

What Values in Design? The Challenge of Incorporating Moral Values into Design

Noëmi Manders-Huits

Received: 2 April 2009 / Accepted: 18 February 2010 / Published online: 12 March 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract Recently, there is increased attention to the integration of moral values into the conception, design, and development of emerging IT. The most reviewed approach for this purpose in ethics and technology so far is Value-Sensitive Design (VSD). This article considers VSD as the prime candidate for implementing normative considerations into design. Its methodology is considered from a conceptual, analytical, normative perspective. The focus here is on the suitability of VSD for integrating moral values into the design of technologies in a way that joins in with an analytical perspective on ethics of technology. Despite its promising character, it turns out that VSD falls short in several respects: (1) VSD does not have a clear methodology for identifying stakeholders, (2) the integration of empirical methods with conceptual research within the methodology of VSD is obscure, (3) VSD runs the risk of committing the naturalistic fallacy when using empirical knowledge for implementing values in design, (4) the concept of values, as well as their realization, is left undetermined and (5) VSD lacks a complimentary or explicit ethical theory for dealing with value trade-offs. For the normative evaluation of a technology, I claim that an explicit and justified ethical starting point or principle is required. Moreover, explicit attention should be given to the value aims and assumptions of a particular design. The criteria of adequacy for such an approach or methodology follow from the evaluation of VSD as the prime candidate for implementing moral values in design.

This article originates from a co-authored paper with Paul Sollie, which was published in the proceedings of Ethicomp 2007. I have substantially rewritten the paper. Here I would like to express my gratitude to Paul Sollie for our initial collaboration.

N. Manders-Huits (✉)
Philosophy Section, Delft University of Technology, P.O.Box 5015, 2600 GA Delft, Netherlands
e-mail: N.L.J.L.Manders-Huits@tudelft.nl
URL: <http://www.ethicsandtechnology.eu>

Keywords Applied ethics · Technology · Ethics of technology · Methodology · Values · Design · Value-sensitive design · Value-conscious design · Awareness

Introduction: New Technologies Challenging Ethics

Until the late twentieth century, technology was commonly regarded as value-neutral, i.e. mainly of instrumental value to human endeavours and activities (see, e.g., Florman 1987). Recently, this conception of technology has been extensively challenged. One of the first prominent contemporary critics of the natural view of technology was Langdon Winner. He argued that technology is not value-neutral, but instead exhibits moral and political choices. He provided a compelling example of the highway overpasses built by Robert Moses, a famous and influential New York architect of urban planning (Winner 1980). In the 1930s Robert Moses was asked to design overpasses for the only highway connecting New York to Long Island, leading to Long Island Beach. Remarkably, Moses designed these bridges very low. The way they were built only allowed for cars to pass under whereas public transport (busses) passing under was rendered impossible. Supposedly, this implied Robert Moses favored highways over public transport and community needs in his design. Some have argued that the highway overpasses were intentionally designed low so as to prevent public busses, the main means of transportation for the least well-off, including racial minorities, from getting to Long Island beach (Winner 1980).

Although there has been some dispute about whether this consequence was actually intended in the design (Joerges 1999), thanks to its strong illustrative power the example of Moses' low hanging bridges is often used to exemplify the possible political and moral import of design choices. Other examples of the inherent moral and political import of technology include the discussion of online search engines (Zimmer 2008; Introna and Nissenbaum 2000), soft- and hardware codes (Lessig 1999), databases and classification systems (Bowker and Star 1999), cookies on personal computers (Elmer 2004) and many more. A related value-laden phenomenon concerns the way in which information is presented, such as the presentation of risks associated with technology (Asveld and Roeser 2009) or the degree of reliability of information on the Internet (Vedder and Wachbroit 2003). The political and moral dimensions of technology have been recognized and have become important research topics in the philosophy and ethics of technology. Proponents of this account not only weigh and assess the risks and benefits of (new) technologies, but also address the issue of how technologies impact upon our moral decisions, actions, and ultimately, upon our lives. In light of the value-ladenness of technology it is highly desirable to have at our disposal means for ethically evaluating and justifying decision-making during technology design.

The increased awareness of possible implications of our design and our design choices for the expression, support, or undermining of our (political and) moral values contributes to an understandable desire to control and influence this process: Would it be possible to design our buildings, technologies, and institutions, so that

they reflect, express, and enhance our moral and political views? Let us, for the purpose of this paper, set aside political values and focus on the practice of enhancing and incorporating the integration of moral values into the design of Information Technology.¹

In this paper I explore the criteria of adequacy for a methodology of ethics of technology while considering which normative considerations are implemented in design. I will consider and discuss Value-Sensitive Design (VSD), the most reviewed approach pertaining to values in technology design so far, as the prime candidate for such a methodology. First I describe the methodology as put forward by VSD in section 2. In the next section this approach is critically discussed from a conceptual, analytical perspective. The focus of this discussion is on the suitability of the VSD approach for integrating moral values into the design of technologies in a way that joins in with an analytical perspective on ethics of technology. From this follow the criteria of adequacy for an approach or methodology to implement moral values in design in the section “Towards Value Conscious Design.”

