Conflicting Values in Design for Values

Ibo van de Poel

Contents

Abstract

Designers are regularly confronted with conflicting values in design: different values select different design options as best. This contribution deals with how one can deal with such value conflicts in design for values. A characterization of value conflict in design is given, and the notion is compared with the notion of moral dilemmas. It is further argued that value conflicts in design entail a kind of multi-

I. van de Poel (✉)

Department of Ethics and Philosophy of Technology, Delft University of Technology, Delft, The Netherlands
e-mail: i.r.vandepoel@tudelft.nl

criteria decision problems to which Arrow's impossibility theorem applies. This theorem says that there is no procedure to aggregate scores on individual criteria (values) into an overall score unless one is willing to violate one of more minimally reasonable conditions for any such an aggregation procedure. Six methods to deal with value conflicts (cost-benefit analysis, direct trade-offs, maximin, satisficing, judgment, and innovation) are discussed. Three of these avoid Arrow's theorem by assuming a form of value commensurability, although they may be too informationally demanding and have other disadvantages as well. The other three are non-optimizing methods that do not result in one best solution and therefore do not entirely solve the value conflict, although they are a way forward in some respects. In conclusion, an approach that combines the several methods is proposed as a way to deal with cases of conflicting *moral* values in design and which avoids many of the disadvantages of the single methods.

Introduction

In design, we often try to respect or promote a range of values such as safety, sustainability, human welfare, privacy, inclusiveness, and justice. When we design not for one value but for a range of values, it will regularly occur that design options that score good on one value score less on another. In other words, if we use the values as choice criteria, for example, during concept selection, one value will point in the direction of one particular design and another value in the direction of another. How can we deal with such value conflicts in design?

There are lots of methods in traditional engineering to deal with trade-offs and conflicts between evaluation criteria which are in principle also relevant for value conflicts. This includes such methods as multi-criteria design analysis, the method of weighted objectives, Pugh charts, and quality function deployment (QFD). Most of the methods, however, do not pay explicit attention to their value dimension. Moreover, it has been argued that many of these methods are methodologically flawed. Franssen (2005) has shown that Arrow's impossibility theorem also applies to multi-criteria decisions in engineering design. Applied to design, this theorem implies that it is impossible to aggregate scores of options on individual criteria into an overall ordering of the options on all criteria without violating one or more axioms, which any reasonable aggregation procedure should minimally meet.

Value conflict, then, is a persistent problem in engineering design. The aim of this chapter is to explore ways in which we can deal with value conflict if one designs for values. I start with further exploring the notion of value conflict. In doing so, I will also relate value conflict to moral dilemmas and will provide a number of examples of value conflict in engineering. In the next section, I explain

how Arrow's impossibility theorem applies to multi-criteria decision problems in design, and I discuss its consequences and conditions under which it may be avoided. In section "Approaches for Dealing with Value Conflict," I explore six different methods for dealing with value conflicts in design. In the final concluding section, I compare the methods and propose an approach that combines different methods to deal with conflicts between moral values in design.

Value Conflict in Engineering Design

Value Conflict and Moral Dilemmas

A value conflict may be defined as the situation in which all of the following conditions apply (Van de Poel and Royakkers 2011):

1. A choice has to be made between at least two options for which at least two values are relevant as choice criteria.
2. At least two different values select at least two different options as best.
3. There is not one value that trumps all others as choice criterion. If one value trumps another, any (small) amount of the first value is worth more than any (large) amount of the second value.

The reason for the second condition is that if all values select the same option as the best one, we can simply choose that one, so that we do not really face a value conflict. The reason for the third condition is that if one value trumps all others, we can simply order the options with respect to the most important value, and if two options score the same on this value, we will examine the scores with respect to the second, less important, value, and so on. So if values trump each other, there is not a real value conflict.

Value conflicts are somewhat similar, though not entirely the same as moral dilemmas. Williams (1973, p. 180) provides the following general characterization of moral dilemmas:

1. The agent ought to do a.
2. The agent ought to do b.
3. The agent cannot do a and b.

Moral dilemmas are thus formulated in terms of conflicting "oughts" which cannot be followed at the same time. It is instructive to try to formulate value conflicts in terms of moral dilemmas to see the similarities and differences between both types of conflict. On the basis of the earlier conditions, we might give the following characterization of a value conflict for the case of two values (v and w) and two options (a and b):

1. Value v selects option a as best.
2. Value w selects option b as best.

3. The values v and w do not trump each other.
4. It is impossible to choose both a and b.

This formulation would amount to a moral dilemma if the following two conditions are also met:

5. Option a ought to be chosen because value v selects it as best.
6. Option b ought to be chosen because value w selects it as best.

Statements (5) and (6) are, however, far from uncontroversial. There are in fact two independent objections possible against (5) and (6). The first objection is that if in a choice situation two values v and w are relevant, an "ought" cannot follow from considering only one of the values unless that value trumps the other (which is in fact denied by (3)). The reason for this is that "ought" judgments are all things considered judgments that take into account all relevant considerations in the (choice) situation.[1] A second possible objection is that (5) and (6) seem to presuppose that we ought to choose what brings about most value. This maximizing assumption is indeed present in many choice methodologies and in some ethical theories. It is, however, not an uncontroversial assumption.[2] Therefore, I think it is better to avoid this assumption in a characterization of value conflict.

If a value conflict is indeed characterized by (1)–(4), a value conflict does not entail a moral dilemma, although value conflict may occasionally amount to moral dilemmas. But even if value conflicts do not necessarily include hard choices in which two (or more) "oughts" conflict, they still involve difficult choices, and it may be hard to know not only what to choose but also how to choose. In the remainder of the chapter, I will focus on value conflicts but I will come back to moral dilemmas in the conclusion.

