on agents to revise their perspective. It is important to be aware that (value) conflict can also be constructive in the sense that it may reveal hidden or marginalized values or perspectives on the issue (Rip 1987; see also Chapter 45, 'Engineering as a Political Practice' by Govert Valkenburg, this volume).

The three discussed strategies are ideal types, in the sense that real-world approaches and solutions to value conflicts will contain elements of all three approaches. Moreover, in different phases of the design process, different approaches may be more appropriate (Van de Poel 2017a). For example, once a first formulation of the design problem has settled and a value conflict surfaces, designers should first look for possible win/win solutions. This invites designers to think out of the box and to look for possibly better design options than they were initially considering. This may result in better but not yet ideal solutions, and a next phase may involve further deliberation to see whether an integrated solution can be found. When this is not possible, the designer may settle for a compromise. The point is that even if the eventual solution is a compromise it is likely to be a better compromise (in terms of values and agent perspectives respected) than if the designers had from the start aimed at a compromise.

22.5. Value Change

Studies in the philosophy of technology have shown that sometimes new technologies induce value change (Boenink et al. 2010; Swierstra 2013). The classical example is contraceptives that have triggered a change in sexual morality (Boenink et al. 2010; Swierstra 2013). New values may also emerge in order to deal with the sometimes unexpected consequences of human action and technology. An example is sustainability that may be seen as a response to the human endangerment of the natural environment.

Value changes after a technology has been designed may create a mismatch between the values for which a technical artifact or a sociotechnical system was originally developed and which are embedded in the technology, and the values we currently consider important. An example is that of energy systems, which according to many insufficiently reflect the value of sustainability, which was a less important, or even absent, value when these systems were originally developed. However, due to the large technological and institutional momentum of energy systems, these systems cannot easily be adapted to better meet the value of sustainability. This has led to discussions about the need for an energy transition, which not only is an economic, technical and institutional challenge but also requires a change in the values embedded in these systems (Demski et al. 2015).

To understand how values may change, we may understand values as emerging from earlier responses to moral problems. In line with pragmatist philosophers like Dewey, values can be seen as generalized responses to earlier moral problems. In many situations, existing values are adequate as a response to (morally) problematic situations people encounter. However, in new types of situations or due to new experiences, current values may no longer be adequate or sufficient. Such situations may require an adaptation of current values or the adoption of new values (Dewey 1922).

This suggests that values may change in response to new problematic situations or new experiences. One might, for example, think of the following possibilities:

- Technologies lead to new types of *consequences* that require new evaluative dimensions and therefore new values (e.g. privacy, sustainability) to evaluate sociotechnical systems.
- Technologies offer new *opportunities* (e.g. to protect homes against earthquakes) that lead to new moral obligations and therefore new values.
- Technologies create new *moral choices and dilemmas* where previously there were no choices (e.g. predictive genetics) that require new values.
- Technologies lead to new *experiences* (e.g. friendship online) that lead to new values or change existing values.

In the literature, there has until now been little attention for how to deal with value change, but one can think of different possible strategies. Anticipatory strategies try to predict or anticipate value changes. Currently most approaches do not try to predict value changes but rather develop, for example, technomoral scenarios to sketch different possible futures (Boenink et al. 2010). These can then be the basis for deliberation, and technologies may be so designed that they can deal with certain anticipated value changes or they may be designed in an attempt to steer value change in a certain direction. However, not all value changes can be predicted or even anticipated (Van de Poel 2017c). Not only may values change in other ways than anticipated, also how we perceive and morally evaluate value changes may be hard to anticipate beforehand. This means that anticipatory strategies alone are not enough to deal with value change. Partly we may deal with unanticipated value change by adaptive strategies. Possible adaptive strategies include the redesign of some of the technical parts of the sociotechnical system, redesign of some of the relevant institutions, adaptation of the operation of the sociotechnical system (operational strategies) or changing the behaviour of the actors in the system (behaviour strategies).

However, often it is better to be prepared for adaptation beforehand, even if the specific to-be-expected value changes cannot be anticipated. This may be done by employing a range of proactive design strategies, technical as well as institutional. One might think of design strategies or principles like robustness, flexibility, adaptability, and modularization.