Value-Sensitive Design

Human beings are confronted on a daily basis with the design, development and use, of (new) technologies that in some way impact upon them. Technological innovations in the field of information and communication technology (IT) have enabled the expansion of our social and political lives and activities to a global level—one of tremendous networks such as Instant Messaging services (MSN or Trillian), Skype, YouTube, MySpace, Flickr, Hyves, Twitter, and Facebook.

One of the prominent approaches for evaluating values in technology design—and information technologies especially—is VSD.² VSD propounds a proactive approach regarding (the incorporation of) values in design, characterized by Van den Hoven as a way to ‘frontload ethics’ (2005, 2008).

VSD emerged in the 1990s proceeding from the insights of “human-computer interaction studies.”³ It started from the recognition that when designing information technologies, the predominant, traditional focus of engineers is on functionality; the primary interests of engineers concern usability, efficiency, reliability, and affordability of (new) technologies. An engineer’s principal concern is to make a technology which has the required functionality. Yet there is more to technological innovation. As mentioned above, technology should not be considered to be value-neutral, but rather to have moral (and political) impacts on humans and their environment. Decisions made during the design process have value implications

¹ The main methodological point of this paper however, holds for (values in) design in general.

² Several authors in the field of values in design draw attention to human and moral values as an integral part of the conception, design, and development of technological artifacts and systems. These include Design for Values (Camp 2003), Values at Play (Flanagan et al. 2005; Flanagan et al. 2008), Value-Sensitive Design (Friedman 2004; Friedman et al. 2006), Values in Engineering Design (Van de Poel 2009) and Disclosive Computer Ethics (Brey 2001).

³ For an academic overview of the field see Jacko and Sears (2007).

(Van de Poel 2009). The underlying idea is that technology is not merely enabling, but constitutive: it shapes our practices and institutions in important ways, such as those of health care and transportation, as Van den Hoven argues (2005, 2008) Possibilities pertaining to (new) technologies are created as a direct or indirect consequence of design decisions whereas other possibilities are taken away. As a result, technologies can promote or undermine specific human values.

This is not to say values are solely designed into technology, nor solely conveyed by social drivers and forces. Influence is exerted in various ways, from the direction of the designers of a technology as well as from its users and stakeholders. When new technologies are designed and introduced, users may apply them for purposes other than those intended in design.⁴ By doing so, technologies evolve, i.e. their functionality is adjusted and changed. This is what is referred to in VSD as an interactional position; design and social context and the interaction between both matter (Friedman and Kahn 2003).

Special VSD projects have been carried out for companies like Microsoft (on the reusing and sharing of information, such as code) and Intel. Numerous other projects have been taken up in the academic sphere, such as the RAPUNSEL project (involving the design of games for girls), projects dealing with informed consent online (via cookies and web browsers), network security, and projects in the field of human-robotic interaction or in the field of display technology. VSD assumes that:

[h]uman values and ethical considerations no longer stand apart from the human-computer interaction (HCI) community but are fundamentally part of our practice. This shift reflects, at least in part, the increasing impact and visibility that computer technologies have had on human lives. Technology may support/enhance and/or undermine/corrupt human values. E.g. information available on the Internet increases access to and use of information, but might entail infringement of privacy or the dissemination of incorrect, false information (Friedman 2004).

In order to do justice to these moral (and political) implications, VSD is employed as a methodology that ‘seeks to design technology that accounts for human values in a principled and comprehensive manner throughout the design process.’ (Friedman et al. 2000) What is more, VSD is a ‘way of doing ethics that aims at making moral values part of technological design, research, and development,’ says Van den Hoven (2005). In this paper this claim is investigated, by exploring whether VSD is indeed suitable as a methodology for designing values into technology. I start by describing the VSD methodology below.

⁴ Although the intended instrumental value of a technology or technological artifact is inherent to its intentional history (Vermaas and Houkes 2004), in complex technology development we are not able to foresee or predict the exact use and consequences of the (overall, final) technology due to uncertainty, i.e. a lack of information. This relates to what Anders Albrechtslund coins the ‘positivist problem,’ the fact that there is a problematic relation between intended design and final use (Albrechtslund 2007). The use of a technology is not linear-causally determined by design, but plays out in practice, in a specific context of use. In consequence, there are many possible ways (multi-stability) in which a technology can be used in practice.

Methodology

VSD is characterized by Friedman et al. (2002a) as an approach that:

brings forward a unique constellation of features. First, Value-Sensitive Design seeks to be proactive: to influence the design of information technology early in and throughout the design phase. Second, Value-Sensitive Design enlarges the arena where values arise to include not only the work place [...] but also education, the home, commerce, online communities, and public life. Third, Value-Sensitive Design enlarges the scope of human values beyond those of cooperation [...] and participation and democracy [...] to include all values, especially those with moral import. Fourth, Value-Sensitive Design contributes a unique integrative methodology that involves conceptual, empirical, and technical investigations. [...]

VSD evaluates and informs (the development of) technologies by taking into account human values. It sets out an integrative and iterative tripartite methodology, consisting of conceptual, empirical, and technical investigations (see, Friedman et al. 2006 or Nissenbaum 2005). Each of the conceptual, empirical, and technical investigations and analyses are carried out iteratively, mutually informing and being informed by the other investigation. These interdependencies are metaphorically described by Nissenbaum as “‘balls in the air’”: Conscientious designers must juggle and keep in play the results of at least three modes, i.e. the results of empirical, conceptual, and technical research (Nissenbaum 2005). The interaction among these three distinct analyses—and so the assembly of these separate methodological perspectives—is part of what makes the VSD an attractive enterprise.