Examples of Value Conflict Engineering Design

Let me now turn to a number of examples of value conflict in engineering design. This will both make the above discussion more tangible and, at the same time, it will show how this is relevant for design for values.

Example 1: Safety Belts Suppose you are a car designer and you are designing a safety seat belt system for a car. You know that the use of seat belts in cars reduces

---

[1]This is exactly why moral dilemmas have been characterized as hard choices (Levi 1986) and, also, why some philosophers deny the possibility of moral dilemmas because according to them there is always an overall point of view in which only one ought applies.

[2]What makes the assumption problematic is not just the assumption that "more value is better than less" but also the assumption that there is an obligation (an "ought") rather than just a recommendation to choose the highest value. Even if "more value is better than less," we might not be obliged to choose the object with higher value.

Table 1 Different seat
belt systems for cars

	Safety	Freedom
Traditional seat belt	Lowest	Highest
Seat belt with warning signal	Moderate	Moderate
Automatic seat belt	Highest	Lowest

the number of fatalities and injuries, that is, a car driver or other car occupants have a lower probability to get killed or injured (or they are less severely injured) in case of a car accident if they wear a seat belt than if they do not. You also know, however, that some people tend to forget the use of seat belts or find them unpleasant to use or do not use them for other reasons.

Two values that are relevant in the design of a seat belt system then are safety and freedom. Safety is here mainly understood as lower probability of fatality or injury, or less severe injuries, in case of an accident for car drivers or occupants. Freedom is by and large understood as the presence of a free and uninfluenced choice in using a safety belt or not.

Let us suppose that there are three options that you have selected to choose between as the designer of a safety belt system: (1) a traditional seat belt; (2) a so-called automatic seat belt that enforces its use, for example, by making it impossible to enter the car without using the seat belt or making it impossible to start and drive the car without using the seat belt; and (3) a system with a warning signal that makes an irritating noise if the seat belt is not used.

Table 1 represents these three options and their scores on the values of safety and freedom. The choice situation is an example of a value conflict as I defined it above. The question is how the designer who wants to design a seat belt for both the value of safety and the value of freedom can choose between the three options.

Example 2: The Storm Surge Barrier in the Eastern Scheldt After a huge flood disaster in 1953, in which a large number of dikes in the province of Zeeland, the Netherlands, gave way and more than 1,800 people were killed, the Delta plan was drawn up.[3] Part of this Delta plan was to close off the Eastern Scheldt, an estuary in the southwest of the Netherlands. From the end of the 1960s, however, there was growing societal opposition to closing off the Eastern Scheldt. Environmentalists, who feared the loss of an ecologically valuable area because of the desalination of the Eastern Scheldt and the lack of tides, started to resist its closure. Fishermen also were opposed to its closure because of the negative consequences for the fishing industry. As an alternative, they suggested raising the dikes around the Eastern Scheldt to sufficiently guarantee the safety of the area.

In June 1972, a group of students launched an alternative plan for the closure of the Eastern Scheldt. It was a plan that had been worked out as a study assignment by students of the School of Civil Engineering and the School of Architecture of the Technical University of Delft and the School of Landscape Architecture of the

---

[3]A more detailed discussion can be found in van de Poel (1998).

Table 2 Options for Eastern Scheldt

	Safety	Ecology
Closing off Eastern Scheldt	Highest	Lowest
Storm surge barrier	Moderate	Moderate
Raising dikes around the Eastern Scheldt	Lowest	Highest

Agricultural University of Wageningen. The values the students focused on were safety and ecological care. On the basis of these values, they proposed a storm surge barrier, i.e., a barrier that would normally be open and allow water to pass through but that could be closed if a flood threatened the hinterland.

Table 2 lists the three abovementioned options. The original plan to close off the Easter Scheldt would be the safest (in terms of probability of flooding and number of fatalities in case of flooding) but scores the worst in terms of ecology. Heightening the dikes would most likely be the least safe (although this was not entirely beyond debate) and the best in terms of ecology. The storm surge barrier was a creative compromise between both values.

Example 3: Refrigerants for Household Refrigerators As a consequence of the ban on CFCs in the 1990s, an alternative to CFC 12 as refrigerant in household refrigerators had to be found.[4] I will focus here on three (moral) values that played an explicit role in the search for alternative coolants: safety, health, and environmental sustainability. In the design process, safety was mainly understood as nonflammability, and health as nontoxicity. Both understandings were based on existing codes, standards, and testing procedures like the ASHRAE Safety Code for Mechanical Refrigeration. ASHRAE is the American Society of Heating, Refrigerating, and Air-Conditioning Engineers. Environmental sustainability was typically formulated in terms of a low ODP (ozone depletion potential) and a low GWP (global warming potential). Both ODP and GDP mainly depend on the atmospheric lifetime of refrigerants. In the design process, a conflict between those three considerations arose. This value conflict can be illustrated with the help of Fig. 1.