22.6. Conclusions

That technology is value-laden is hardly controversial nowadays. This seems also to bring the obligation for engineers and designers to address values more systematically in the design processes of new technologies. I have discussed four (philosophical) challenges that such attempts face, i.e. (1) deciding what values to include in a design, (2) to make these values operational in the design, (3) to deal with conflicting values and (4) to deal with value change. For each, I discussed earlier work that has been done on the issue, and I presented some conceptual tools and approaches that might be useful for better dealing with these issues, theoretically as well as practically. It is, however, good to be aware that each of these issues cannot be completely solved in an abstract theoretical way. Each will require some form of moral judgement and (political) decision-making by those involved (i.e. the designers and relevant stakeholders) in the context of a concrete design process. Nevertheless, such judgement and decision-making are likely to be better if they are informed by some of the considerations offered in this contribution.

Acknowledgements

This publication is part of the project ValueChange that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No 788321.

Related Chapters

Chapter 7: Engineering Design and the Quest for Optimality (Maarten Franssen)
Chapter 20: Towards an Ontology of Innovation (Vincent Blok)
Chapter 21: Engineering Design (Peter Kroes)
Chapter 24: Human-Centred Design and Its Inherent Ethical Qualities (Marc Steen)
Chapter 25: Sustainable Design (Steven A. Moore)
Chapter 33: Values in Risk and Safety Assessment (Niklas Möller)
Chapter 34: Engineering and Sustainability: Control and Care in Unfoldings of Modernity (Andy Stirling)
Chapter 38: Health (Marianne Boenink)
Chapter 40: Ethical Considerations in Engineering (Wade L. Robison)
Chapter 43: Professional Codes of Ethics (Michael Davis)
Chapter 44: Responsibilities to the Public—Professional Engineering Societies (Joseph Herkert and Jason Borenstein)
Chapter 45: Engineering as a Political Practice (Govert Valkenburg)

Further Reading

Friedman, Batya and Kahn, Peter H. Jr. (2003). Human Values, Ethics and Design. In J. Jacko and A. Sears (eds.), *Handbook of Human-Computer Interaction*. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 1177–1201. (Introduction to the VSD approach.)

Kroes, Peter, and Verbeek, Peter-Paul (eds.) (2014). *The Moral Status of Technical Artefacts*. Dordrecht: Springer. (Overview of different approaches how to understand the embedding of values in technology.)

Van den Hoven, Jeroen, Vermaas, Pieter E. and Van de Poel, Ibo (eds.) (2015). *Handbook of Ethics and Values in Technological Design. Sources, Theory, Values and Application Domains*: Springer. (Extensive overview of different approaches of designing for values for a variety of values in a multiplicity of domains.)

Winner, Langdon. (1980). Do Artifacts Have Politics? *Daedalus* (109), 121–136. (A classical on the value-ladenness of technology.)

Notes

1. One could interpret this as a definition of value in which there can be morally good and morally bad values; values are 'simply' what people consider important in life. In what follows, I will propose a definition of value that is more closely tied to what is good.
2. It should be noted that in such cases, distributive justice may be one of the values, so that ethical considerations about the distributions of advantages and disadvantages among stakeholders are taken into account even if the conflict is approached as a multiple-value conflict.
3. More formally, we may want to employ the notion of Pareto Improvement from economics and choice theory (see also Chapter 12, 'Engineering Design and the Quest for Optimality' by Christian Dieckhoff and Armin Grunwald, this volume). For a multiple-agent setting, a Pareto Improvement is defined as an

improvement that makes at least one agent better off and the other agents not worse off. Similarly for multiple-value issues, we may define it as an improvement in which at least one value is better achieved and no values are compromised.

References

Akao, Y. (ed.) (1990). *Quality Function Deployment. Integrating Customer Requirements Into Product Design*. Cambridge, MA: Productivity Press.

Bhamra, Tracy, and Lofthouse, Vicky (2007). *Design for Sustainability: A Practical Approach*. Aldershot: Gower.

Birkeland, Janis (2002). *Design for Sustainability. A Source Book for Ecological Integrated Solutions*. London: Earthscan.

Boenink, Marianne, Swierstra, Tsjalling and Stemerding, Dirk (2010). Anticipating the Interaction Between Technology and Morality: A Scenario Study of Experimenting With Humans in Bionanotechnology. *Studies in Ethics, Law, and Technology,* 4(2), 1–38.

Bogetoft, Peter, and Pruzan, Peter (1991). *Planning With Multiple Criteria: Investigation, Communication, Choice*. Amsterdam: Elsevier.

Borning, Alan, and Muller, Michael (2012). Next Steps for Value Sensitive Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Austin, TX, USA.

Carens, Joseph H. (1979). Compromises in Politics. *Nomos,* 21, 123–141.