The first methodological part involves a philosophically informed conceptual analysis. Its goal is to identify and articulate, on the one hand, the central values at stake in a particular design context, and, on the other hand, the stakeholders that are affected by this (technology) design.

Since this is its main goal or *raison d'être*, the lion's share of conceptual work in VSD evidently involves values. VSD is related to values in a two-fold way. First, it starts from the observation that design and technology may impact upon values: (new) technology developments may enhance, threaten, or transform existing values. For example, the advent of social networking sites has changed existing conceptions of privacy, especially for younger generations (Livingstone 2008). Secondly, values are fostered and built into design by means of VSD; it seeks to identify values considered to be of importance for a target group, say a certain company, society, or user group, and subsequently safeguarding these values by designing them into technology.

Friedman maintains (2004; but also Friedman and Kahn 2003 and Friedman et al. 2006) that VSD particularly focuses on values with moral import that are often implicated in technological developments, such as the values of human dignity, justice, welfare, human rights, privacy, trust, informed consent, respect for intellectual property rights, universal usability, environmental sustainability, moral responsibility, accountability, honesty, and democracy. According to Mary Cummings, ‘[w]hile neither independent nor exclusive, these values were selected in the

development of VSD because they represent the broad values generally discussed in technology and ethics literature, as well as those that have become more important with the increasing use of computer technologies in everyday life' (2006). She argues that VSD does not employ 'values' as referring to something being of economic worth, but that VSD takes up a broader sense of value: values refer to what persons, either singularly or collectively, consider as important to their lives.

To complement the value investigation, VSD sets out to identify the stakeholders of the technology in question. For VSD it is insufficient to solely articulate the central constructs of a design practice. This needs to be carried out recognising who is affected and to what extent. Direct and indirect stakeholders are distinguished; the former being those who interact directly with the new technology, the latter being those that are affected in a less straightforward way (Friedman 2004).

The first, rather abstract conceptual part of VSD's methodology is complemented with empirical investigation in the second part. Here the focus is on the way stakeholders assess the technology in question. Both qualitative and quantitative research from the empirical sciences are used to inform this part of the deliberation process (Friedman 2004) + focus groups, surveys, interviews, and measurements of user behaviour are some of the methods employed to investigate the stakeholders' assessment of the technology. The point of this research is to find out how stakeholders experience (new) technologies with regard to the values they consider important in relation to their social environment and reference groups.⁵ Moreover, it concerns how they cope with emerging value conflicts. A primary consideration in this phase of analysis is investigating in what way design trade-offs affect perceptions, behaviours, and prioritization of competing values. This also includes the way in which the designer can support or detract from a value conflict (Cummings 2006). The results of these studies are applied for improving the design of (new) technologies by making them more sensitive and compliant with stakeholders' values.

Whereas the focus in both conceptual and empirical investigations is on the stakeholders' perspective with regard to who they are, the values they hold and the value implications of a particular (new) technology, the attention of the third part of the methodology—the technical analysis—is directed more specifically to the design and performance of the technology in question. As mentioned earlier, decisions during the design process knowingly or unknowingly determine to a large extent the way in which a given technology can be used in practice. Each particular design creates certain possibilities for action and application, while at the same time obstructing other possibilities. The technical part of the methodological investigation focuses primarily on how the technology can and will support, or compromise, the human and moral values identified in the other parts. This is where VSD has a pre-emptive stance with regard to ethics and technology: it purports to incorporate the results of the conceptual and empirical phases into design in a proactive way (Friedman 2004). As expounded by Van den Hoven, 'Value-Sensitive Design

⁵ '[V]alues cannot be motivated only by an empirical account of the external world, but depend substantively on the interests and desires of human beings within a cultural milieu', Friedman et al. (2006).

provides us with the opportunity to deal with [...] ethical issues in a new and fresh way: by “frontloading ethics” and by all means the proactive integration of ethical reflection in the stage of design’ (2005). The three parts⁶ of the VSD methodology are developed to secure the objective of not only making technology work, but making it work sensitive to human and moral values.

Will VSD do?

A Discussion of VSD as a Methodology for Implementing Values in Design from an Ethics of Technology Perspective

The potential of VSD as a proactive approach to ethics of technology is well appreciated. It recognises the importance of designing technology conscious of human and moral values over a mere retrospective perspective of discussing and dealing with value considerations after a technology has already been introduced and embedded in society. In contrast with retrospective analyses and the tailoring of technology to the requirements and needs of users with hindsight, VSD supports designing so that values are pre-emptively taken into account. What is more, VSD seems especially appropriate in dealing with a potentially diverse user population and the values they hold. However, in analysing whether VSD can meet the requirements associated with a (normative) methodology for implementing values in design (as part of ethics of technology), I argue that VSD falls short in a variety of ways that I will now discuss.

Who are the Stakeholders?