Figure 1 is a graphic representation of CFCs based on a particular hydrocarbon. In the top, there is methane or ethane or another hydrocarbon. If one moves to the bottom, hydrogen atoms are replaced by either chlorine atoms (if one goes to the left) or fluorine atoms (if one goes to the right). In this way, all CFCs based on a particular hydrocarbon are represented. The figure shows how the properties flammability (safety), toxicity (health), and atmospheric lifetime (sustainability) depend on the exact composition of a CFC. As can be seen, minimizing the atmospheric lifetime of refrigerants means maximizing the number of hydrogen atoms, which increases flammability. This means that there is a fundamental trade-off between

---

[4]A more detailed discussion can be found in van de Poel (2001). The data for GWP given in Table 3 are based on a more recent IPCC report (Solomon et al. 2007, Table 2.14) and therefore deviate from the data in van de Poel (2001). The GWP for a 100-year time horizon is given in the table.
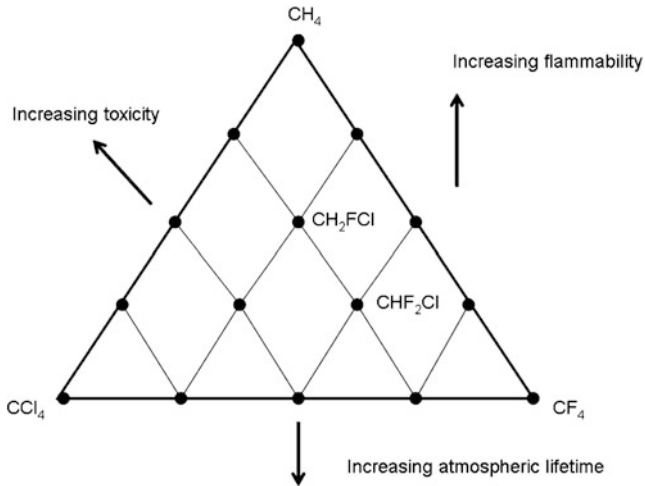
Fig. 1 Properties of refrigerants (Based on McLinden and Didion (1987))

Table 3 Properties of refrigerants

	Environmental sustainability		Health	Safety
	ODP	GWP	Toxicity class	Flammability class
CFC 12	1	10,900	A	1
HFC 134a	0	1,430	A	1
HFC 152a	0	124	A	2
HC 290 (propane)	0	3	A	3
HC 600a (isobutane)	0	3	A	3

flammability and environmental effects, or between the values of safety and sustainability.

Table 3 lists some of the options that were considered as replacements for CFC 12 as coolant in household refrigerators. The ODP (ozone depletion potential) is measured relative to CFC 12, the global warming potential (GWP) relative to CO_2. For health, two toxicity classes have been defined in relevant codes and standards; class A is considered toxic and class B nontoxic. The same codes and standards define three flammability classes; class 1 is considered nonflammable, class 3 highly flammable, and class 2 moderately flammable.[5] The coolants listed in Table 3 exemplify a value conflict specifically between the value safety and environmental sustainability.

---

[5]It should be noted that coolants in the same flammability class are not necessarily equally flammable; neither are coolants in the same toxicity class necessarily equally toxic. Membership of certain flammability or toxicity class is determined by certain threshold values, and therefore does not reflect degrees of flammability or toxicity that follow the patterns shown in Fig. 1.

Arrow's Theorem and Multi-criteria Decision-Making

It has been shown that the famous Arrow's theorem from social choice (Arrow 1950) also applies to multi-criteria decision-making (May 1954; Arrow and Raynaud 1986; Franssen 2005). Since value conflicts are a kind of multi-criteria decision problems, it also applies to value conflicts. Arrow's theorem establishes the impossibility of certain solutions that meet a number of minimally desirable characteristics. It thus also sets serious limits to ways to deal with value conflicts. This section discusses these limitations and how they follow from Arrow's theorem. The next section will, then, on the basis of this information, discuss approaches to deal with value conflicts.

The Decision Problem for the Designer

Let me begin with formulating the decision problem that a designer or a design team faces when a value conflict occurs in design. I will assume here that there is one designer who faces a value conflict and makes a decision. This is of course a simplification compared to reality where often a design team is involved and where the additional complexity faced is how to reach a decision together.[6] I will further assume that the designer aims to design for values. So in dealing with the value conflict, the designer does not aim at the design that meets his/her own personal preferences best but rather he/she looks for a design that best meet the relevant values at stake. The designer may be said to take an ethical point of view.[7]

The decision problem faced by the designer may now be modeled as follows:

1. In the choice situation S, n values $v_1 \ldots v_i \ldots v_n$ are relevant.
2. In the choice situation S, m options $o_1 \ldots o_j \ldots o_m$ are feasible.
3. For each value v_i, a corresponding ordinal value function exists so that $v_i(o_a) \geq v_i(o_b)$ implies that option o_a is at least as good (or better) as option o_b with respect to value v_i.

In lay terms, (3) says that it is possible to order the options for each relevant value on a scale from better to worse. This implies at least an ordinal measurement of the options on each of the relevant values. Table 4 explains the difference between different measurement scales.

Below, we will consider the question whether it is possible to derive on the basis of the information contained in (1)–(3) a value function $v(o_j)$ that orders all options

---

[6]To this social choice problem, the Arrow's theorem also applies (Kroes et al. 2009).

[7]According to some ethical theories (e.g., Kantianism), all people would come to the same conclusion if they take an ethical point of view. In that case, the earlier assumption of there being just one decision-maker would be justified.

Table 4 Measurement scales

Measurement scale	Invariant	Allowed transformation	Degree of freedom	Example
Ordinal	Order	Monotonic	Infinite	Mohs scale of mineral hardness[a]
Interval	Ratio between differences	Positive linear	Two (zero point and unit)	Temperature measured in Celsius
Ratio	Ratio	Scalar	One (unit)	Distance measured in meters

[a]On this scale, the hardness of a material is measured by determining the hardest material that the given material can scratch and/or the softest material that can scratch the given material

on (at least) an ordinal scale with respect to all values $v_1 \ldots v_i \ldots v_n$ *in combination*. Since we are looking for the general possibility to construct such a value function from (1)–(3), this includes cases when the designer faces a value conflict.