Chan, Jeffrey Kok Hui, and Jean-Pierre Protzen (2018). Between Conflict and Consensus: Searching for an Ethical Compromise in Planning. *Planning Theory,* 17(2), 170–189. doi:10.1177/1473095216684531.

Crisp, Roger (2013). Well-Being. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy,* Summer 2013 ed. https://plato.stanford.edu/archives/sum2018/entries/it-privacy

Dancy, Jonathan (1993). *Moral Reasons*. Oxford: Blackwell Publishers.

Dantec, Christopher A. Le, Shehan Poole, Erika and Wyche, Susan P. (2009). Values as Lived Experience: Evolving Value Sensitive Design in Support of Value Discovery. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems*, Boston, MA, USA.

Davis, Janet, and Nathan, Lisa P. (2015). Value Sensitive Design: Applications, Adaptations and Critiques. In Jeroen van den Hoven, Pieter E. Vermaas and Ibo Van de Poel (eds.), *Handbook of Ethics and Values in Technological Design*. Dordrecht: Springer, pp. 11–40.

Demski, Christina, Butler, Catherine, Parkhill, Karen A., Spence, Alexa and Pidgeon, Nick F. (2015). Public Values for Energy System Change. *Global Environmental Change,* 34, 59–69. doi:10.1016/j.gloenvcha.2015.06.014.

Desmet, Pieter, and Hekkert, Paul (2007). Framework of Product Experience. *International Journal of Design,* 1(1), 57–66.

Dewey, John (1922). *Human Nature and Conduct: an Introduction to Social Psychology*. New York: Holt.

Ellul, Jacques (1964). *The Technological Society*. Translated by John Wilkinson. New York: Alfred A. Knopf. Original edition, La Technique.

Flanagan, Mary, Howe, Daniel C. and Nissenbaum, Helen (2008). Embodying Values in Technology. Theory and Practise. In Jeroen Van den Hoven and John Weckert (eds.), *Information Technology and Moral Philosophy*. Cambridge: Cambridge University Press, pp. 322–353.

Friedman, Batya, Hendry, David G. and Borning, Alan (2017). A Survey of Value Sensitive Design Methods. *Found. Trends Hum.-Comput. Interact,* 11(2), 63–125.

Friedman, Batya and Kahn, Peter H. Jr. (2003). Human Values, Ethics and Design. In J. Jacko and A. Sears (eds.), *Handbook of Human-Computer Interaction*. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 1177–1201.

Friedman, Batya, Kahn, Peter H. Jr. and Borning, Alan (2006). Value Sensitive Design and Information Systems. In Ping Zhang and Dennis Galletta (eds.), *Human-computer Interaction in Management Information Systems: Foundations*. Armonk, NY: M.E. Sharpe, pp. 348–372.

Griffin, James (1986). *Well-being: Its Meaning, Measurement, and Moral Importance*. Oxford [Oxfordshire]; New York: Clarendon Press.

Hauser, John R. and Don Clausing (1988). The House of Quality. *Harvard Business Review,* 66(3), 63–73.

Hechter, Michael (1993). Values Research in the Social and Behavioral Sciences. In R. Michod, L. Nadel and M. Hechter (eds.), *The Origin of Values*. Berlin: Aldine de Gruyter, p. 28.

Heidegger, Martin (1962). *Die Technik und die Kehre*. 7. Aufl. ed. Pfullingen: Neske.

Hirose, Iwao, and Olson, Jonas (2015). *The Oxford Handbook of Value Theory*. New York: Oxford University Press.

Keates, Simeon, and Clarkson, John (2003). *Countering Design Exclusion: An Introduction to Inclusive Design*. London: Springer.

Koops, Bert-Jaap, Bryce Clayton Newell, Tjerk Timan, Ivan Škorvánek, Tomislav Chokrevski, and Maša Galič (2017). A Typology of Privacy. *University of Pennsylvania Journal of International Law,* 38(2), 483–575.

Koskinen, Ilpo, Battarbee, Katja and Mattelmäki, Tuuli (eds.) (2003). *Emphatic Design. User Experience in Product Design*. Helsinki: IT Press.

Kroes, Peter, and Poel, Ibo van de (2015). Design for Values and the Definition, Specification, and Operationalization of Values. In Jeroen van den Hoven, Pieter E. Vermaas and Ibo Van de Poel (eds.), *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains*. Dordrecht: Springer, pp. 151–178.

Kroes, Peter, and Verbeek, Peter-Paul (eds.) (2014). *The Moral Status of Technical Artefacts*. Dordrecht: Springer.