Prior to the deployment of the VSD methodology, there is one basic issue that needs to be resolved first: How do we identify the stakeholders? This matter is of course not specific to VSD. In any stakeholder analysis the identification of stakeholders is paramount (see e.g., Freeman 1994; Mitchell et al. 1997). However, the objective of designing technology sensitive to - or even consciously and deliberately designed for—human and *moral* values, demands a much more exhaustive and comprehensive identification of stakeholders in order to do justice to underlying issues of fairness and equality. Whereas to meet the objective of designing a technology that is profitable, useful, or that warrants the inclusion of patently relevant stakeholders for a large enough group, incorporating moral values calls for further-reaching research in order to recognize who is affected by the technology in what way and to what extent.

This may seem relatively unproblematic at first sight, but in the case of more complex technologies, identifying stakeholders becomes increasingly difficult - let alone identifying the *indirect* stakeholders. This by implication raises doubts as to how the overall input provided by stakeholders, concerning their evaluation of a

⁶ These methodological parts can take part concurrently, though referred to by Flanagan et al. (2008) as ‘phases’.

particular technology, is to be valued or interpreted. What is the threshold for stakeholder input in order to obtain a sufficient and well-informed understanding of stakeholder values?

Once the issue of the identification of stakeholders is resolved, questions arise such as how these stakeholders can be reached, how their input is gathered and finally, how to deal with conflicting values. These sorts of issues are very important and will be further addressed in the following sections.

How to Integrate Empirical Methods?

The second point of attention concerns the use of empirical methods in VSD. VSD holds that it employs multiple methodologies of the social sciences to study how both direct and indirect stakeholders assess new technologies, i.e. the values these stakeholders render important in relation to a particular technology. The empirical investigation results in knowledge concerning the values and opinions of the stakeholders.

Consider the design of a medical system that supports decision-making for patients suffering from aneurysms. The system advises patients regarding their healthcare. In order to receive advice, the patients have to go online and fill out a questionnaire. What patients (and others) do not know is how the system comes to its conclusions. Therefore, the system is said to be 'opaque.' In addition, the patients are unaware of the available alternatives. Suppose we were to use VSD to evaluate and improve the values in the design of this medical system by means of empirical investigation. Now, imagine that one of the outcomes would be that a large part of the stakeholders fear that their autonomy is seriously threatened. What does this tell us and what can we learn from this information?

First of all, current technology development is often so complex that one can question whether stakeholders are in a position to fully assess new technologies. (New) Technologies are often so complex that they are unintelligible to most people, for example due to incomplete information availability or for lack of competence to assess the technology (how it functions, the possible effects on people, the consequences it yields for the environment, and so on). In the case of the medical decision support system described above, people may feel insecure because the system is opaque and they are unaware of possible alternatives. Moreover, people may err or have mistaken beliefs about factual and normative issues concerning the technology. Secondly, and more fundamentally, it is unclear what stakeholders actually intend to say when they are speaking of particular values.⁷ Not only is this dependent on the extent to which values are clearly defined in the (empirical) research, but also on the way the values are experienced and interpreted in each particular situation or assessment. Consider the value of autonomy in the example described above. Even if the definitions of values of stakeholders, designers, and researchers were to correspond, still these values could be interpreted differently, thereby generating different norms and actions.

⁷ This relates to the conceptualization of values cf. the subsection "[What are Values.](#)"

Without giving ground and substance to values, interviewing stakeholders seems to be on loose grounds; as a result the values are too abstract and multi-interpretible, ultimately undermining the legitimacy of the empirical component of VSD. The outcome of research carried out by Friedman et al. (2002b) on users' conceptions of risks and harms on the web effectively illustrates this point. The study indicates that users emphasise security, privacy, and threats to computer systems as the key harms and risks associated with the web. These values and concepts, however, are not substantially defined in the research. Because, what does it mean if 67% of the users say that they fear security issues? And do they all refer to the same conception of security? In addition, how does the 67% fear of security problems relate to the 49% fear of privacy issues? Hence, despite the fact that the associated empirical research generates numerous results, Friedman, Nissenbaum et al. acknowledge that it is difficult to draw conclusions based on them (2002b). Lacking in this research and in VSD in general, I claim, is reflection on how to deal with the (used) empirical investigations concerning values. As a result it is unclear how the integration of empirical studies relates to the overall—and especially the conceptual—investigation of VSD.

In addition, building on stakeholder opinions or evaluations as done in VSD implies a *shared* and *fixed* point of view:⁸ it is assumed that stakeholders will *have* an opinion and that this opinion is relatively stable. However, people's opinions and beliefs (including moral beliefs) change because of new information, insights, and experiences. For example, people may be unaware of the implicit normativity of a technology until this is experienced, as a result of which they may come to disapprove. What is more, there is a plurality of values and normative positions in the public domain, and VSD needs to reflect on the issue of dealing with this, e.g. in case of persistent disagreement. Hence, this part of the methodology should be deliberative rather than taking the form of a survey so that issues can be made explicit and critically discussed.

The Naturalistic Fallacy

A third problem for VSD as a potential methodology for ethics of technology concerns the way in which choices for specific design options are (normatively) justified. As discussed, part of the VSD methodology consists of an empirical investigation, i.e. exposing the values of stakeholders involved. There is an implicit assumption in the methodology of VSD that one will know what to do *in a normative sense*, once these values are known. This, I contend, is where VSD runs the risk of committing the naturalistic fallacy, i.e. by reducing an 'is' to an 'ought.' Indeed, Friedman et al. (2006) explicitly state that.