Arrow's Theorem

Let me now turn to Arrow's theorem. Arrow considered the following social choice situations:

1. In the social choice situation S, n individuals $w_1 \ldots w_i \ldots w_n$ are involved in the decision.
2. In the choice situation S, m options $o_1 \ldots o_j \ldots o_m$ are feasible.
3. For each individual w_i, a corresponding ordinal value function exists so that $w_i(o_a) \geq w_i(o_b)$ implies that option o_a is at least as good (or better) as option o_b in terms of the preferences of individual w_i.

Again, (3) says that each individual can order all options on a scale from best to worst (allowing for indifferences between some options).

Arrow shows that if $n \geq 2$ and $m \geq 3$, it is impossible to find a function or decision procedure that meets a number of minimally reasonable conditions to translate the individual preferences into a collective preference. These minimal conditions are[8]:

- *Collective rationality*. This condition implies that the collective preference ordering must be complete and transitive. A preference ordering is complete if all alternatives are ordered by it. Transitivity requires that if o_a is ordered over o_b and o_b is ordered over o_c, o_a is also ordered over o_c.

---

[8]The requirements given are somewhat weaker than those originally formulated by Arrow. See, e.g., Sen (1970). See also Franssen (2005).

- *Unrestricted domain*. This condition implies that there are no restrictions with respect to how an individual orders the alternatives, apart from conditions of completeness and transitivity for the individual preference orderings.
- *Pareto principle*. This condition implies that if everyone prefers o_a over o_b, the collective preference ordering should order o_a over o_b.
- *Independence of irrelevant alternatives*. The ordering of alternative o_a relative to alternative o_b may not depend on the inclusion or exclusion of a third alternative in the set of alternatives.
- *Absence of a dictator*. This condition implies that there is no individual whose preferences determine the collective preference.

Arrow's theorem means that no *general* procedure exists to translate individual preferences into a collective preference ordering unless one is willing to breach one of the abovementioned conditions.

Application of Arrow's Theorem to Multi-Criteria Decision Problems

It can easily be seen that the choice situation faced by the designer described above is structurally similar to the social choice situation to which Arrow's theorem applies. The difference is that where in the original Arrow case, individuals order the alternatives, in the design choice situation, values order the alternatives.[9] In both cases, the possibility of an aggregation procedure that meets some minimally desirable characteristics is at stake.

Franssen (2005) has argued that Arrow's theorem also applies to multi-criteria choices, and he also argued that all conditions listed above that play a role in Arrow's theorem are still reasonable in the case of multi-criteria decision problems in engineering (see also Jacobs et al. 2014). I will not repeat all of his arguments but will focus on some of the main issues with respect to the applicability of Arrow's conditions to choices with various values in engineering design.

One possible objection against the requirement of *collective rationality* is that in design, we only want to select the best design and we have no interest in ordering all the other designs. It can, however, be shown that if we release the condition accordingly, an impossibility theorem that is rather similar to the original Arrow's theorem can be proven (Mas-Colell and Sonnenschein 1972).

With respect to the condition of *unrestricted domain*, one might argue that *given* a specific value and a specific range of options, the ordering of those options on that value is not unrestricted. We cannot suddenly say in the safety belt example (example 1 in section "Value Conflict in Engineering Design") that the traditional

---

[9]A difference is that the values have different degrees of importance, whereas in the original Arrow choice situation, each individual has equal weight. We can, however, repair this by replacing each value by x values where x is the (relative) degree of importance of that value (cf. Franssen 2005).

seat belt is best in terms of safety. The condition of unrestricted domain is, however, to be understood as expressing that we look for a procedure that is able to deal with any way, a value may order the alternatives (as long as the conditions of completeness and transitivity are met). The Arrow's theorem thus shows the impossibility of a generally applicable procedure, which does not imply the impossibility of solving one particular case.

The *Pareto principle* says that if all values select an option as best that option should be ordered as best in the overall ordering. This seems hardly contestable. Still, there are two possible objections. One is that more value is not always better; sometimes we want to minimize a value (or a criterion for a value), or sometimes we might strive for a target rather than for as much as possible. However, such cases can usually be mathematically converted into a new criterion in which more is indeed better. A second objection is that sometimes the desirable degree of attainment of one value may be dependent on the actual attainment of another value. For example, suppose that two values in the design of a car are "safety" and "looks robust." It might be that for a not so safe car, one prefers a less robust looking design over a more robust looking design, while for the safer car, one prefers the more robust looking design over the less robust looking design (e.g., because one believes that car look should represent the underlying safety features). In cases like this, the Pareto principle does not apply.

The condition of *independence of irrelevant alternatives* seems a quite reasonable condition again. Also, in design, one does not want the choice between two alternatives to depend on the inclusion of a third in the overall choice. Underlying this condition, there are, however, two assumptions in the case of collective choice that Arrow made that have been contested. One assumption is that individual preferences can only be measured on an ordinal scale (and not on an interval or ratio scale). The other is the assumption of the impossibility of interpersonal utility comparison: we cannot compare the utility (also not on an ordinal scale) of one person with that of another.

It has been shown that if the first assumption is somewhat released and we allow preference or utility measurement for individuals on an interval scale, impossibility theorems similar to that of Arrow can be formulated (Hylland 1980). However, it has also been shown that under stronger assumptions about the informational base, aggregation procedures that meet axioms comparable to the ones proposed by Arrow are possible (Roberts 1980; Jacobs et al. 2014). If these assumptions are translated to the context of engineering design, aggregation would be possible in each of the following cases:

1. The score of all options on all individual values (criteria) can be measured on a ratio scale with respect to preference (utility) or value[10] (ratio measurability).