Manders-Huits, Noëmi (2011). What Values in Design? The Challenge of Incorporating Moral Values Into Design. *Science and Engineering Ethics,* 17(2), 271–287. doi:10.1007/s11948-010-9198-2.

Moore, G.E. (1922). *The Conception of Intrinsic Value. Philosophical Studies*. London: Kegan Paul, Trench, Trubner, pp. 253–275.

Pugh, Stuart (1991). *Total Design: Integrated Methods for Successful Product Engineering*. Wokingham, England and Reading, MA: Addison-Wesley Pub. Co.

Rawls, J. (2001). *Justice as Fairness. A Restatement*. Cambridge, MA: The Belknap Press of Harvard University Press.

Raz, Joseph (1999). *Engaging Reason. On the Theory of Value and Action*. Oxford: Oxford University Press.

Rip, A. (1987). Controversies as Informal Technology Assessment. *Knowledge: Creation, Diffusion, Utilization,* 8(2), 349–371.

Rokeach, M. (1973). *The Nature of Human Values*. New York: The Free Press.

Schwartz, Shalom H. and Wolfgang Bilsky (1987). Toward a Universal Psychological Structure of Human Values. *Journal of Personality and Social Psychology,* 53(3), 550–562. doi:10.1037/0022-3514.53.3.550.

Swierstra, Tsjalling (2013). Nanotechnology and Technomoral Change. *Etica & Politica/Ethics & Politics,* XV(1), 200–219.

Van de Poel, Ibo (2009). Values in Engineering Design. In Anthonie Meijers (ed.), *Handbook of the Philosophy of Science. Volume 9: Philosophy of Technology and Engineering Sciences*. Oxford: Elsevier, pp. 973–1006.

Van de Poel, Ibo (2012). Can We Design for Well-being? In Philip Brey, Adam Briggle and Edward Spence (eds.), *The Good Life in a Technological Age*. London: Routledge, pp. 295–306.

Van de Poel, Ibo (2013). Translating Values into Design Requirements. In D. Mitchfelder, N. McCarty and D.E. Goldberg (eds.), *Philosophy and Engineering: Reflections on Practice, Principles and Process*. Dordrecht: Springer, pp. 253–266.

Van de Poel, Ibo (2015). Conflicting Values in Design for Values. In Jeroen van den Hoven, Pieter E. Vermaas and Ibo Van de Poel (eds.), *Handbook of Ethics, Values, and Technological Design*. Dordrecht: Springer, pp. 89–116.

Van de Poel, Ibo (2017a). Dealing with Moral Dilemmas through Design. In Jeroen van den Hoven, Seumas Miller and Thomas Pogge (eds.), *Designing in Ethics*. Cambridge: Cambridge University Press, pp. 57–77.

Van de Poel, Ibo (2017b). Design for Sustainability. In David M. Kaplan (ed.), *Philosophy, Technology, and the Environment*. Cambridge, MA: MIT Press, pp. 121–142.

Van de Poel, Ibo (2017c). Moral Experimentation with New Technology. In Ibo Van de Poel, Donna C. Mehos and Lotte Asveld (eds.), *New Perspectives on Technology in Society: Experimentation Beyond the Laboratory*. London: Routledge.

Van de Poel, Ibo, and Zwart, Sjoerd D. (2010). Reflective Equilibrium in R&D Networks. *Science, Technology & Human Values,* 35(2), 174–199.

Van den Hoven, Jeroen, Lokhorst, Gert-Jan and Van de Poel, Ibo (2012). Engineering and the Problem of Moral Overload. *Science and Engineering Ethics,* 18(1), 143–155. doi:10.1007/s11948–011–9277-z.

Van den Hoven, Jeroen, Vermaas, Pieter E. and Van de Poel, Ibo (eds.) (2015). *Handbook of Ethics and Values in Technological Design. Sources, Theory, Values and Application Domains*. Dordrecht: Springer.

von Wright, Georg Henrik (1963). *The Varieties of Goodness*. London: Routledge & Kegan Paul.

Warren, Samuel D. and Brandeis, Louis D. (1890). The Right to Privacy. *Harvard Law Review,* 4(5), 193–220. doi:10.2307/1321160.

Winner, Langdon (1980). Do Artifacts Have Politics? *Daedalus* (109), 121–136.

Zimmerman, Michael J. (2004). Intrinsic vs. Extrinsic Value. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy,* Fall 2004 ed. http://plato.stanford.edu/archives/fall2004/entries/value-intrinsic-extrinsic/.