⁸ This relates to the issue whether the context for deciding between options is 'static' or 'dynamic:' in static contexts, all options are known with a certain degree of probability regarding their consequences (Schmidtz 1995, discussed by Van de Poel 2009). However, as pointed to before, uncertainty is inherent to complex technology (development) (Sollie 2007, 2009).

it is usually agreed (with Moore) that values should not be conflated with facts (the fact/value distinction) especially insofar as facts do not logically entail value. In other words, ‘is’ does not imply ‘ought’ (the naturalistic fallacy). In this way, values cannot be motivated only by an empirical account of the external world, but *depend substantively on the interests and desires of human beings within a cultural milieu* (italics added)

Yet the conflation of facts and values is exactly what happens when the value stances of stakeholders are taken as the normative input for the VSD of a technology. The final (italicized) part of the quotation points exactly to the weak spot in this case: when Friedman et al. claim that values depend substantively on the interests and desires of human beings within a cultural milieu, this implies a sociological conception of values rather than an ethical one. Hence, this is a naturalistic interpretation of values that is running straight against Moore’s account.

What are Values?

Fourth, the concept of ‘values’ is central to VSD: VSD claims that human values and ethical issues do not stand apart from technology, but are fundamentally part of our technological practices. Technology, especially IT, has a rising impact on society and human lives; it may either support or undermine certain values. As explained, the starting point for VSD is to identify stakeholders and values in technology (design), with the ultimate objective of incorporating values into technology by means of design decisions. Given that values are of paramount importance for VSD, I claim that more attention should be given to the conceptualisation of this notion.

As the starting point for exploring ethics and values in relation to technology design, Friedman and Kahn (2003) raise a series of questions. They ask,

if human values, especially those with ethical import [...] are important, they are no less controversial. What values count? Who decides? Are values relative? Or are some values universal, but expressed differently across culture and context? Does technology even have values? [...] [O]n what basis do some values override others in the design of, say, hardware, algorithms, databases, and interfaces?

Although this outset seems promising from the perspective of ethics of technology, Friedman et al. remain descriptive with regard to the conceptualization of values. They describe a set of values “often impacted upon by technology”. These are concepts that are certainly central to human existence, including major moral concepts such as autonomy, justice, and trust. (2006) However, the definitions of these concepts remain abstract. Friedman et al. characterize the values they mention only briefly. Consider the value of autonomy, which Friedman and Kahn (2003) define as ‘people’s ability to decide, plan, and act in ways that they believe will help them to achieve their goals’ This (rather instrumental) definition of autonomy is formulated in such general terms that it can hardly function as a practical concept for implementing values in design from an ethics of technology

perspective. Hence, Friedman et al. avoid the meta-ethical question about the nature of values just as much as the normative ethical question concerning which values matter, why, and to what degree. Maybe they do so because the ultimate conceptualization of each specific value should be left to each specific VSD.⁹ Yet Friedman et al. remain rather vague on the conceptual level.

Illustrative of this is the way Friedman et al. discuss ‘UrbanSim,’ a project on integrated land use, transportation, and environmental simulation. They notice that stakeholders raise widely divergent values, and then ask how to prioritise and weigh these values. The solution offered in this case is a web-based interface in which stakeholders can ‘select indicators that speak to values that are important for them’ (Friedman et al. 2006). Notwithstanding the attractiveness and praiseworthiness of this undertaking, it seems VSD offers no methodological account for distinguishing genuine moral values from mere preferences, wishes and whims of those involved in the design process.

Notwithstanding this critique, Friedman and Kahn take a position with respect to values (in general) as they state (2003) that VSD defends what they call a ‘middle position’ between moral universality and moral variability. Both extremes, they claim, are surrounded by problems: moral universality cannot find or ground its starting point for morality, i.e. a core set of values central to its moral conviction(s), whereas moral variability has problems regarding values shared among different groups or cultures (as a result of which values lose their meaning, cf. the problem of moral relativism). This leads Friedman and Kahn (2003) to adopt a middle ground, which:

allows for an analysis of universal moral values, as well as allowing for these values to play out differently in a particular culture at a particular point of time. [...] The general principle then is that designs need to be robust enough to substantiate the value under consideration and yet adaptable enough so that different cultures (or subcultures) can use the design in their own way.

Flanagan et al. (2006) argue by the same token that:

the study of values in technologies pushes us to ask questions about origins and sources that are far from settled in the general study of moral and political values. In the absence of clear answers to these questions, recourse to a pragmatic middle-ground seems reasonable. This means going along with the idea of a hierarchy of values in which a “thin” set are taken as common to all humanity and thicker sets associated with groups and subgroups of varying sizes and compositions [...] On the issue whether values can be derived, analytically, or must be discovered, empirically, a middle-ground sees virtue in both, drawing conceptual clarity and normative justification from theoretical works in moral and political philosophy, while supplementing these with knowledge about actual interpretations and value commitments of populations relevant to technologies under study.

⁹ I thank Ibo van de Poel for drawing my attention to this point.