---

[10]If we assume only nonnegative utilities, ratio scale measurement is enough. However, if we also allow negative utilities, an additional commensurability assumption is needed for a reasonable aggregation procedure to be available (Tsui and Weymark 1997).

2. The score of all options on all individual values (criteria) can be measured on interval scales which share a common unit of measurement (unit commensurability).
3. The score of all options on all individual values (criteria) can be measured on a common ordinal scale, so that the score of the xth option on the ith criterion can be ordinally compared with the score of the yth option on the jth criterion (level commensurability).

If any of these three conditions apply, Arrow's theorem can be avoided. It should be noted that while the second and third conditions list a commensurability condition, the first condition does not require commensurability. It only requires that each individual value can be measured on a ratio scale, but this needs be a common scale, so that no commensurability is required. It should further be noted that unit commensurability and level commensurability are independent from each other; you can have unit commensurability without level commensurability, and vice versa. In the remainder I will speak of *value commensurability*, if either unit or level commensurability applies, or if both apply. If both apply, this is also sometimes called full commensurability. I will call two values incommensurable if neither unit nor level commensurability applies.

Franssen (2005) argues that ratio scale measurements of preferences or value (the first condition above) are impossible since ratio scales require extensive measurement, which he believes to be impossible for mental constructs like preference or value. Nevertheless, an often used approach, cost-benefit analysis, may be said to be based on the assumption that money can measure utility or value on a ratio scale (although this assumption is by no means unproblematic). I will discuss cost-benefit analysis in the next section as a possible method to deal with value conflicts; there, I will also discuss two methods that are available if one assumes either unit commensurability (direct trade-offs) or level commensurability (maximin).

The condition *absence of a dictator* in the case of multi-criteria problems implies that there is not one criterion or value that dictates the overall ordering of options. The criterion is the same as the third criterion that I formulated for value conflicts that there is no one value that trumps all others as choice criterion.

Approaches for Dealing with Value Conflict

In this section, I will discuss the main approaches to value conflict and their advantages and disadvantages. The methods that will be discussed are:

– Cost-benefit analysis
– Direct trade-offs
– Maximin
– Satisficing (thresholds)
– Respecification
– Innovation

The first three methods each suppose a specific form of value commensurability through which Arrow's theorem might be avoided as we have seen in the previous section.[11] The other three methods are so-called non-optimizing methods (Van de Poel 2009).[12] They do not aim for one best option, and they do not, or at least not always or necessarily, result in one option that is to be chosen. They therefore do not meet Arrow's condition of collective rationality. Still, they may be useful in dealing with value conflicts in certain circumstances as will become clear.

Cost-Benefit Analysis

In cost-benefit analysis, all relevant considerations are expressed in one common monetary unit, like dollars or euros. Because all values are measured on a common ratio scale (money), cost-benefit analysis assumes both ratio measurability and value commensurability. The advantage of this assumption is that Arrow's theorem is avoided and that it becomes possible to select the best alternative by expressing the score of options on a range of values in a common measure: money.

If we want to apply cost-benefit analysis to value conflicts in engineering design, we somehow need to express gains and losses in values, like freedom, safety, sustainability, etc., in monetary terms. A glance at the examples in section "Value Conflict in Engineering Design" shows how difficult this is. Take the safety belt example: is there a way to express the different degrees of freedom and safety realized by the various designs in monetary terms, and if so, can it be done in a reliable and uncontroversial way? If we look at the second example (the Eastern Scheldt barrier), a cost-benefit analysis was done for the original Delta plan, which still assumed a closed barrier in the Eastern Scheldt (Tinbergen 1959). In this cost-benefit analysis, safety was treated as an imponderable value, i.e., as a value that cannot be expressed in monetary terms.[13] However, ecology and environmental

---

[11]Cost-benefit analysis does not only suppose ratio measurability but also value commensurability because the various values are measured on a common scale (i.e., money). Theoretically, it would be possible to construct a method that only supposes ratio measurability (but see also note 10). Such a method could, for example, proceed by multiplying the score on one value dimension (measured on a ratio scale) with the score on another value dimension (measured on another ratio scale) and then selecting the alternative with the highest score so attained. It seems, however, doubtful whether that leads to a meaningful decision procedure for design, and as far as I know, no such decision methods have been proposed in the design literature. The multiplication of, for example, safety and sustainability, both measured on a (positive) ratio scale, for example, does not seem meaningful as a decision criterion. Note that in physics, such multiplications are sometimes meaningful, e.g., mass times velocity is a measure of momentum.

[12]The three non-optimizing methods discussed are the same as in Van de Poel (2009). There, I also discuss a fourth (diversity) that is not discussed here as it cannot be employed in a single design process (which I take to be the focus of this chapter). The discussion follows Van de Poel (2009) but has been updated and somewhat revised in several respects.

[13]The resulting costs can then be interpreted as the amount of money that one should be willing to pay for the increase in safety that is obtained by realizing the Delta plan.

concerns were not taken into account in the original cost-benefit analysis. It might be argued that these values are also imponderable. However, if one treats both the (conflicting) values of safety and ecology as imponderable, a cost-benefit analysis is of no help in example 2.