This position ties in with the method of Reflective Equilibrium, which has been put forward by Van den Hoven as most practicable for ethics of information technology. [Van den Hoven 2008, referring to Norman Daniels, 1979, James Griffin, 1993, and John Rawls] Just like the middle position of VSD however, this method is contested due to its lack of clarity. It results in questions regarding the tenability of the ‘thin’ or universal values VSD puts forward: if values are universally valid for one thing but play out differently in various contexts for another, how can we avoid relativist arguments? Consider the value of human dignity: it is true that any culture can subscribe to this value. Yet what is left of its (universalist) meaning if the substance given to it differs in each context? In the one culture or context the value of human dignity may be taken as the basic principle to treat people equally, regardless of their race, skin colour, or religion. Yet in another context, the very same value may have an oppressing effect on people (even if unintentionally). Just consider historical and cultural examples regarding differences in dignity, e.g. of women vs men, or ‘black’ vs. ‘white’ people. In the case of (new) technologies with global impact, such as the Internet, mobile communication, and airplanes, the differences among various interpretations of values become even more pressing. A technology that is introduced in two or more cultures with different conceptions of a value affects the moral emotions regarding the technology, including its uses and consequences, in different ways. For this reason I think that the ‘middle position’ Friedman and Kahn adhere to, is subject to what is called ‘buck-passing’: at the end of the day, the problem is not solved but resurfaces elsewhere. What is lacking is the complementation of VSD with a justificatory theory of values, as will show below.

To recapitulate: the main criticism concerning the notion of values in VSD is that this notion remains underdeveloped. VSD should, in view of a normative approach for ethics of technology, explicate the difference between descriptive and normative values, i.e. what stakeholders factually consider important versus what they should regard as important. If this is lacking, the list of values runs the risk of becoming arbitrary, and any value serving the particular interests of designers (including the initiators and potential customers) may become a serious threat for VSD.

In Need of Ethical Theory

The final point of attention for VSD from an ethics of technology perspective concerns the integration of ethical theory and analysis. Mary Cummings (2006) suggests ‘VSD provides a road map [...] on how to include ethical considerations into design projects, even if they are not trained in philosophy or ethics.’ Although I agree that VSD *might* be able to include ethical considerations in design, I claim that in order to successfully meet this objective, VSD requires complimentary ethical theory and expertise.

Above, I address the conceptualisation of values. Now I will discuss the ordering and weighing of values: What values have priority over others? Are certain values paramount? From the literature it is not clear how VSD purports dealing with conflicting and incommensurable values (see e.g. Van de Poel 2009). This leads to questions of how decisions should be made—and justified accordingly—in case of

conflicting values. Notably, there are two sources of conflicting values: persistent disagreement between people (epistemological source) or dilemmas where trade-offs have to be made (ontological source).

Now let us return to the example of the medical support-system; as mentioned this system is opaque, so we do not know exactly how it calculates or comes to its results. Nor do we, as patients, know the alternative treatments available to us, besides those suggested by the system. The design of such a (medical) system involves a variety of human values and associated questions, e.g.:

- Trust—is it possible to design a system that is technologically safe and sound, and that functions well so that it can be trusted both by physicians and patients?
- Privacy—is all the information collected necessary in order to generate the output, and is the database that contains the data about the questions and results protected against unauthorised access?
- Bias—is the medical system free from bias towards a specific preference for one treatment over the other; is the system not secretly a system produced or sponsored by a pharmaceutical company favouring one of their products?
- Autonomy—to what extent can the patient express his or her preferences regarding life and well-being?

In this example, as in most cases of technology (if not all), multiple values play a role. How do we compare and order these different values?

It becomes clear that VSD cannot be taken as a normative account in which competing values can be balanced and traded off as per a certain principle. In complex design situations, there is no clear-cut way to balance competing values such as autonomy and trust, or freedom from bias and privacy, This raises an important question: who makes the final decision on how to prioritize these competing values? Is this left up to the majority of stakeholders to decide, i.e. when a certain value is ranked as important by a certain percentage of people, or is it left up to the designers? As argued by Van den Hoven (2010), we engage in ethics to come to reasoned solutions and clarifications to practical problems, such as value trade-offs. Although different conceptual ethical frameworks lead to different descriptions of the situation, ethical theory nonetheless provides for sources of moral arguments and moral considerations (Van den Hoven 2010). Therefore, I claim that VSD should be complemented with an explicit ethical theory.

Concluding Remarks for VSD

VSD is an attractive approach for assessing values in (technology) design by means of an iterative process between conceptual, empirical, and technical investigations. In this section, VSD has been discussed as a potential methodology for ethics of technology. Despite its promising character, it turns out that VSD falls short in several respects. It cannot yet be considered as a fruitful methodology for implementing values in design from a normative perspective. Therefore, for VSD to be applied as an eligible methodology for ethics of technology, the issues raised in this article need to be dealt with first. At present VSD suffers from the following deficiencies: (1) VSD does not have a clear methodology for identifying

stakeholders, (2) the integration of empirical methods with conceptual research within the methodology of VSD is obscure, (3) VSD runs the risk of committing the naturalistic fallacy when using empirical knowledge for implementing values in design, (4) the concept of values, as well as their realization, is left undetermined and (5) VSD lacks complimentary or explicit ethical theory for dealing with value trade-offs.

I claim that in order for VSD to function as a methodology for implementing values in design from an ethics of technology perspective, it needs to give more explicit attention to its value aims and assumptions. At this point the value analysis of VSD is descriptive rather than normative, leading to the question of whether VSD can reach its original objective of enlarging the scope of human values to include all values, especially those with moral import. For the normative evaluation of a technology an explicit and justified ethical starting point or principle is required.

Towards Value Conscious Design

This paper was meant to explore the criteria of adequacy for a methodology of ethics of technology for implementing moral values into technology. VSD was taken as the prime candidate for such a methodology. However, what has been shown is that VSD lacks a normative, ethical component in order for it to meet the criteria for implementing (moral) values into design in an ethically justified manner. Taking values into account is not the same as normative reasoning. For this reason, I claim that applying VSD as it currently stands is a bridge too far for ethics of technology. For our purpose, technology should not only be made 'sensitive' to values, in other words, account for value considerations, but instead the objective should be to have technology consciously and deliberately designed to include ethical value considerations. I therefore propose to complement VSD with the following criteria:

- (1) The clarification and explanation of the overall (ethical) objective in view of a methodology of integrating moral values in design. In other words, the methodology to be used for implementing moral values into design needs to be explicit regarding its normative aims. What is the goal we wish to obtain? What entities are considered 'values,' is it desirable to make a technology sensitive to values in general, or primarily to moral values?
- (2) The next step is the explication of the ethical theory to be used. As discussed, VSD is in need of complementary ethical theory. By means of ethical theory, a light can be shed on ethical issues and associated moral arguments and considerations. To my view, making ethical deliberation explicit is crucial for a normative account; this way, relevant ethical concepts and value choices can be clearly identified and design decisions motivated accordingly.
- (3) Related to the previous points is the necessity for a clear-cut delineation of the (practical) objectives in each specific design situation with regard to the values to be considered in design. This includes a specification of *whose* values are to be considered (which stakeholders), what empirical methods are to be used,

how, and why, and how does one come to points of action on the basis of this input without losing sight of ethical objectives (cf. (1)).

- (4) The safeguarding and monitoring of the process of incorporating moral values into design, e.g. by including a ‘values advocate’ in the design team (see Manders-Huits and Zimmer 2007, 2009). This person is concerned with the effects of (1), (2), and (3) by (a) identifying and motivating choices for including certain stakeholders and leaving out others, (b) explicating value conflicts and trade-offs, and (c) questioning and evaluating (design) choices in light of overall (moral) aims.

In conclusion, for an approach as VSD to become eligible as a normative account for integrating and implementing moral values into design, I propose to complement it with the explication of an ethical theory. This way, technology is not only made sensitive to moral values, but can also be consciously and deliberately designed that way.¹⁰ I’d like to term this alternative reading of implementing moral values in design “Value Conscious Design.”¹¹ By means of engaging in Value Conscious Design ethics is implemented in design, thereby activating ethical expertise for the benefit of improving our (technological) environment.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Albrechtslund, A. (2007). Ethics and technology design. *Ethics and Information Technology*, 9, 63–72.
- Asveld, L., & Roeser, S. (Eds.). (2009). *The ethics of technological risk*. London: Earthscan.
- Bowker, G. C., & Star, S. L. (1999). *Sorting things out: Classification and its consequences*. Cambridge: MIT Press.
- Brey, P. (2001). Disclosive computer ethics. In R. A. Spinello & H. T. Tavani (Eds.), *Readings in cyberethics* (pp. 51–62). Boston: Jones and Bartlett Publishers Inc.
- Camp, L. J. (2003). Design for trust. In R. Falcone (Ed.), *Trust, reputation and security: Theories and practice*. Berlin: Springer-Verlag.
- Cummings, M. L. (2006). Integrating ethics in design through the value-sensitive design approach. *Science and Engineering Ethics*, 12, 701–715.
- Elmer, G. (2004). *Profiling machines: Mapping the personal information economy*. Cambridge: MIT Press.
- Flanagan, M., Howe, D., & Nissenbaum, H. (2005). Values at play: Design tradeoffs in socially-oriented game design. In *Proceedings of CHI 2005* (pp. 751–760).
- Flanagan, M., Howe, D., & Nissenbaum, H. (2008). Embodying values in technology: Theory and practice. In J. van den Hoven & J. Weckert (Eds.), *Information technology and moral philosophy* (pp. 322–353). Cambridge: Cambridge University Press.
- Florman, S. C. (1987). *The civilized engineer*. New York: St. Martin’s Press.

¹⁰ Notably the methodological comments of this paper are not directly related to technology; this methodology also applies to design situations apart from technology, e.g. institutions.

¹¹ ‘Value Conscious Design’ was first introduced by Manders-Huits and Zimmer (2009) as an umbrella umbrella term to refer to the collective of initiatives involved with promoting human and moral values as an integral part of the conception, design, and development of technological artifacts and systems. In this paper the term is given more substance, subsuming the other initiatives and adding an ethical component.