Despite the above reservations, approaches and methods like contingent validation have been developed to express considerations like safety, freedom, and ecology in monetary terms. Contingent validation proceeds by asking people how much they are willing to pay for a certain level of safety or for, for example, the preservation of a piece of beautiful nature. In this way, a monetary price for certain safety levels or a piece of nature is determined. Such methods are, however, beset with methodological problems, and it is questionable whether they deliver a reliable measurement for the values at stake. For example, the monetary value of a piece of nature is lower if one asks people how much they are willing to pay for it than if one asks for how much one would want to be compensated for giving it up (Horowitz and McConnell 2002). It has been suggested that such differences may be due to the intrinsic (moral) value of nature (Boyce et al. 1992).

There are a number of more fundamental issues with cost-benefit analysis as well. For one thing, it is questionable whether one could regard money as a good measure for preference or utility (as is assumed, as we saw in section "Arrow's Theorem and Multi-Criteria Decision-Making," if one conceives of money as a way to measure utility on a ratio scale). One problem here is the diminished marginal utility of money. For most people, a gain in income from 100 to 200 euros will imply a larger increase in utility than a gain in income from 10.100 to 10.200 euros, while both increases in utility should be the same if money is to measure utility on a ratio scale. Another problem is that it is questionable whether a similar gain in income, say 100 euros, will realize the same increase in utility for two different persons.

Another fundamental problem is whether we can measure a range of values like safety, sustainability, freedom, justice, etc., in terms of a common measure on a ratio scale (be it in terms of money, utility, or whatever other value measure). This is not just a practical or methodological issue about how to express these values in monetary terms (as discussed above), but it involves a more fundamental assumption about the nature of values. It should be noted that if one assumes that values are commensurable on a ratio scale, a loss in one value can always be compensated by a gain in another value (if the latter gain is large enough). Some authors believe that this assumption is wrong for at least some values. Consider, for example, the following trade-off: for how much money are you willing to betray your friend? It may well be argued that accepting a trade-off between friendship and financial gain undermines the value of friendship. On this basis, it is constitutive of the value of friendship to reject the trade-off between friendship and financial gain (Raz 1986). Such constitutive incommensurability seems especially true of moral values and values that regulate the relations between, and the identities of, people.

Even if some of the above issues are solved (or are just neglected as is often the case in actual cost-benefit analyses), one faces a range of additional methodological

and ethical issues in cost-benefit analysis (Hansson 2007). One issue is how to discount future benefits from current costs (or vice versa). One dollar now is worth more than one dollar in 20 years, not only because of inflation but also because a dollar now could be invested and would then yield a certain interest rate. To correct this, a discount rate is chosen in cost-benefit analysis. The choice of discount rate may have a major impact on the outcome of the analysis. Another issue is that one might employ different choice criteria once the cost-benefit analysis has been carried out. Sometimes all of the options in which the benefits are larger than the costs are considered to be acceptable. However, one can also choose the option in which the net benefits are highest or the option in which the net benefits are highest as a percentage of the total costs.

From the above considerations and reservations, it does *not* follow that one should never use cost-benefit analysis to deal with value conflicts in design. As we will see below, other approaches for dealing with value conflicts have their problems and disadvantages as well. In some design decision contexts, the above concerns may be less serious or we might have reasons to prefer cost-benefit analysis over other approaches. Still, one should be aware of the abovementioned limitations and issues in applying cost-benefit analysis.

Direct Trade-Offs

A second approach to deal with value conflict is to make direct trade-offs between the relevant values. As we have seen in section "Arrow's Theorem and Multi-Criteria Decision-Making," this requires that the individual values are measured on (at least) an interval scale and that there is unit commensurability between the relevant measurement scales. We can then trade off a loss in one value dimension for a gain in another value dimension. The advantage of this approach is that it avoids Arrow's theorem by assuming unit commensurability, and it does so without the need of expressing all values in terms of money, which is an advantage compared to cost-benefit analysis.

It is worth noting that in the examples discussed in section "Value Conflict in Engineering Design," all relevant values are not (yet) measured on an interval scale. In the safety belt example, Table 1 represents measurements of the options on both the value of safety and the value of freedom on an ordinal rather than an interval scale. In this case, it might be possible to measure safety on an interval scale (by expressing it, e.g., in a measure of probability of death or injury); for the value of freedom, this seems much more difficult. When we look at the coolants example (example 3), in Table 3, environmental sustainability is operationalized in a measurement on two ratio scales (ODP and GWP), while health and safety are in the table measured on an ordinal scale. We can, however, also operationalize these latter values in such a way that they can be measured on interval scales (see Table 5).

To make value trade-offs, we do not only need an interval (or ratio) scale measurement of the individual values but also unit commensurability. To achieve

Table 5 Properties of refrigerants. The data for OEL and LFL are based on ASHRAE (2013)

	Environmental sustainability		Health	Safety
	ODP	GWP	OEL (occupational exposure limit)[a]	LFL (lower flammability level)[b]
CFC 12	1	10,900	1,000	None
HFC 134a	0	1,430	1,000	None
HFC 152a	0	124	1,000	48,000
HC 290 (propane)	0	3	1,000	21,000
HC 600a (isobutane)	0	3	1,000	16,000

[a]OEL is "the time-weighted average (*TWA*) concentration for a normal 8-h workday and a 40-h workweek to which nearly all workers can be repeatedly exposed without adverse effect" (ASHRAE 2013, p. 4). It is measured in ppm (parts per million) v/v
[b]The minimum concentration in air at which flame propagation occurs. It is measured in ppm (parts per million) v/v

that, the decision-maker (designer) needs to be able to answer questions like "how many units decrease in GWP compensate for one-unit decrease in LFL?"[14] One problem in answering such questions is that trade-offs may not be constant over the entire domain. Consider, for example, the trade-off between costs and safety in the design of a car. It may well be that at low levels of safety, one is willing to pay more for a one-unit increase in safety than at higher levels of safety. So if one establishes value trade-offs, it should be done, taking into account the current value of values being traded off. Keeney (2002) discusses this and other pitfalls in making value trade-offs.