- Freeman, R. (1994). The politics of stakeholder theory: Some future directions. *Business Ethics Quarterly*, 4, 409–421.
- Friedman, B. (2004). Value sensitive design. In W. S. Bainbridge (Ed.), *Berkshire encyclopedia of human-computer interaction*. Great Barrington: Berkshire Publishing Group.
- Friedman, B. & Kahn Jr., P. H. (2000). New directions: A value-sensitive design approach to augmented reality. In *Conference proceedings of dare 2000: Design of augmented reality environments* (pp. 163–164). New York: Association for Computing Machinery Press.
- Friedman, B., & Kahn, P. H., Jr. (2003). Human values, ethics, and design. In J. A. Jacko & A. Sears (Eds.), *The human-computer interaction handbook* (pp. 1177–1201). Mahwah: Lawrence Erlbaum Associates.
- Friedman, B., Kahn, P. H., Jr., & Borning, A. (2002a). *Value sensitive design: Theory and methods*. UW CSE technical report 02-12-01. Seattle: Department of Computer Science and Engineering, University of Washington.
- Friedman, B., Nissenbaum, H., Hurley, D., Howe, D. C., & Felten, E. (2002b). *Users' conceptions of risk and harms on the web: A comparative study, CHI 2002: Changing the world, changing ourselves*. Accessed April 2, 2009, from <http://doi.acm.org/10.1145/506443.506510>.
- Friedman, B., Kahn, P. H., Jr., & Borning, A. (2006). Value sensitive design and information systems. In P. Zhang & D. Galletta (Eds.), *Human-computer interaction in management information systems: Foundations* (pp. 348–372). New York: Armonk. Reprinted Sharpe, M. E., (2008) In Himma K. E. & H. E. Tavani (Eds.), *The handbook of information and computer ethics* (pp. 69–101). Hoboken, NJ: Wiley.
- Introna, L. D., & Nissenbaum, H. F. (2000). Shaping the Web: Why the politics of search engines matters. *The Information Society*, 16(3), 169–186.
- Jacko, J. A., & Sears, A. (Eds.). (2007). *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications* (2nd ed.). Boca Raton: CRC Press.
- Joerges, B. (1999). Do politics have artefacts? *Social Studies of Science*, 29(3), 411–431.
- Lessig, L. (1999) *Code and other laws of cyberspace*. New York: Basic Books.
- Livingstone, S. (2008). Taking risky opportunities in youthful content creation: Teenagers' use of social networking sites for intimacy, privacy and self-expression. *New Media Society*, 10(3), 393–411.
- Manders-Huits, N. & Zimmer, M. (2007). Values and pragmatic action: The challenges of engagement with technical design communities. In L. Hinman, P. Brey, L. Floridi, F. Grodzinsky, & L. Introna (Eds.), *Proceedings of the seventh international conference of computer ethics: philosophical enquiry* (pp. 238–248). Enschede: Center for Telematics and Information Technology.
- Manders-Huits, N., & Zimmer, M. (2009). Values and pragmatic action: The challenges of introducing ethical intelligence in technical design communities. *International Review of Information Ethics*, 10(2), 37–45.
- Mitchell, R., Agle, B., & Wood, D. (1997). Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. *The Academy of Management Review*, 22(4), 853–886.
- Nissenbaum, H. (2005). Values in technical design. In C. Mitcham (Ed.), *Encyclopedia of science, technology, and ethics* (pp. 66–70). New York: Macmillan.
- Schmidtz, D. (1995). *Rational choice and moral agency*. Princeton: Princeton University Press.
- Sollie, P. (2007). Ethics, technology development and uncertainty: An outline for any future ethics of technology. *Journal of Information, Communication and Ethics in Society*, 5(4), 293–306.
- Sollie, P. (2009). On uncertainty in ethics and technology. In P. Sollie & M. Düwell (Eds.), *Evaluating New Technologies: Methodological Problems for the Ethical Assessment of Technological Developments*. Series of The International Library of Ethics, Law and Technology (Vol. 3, pp. 141–158). Dordrecht: Springer.
- Van den Hoven, J. (2005). Design for values and values for design. *Information Age*, 2005, 4–7.
- Van den Hoven, J. (2008). Moral methodology and information technology. In K. E. Himma & H. T. Tavani (Eds.), *The handbook of information and computer ethics* (pp. 49–69). New York: Wiley.
- Van den Hoven, J. (2010) The use of normative theories in computer ethics. In L. Floridi (Ed.), *The cambridge handbook of information and computer ethics*. Cambridge: Cambridge University Press.
- Van de Poel, I. (2009). Values in engineering design. In A. Meijers (Ed.), *Handbook of the philosophy of science. Volume 9: Philosophy of technology and engineering sciences* (pp. 973–1006). Amsterdam: Elsevier.
- Vedder, A. & Wachbroit, R. (2003). Reliability of information on the Internet: Some distinctions. *Ethics and Information Technology*, 5, 211–215.

- Vermaas, P., & Houkes, W. (2004). Action versus functions. A plea for an alternative metaphysics of artifacts. *Monist*, 87, 52–71.
- Winner, L. (1980). Do artifacts have politics? *Daedalus*, 109(1), 121–136.
- Zimmer, M. (2008). Privacy on planet Google: Using the theory of “Contextual Integrity” to expose the privacy threats of Google’s quest for the perfect search engine. *Journal of Business & Technology Law*, 3(2), 109–126.