Apart from such avoidable pitfalls, the assumption of unit commensurability in making trade-offs raises the more fundamental issue that I also discussed in relation to cost-benefit analysis, namely, can a gain in one value dimension always compensate a loss in another dimension? As indicated, it has been suggested that unit incommensurability and a resistance to certain trade-offs are constitutive of certain values or goods like friendship. It has also been suggested that values may resist trade-offs because they are "protected" or "sacred" (Baron and Spranca 1997). Such trade-offs between protected or sacred values have also been called taboo trade-offs (Tetlock 2003).

Taboo trade-offs create an irreducible loss because a gain in one value cannot compensate or cancel out a loss in the other. The loss of a good friend cannot be compensated by having a better career or more money. One possible explanation for the existence of taboo trade-offs is that protected values correspond with moral obligations (Baron and Spranca 1997), i.e., they express an obligation to meet a certain value to a certain minimal extent. If interpreted thus moral obligations

---

[14]Note that GWP should be as low as possible, while LFL should be as high as possible.

define thresholds for moral values. It seems plausible that below the threshold, the moral value cannot be traded off against other values because the moral obligation is more or less absolute; above the threshold, trade-offs may be allowed.

Maximin

What if we lower our assumptions about what information is available, i.e., if we do not longer assume the possibility of ratio scale measurement of value (as in cost-benefit analysis) or of unit commensurability (as in trade-offs)? If we still assume some form of commensurability, i.e., what we called level commensurability, a decision rule known as the maximin rule becomes possible. This decision rule tells us to select that alternative that scores best, compared to the other alternatives, on its lowest-scoring value. The advantage of this approach is that it avoids Arrow's theorem (by assuming level commensurability) without assuming unit commensurability and, therefore, without the need for trade-offs. This advantage comes, however, at a certain price as we will see.

Consider again the safety belt case. If we were to compare the traditional safety seat belt and the automatic seat belt with the maximin rule, we are to proceed as follows. First, we judge on what value each of the alternatives scores worst (compared to the other values on which we score that alternative). For the traditional safety seat belt, the worst-scoring criterion is most likely safety, and for the automatic seat belt, it is most likely freedom. In comparing the two alternatives, we should now answer the question: What is worse, the score of the traditional seat belt on safety or the score of the automatic seat belt on freedom? If we answer the latter, we should choose the traditional seat belt. (We can then repeat the procedure to compare the winning alternative with the seat belt with warning signal.)

As the example suggests, for the maximin rule, we only need ordinal measurement of the relevant values. In this respect, it is less demanding than the previous two approaches. At the same time, the judgments that this approach asks us to make seem quite complicated, as it asks us to compare the scores of two alternatives on *different* value dimensions; more formally, the method asks us to compare the score of option a on value v with the score of option b on value w. Especially if there are many alternatives, this may be hard and cumbersome, if not impossible.

One may also wonder how sensible the maximin rule is as a decision rule for conflicting values in engineering design. If we try to interpret what the rule means in the context of engineering design, it boils down to what may be called strengthening the weakest link. One selects the design in which the weakest link of that design (i.e., the worst-scoring value) is relatively the strongest compared to the alternatives. Such an approach seems especially sensible if one wants to avoid underperformance on any of the relevant values (or design criteria). We may therefore say that the maximin rule results in a kind of "robust design."

It should be noted, however, that in some situations, the maximin rule leads to seemingly irrational results. Suppose I have a seat belt design that scores worse on safety than on freedom. Now suppose that through some tinkering, I develop a

design that scores only very slightly worse on safety but much better on freedom than the original design. Obviously, this new design will also score less on safety than on freedom. The maximin rule now tells us to prefer the first design over the new design whatever small the loss in safety (as long as there is some nonzero loss) and whatever big the gain in freedom. At least in some occasions, this seems the wrong advice.

Satisficing

All previous approaches aim at selecting the best alternative (although they define the best differently, especially in the case of maximin). In contrast, in satisficing, one does not look for the optimal option, but first sets an aspiration level with respect to the options that are good enough and then goes on to select any option that exceeds that aspiration level (Simon 1955, 1956). Designers are reported to be satisficers in the sense that they set threshold values for the different design requirements and accept any design exceeding those thresholds (Ball et al. 1994). So conceived, satisficing may also be seen as a way of dealing with conflicting values, i.e., by setting thresholds for each value and then selecting any option that exceeds those thresholds.

An example of satisficing is to be found in the earlier-discussed case of the design of new refrigerants (example 2). On the basis of Fig. 1, the engineers McLinden and Didion drew a more specific figure with respect to the properties of CFCs, which is shown as Fig. 2.

According to McLinden and Didion, the blank area in the triangle contains refrigerants that are acceptable in terms of health (toxicity), safety (flammability), and environmental effects (atmospheric lifetime). This value judgment is a type of satisficing because by drawing the blank area in the figure, McLinden and Didion – implicitly – establish threshold values for health, safety, and the environment.
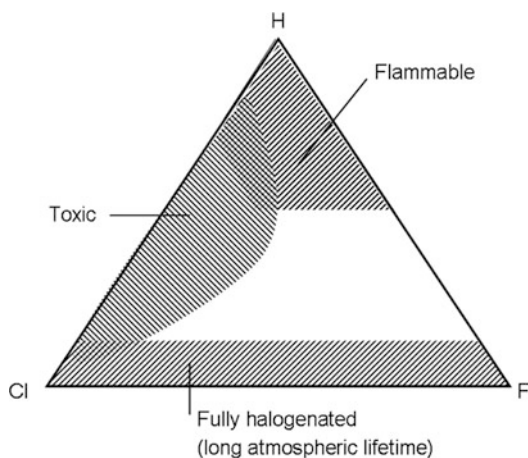

Fig. 2 Properties of refrigerants (Figure from McLinden and Didion (1987))

Table 6 Satisficing thresholds (implicitly) used by McLinden and Didion in drawing Fig. 2 and the score of the various options on these thresholds

Threshold	Environmental sustainability	Health	Safety
	At least one H atom	Toxicity class A (not B)	Flammability class 1 (not 2 or 3)
CFC 12	–	+	+
HFC 134a	+	+	+
HFC 152a	+	+	–
HC 290 (propane)	+	+	–
HC 600a (isobutane)	+	+	–

Table 6 lists the thresholds that they set and it shows that of the earlier-considered alternatives only one, HFC134a, meets all thresholds.

Compared to the earlier-discussed methods, the main advantage of satisficing is that it requires less information, as it does not require any form of commensurability. The price to be paid is that it does not meet all of Arrow's requirements. In particular, it does not meet the condition of global rationality; rather, it orders the total set of alternatives into two sets, one with acceptable alternatives (i.e., those that meet all thresholds) and one with unacceptable alternatives (those that do not meet at least one threshold). Sometimes, the set of acceptable alternatives might consist of one item, as is the case in Table 6, and then it is clear what alternative to choose. However, the set of alternatives may also contain more than one alternative or be empty. If there is more than one acceptable alternative, the decision problem has not been solved yet. To be able to select one alternative, we might opt to satisfice with stricter thresholds or opt for one of the other approaches. If the set of acceptable alternatives is empty, we need to decide whether it is perhaps better not to design something (as no alternative meets all thresholds) or whether the thresholds should perhaps be relaxed.

As the above discussion already suggests, the core issue in satisficing is how to set thresholds. If thresholds are set in an arbitrary way, satisficing can hardly be seen as a rational method for dealing with value conflicts. However, in some situations, thresholds can be based on information external to the decision problem (cf. Van de Poel 2009). They may, for example, be based on technical codes and standards. This indeed happened in the refrigerants case discussed above: the thresholds for both toxicity and flammability were based on equivalence classes and thresholds that were defined in relevant codes and standards such as the ASHRAE Code for Mechanical Refrigeration.[15]

---

[15]For a more detailed discussion, see the chapter "▶ Design for Values and the Definition, Specification, and Operationalization of Values."

Thresholds may also be based on the law or company policy. One particular interesting possibility is to base thresholds on moral obligations. Earlier, I suggested that so-called taboo trade-offs may be due to the fact that some values can, for moral reasons, not be traded off below a certain threshold, as meeting a threshold corresponds with some moral obligation. In such cases, thresholds may thus be based on moral obligations (although it may be hard to define exactly where the threshold is between meeting and not meeting a moral obligation). So applied, the satisficing decision rule has the big advantage that it avoids the choice of a morally unacceptable alternative.

It should be noted that if thresholds are based on external information, it is likely that in many cases, satisficing will not lead to the selection of just one alternative. Especially if more than one alternative is still considered acceptable, between which external thresholds cannot decide, it seems most reasonable to use one of the other discussed approaches. In that sense, satisficing is an approach that is maybe best combined with other approaches.

Finally, just like the maximin rule, satisficing may sometimes result in (seemingly) irrational results. Suppose we have a refrigerant that meets the thresholds for safety (e.g., expressed in LFL) and environmental sustainability (e.g., expressed in GWP). Now suppose we find another refrigerant with a much lower GWP (so much better in terms of environmental sustainability) and a little worse in LFL (i.e., in terms of safety). Now also suppose that the decrease in safety is small but big enough to just fall below the threshold. Satisficing with the given thresholds now tells us that the second option should never be preferred to the first (as it does not meet all thresholds) whatever the gain in terms of environmental sustainability. Again, at least occasionally, this seems the wrong advice.

Judgment: Conceptualization and (Re)specification

We will now look at an approach that emphasizes judgment and reasoning about values. This approach aims at conceptualizing and (re)specifying the values that underlie the conflicting design criteria. The advantage of this approach is that it might solve a value conflict while still doing justice to the conflicting values and without the need to make the values commensurable or to define thresholds for them.

The first thing to do when one wants to exercise judgment in cases of trade-offs is to identify what values are at stake in the trade-off and to provide a conceptualization of these.[16] What do these values imply and why are these values important? Take the value of freedom in the case of safety belts. Freedom can be conceptualized as the absence of any constraints on the driver; it then basically means that people should be able to do what they want. Freedom can, however, also be valued as a necessary precondition for making one's own considered choices; so conceived freedom carries with it a certain responsibility. In this respect, it may be argued that

---

[16]This paragraph draws from Van de Poel (2009).

a safety belt that reminds the driver that he has forgotten to use it does not actually impede the freedom of the driver but rather helps him to make responsible choices. It might perhaps even be argued that automatic safety belts can be consistent with this notion of freedom, provided that the driver has freely chosen to use such a system or endorses the legal obligation for such a system, which is not unlikely if freedom is not just the liberty to do what one wants but rather a precondition for autonomous responsible behavior. One may thus think of different conceptualizations of the values at stake, and these different conceptualizations may lead to different possible solutions to the value conflict. Some conceptualizations might not be tenable because they cannot justify why the value at stake is worthwhile. For example, it may be difficult to argue why freedom, conceived of as the absence of any constraint, is worthwhile. Most of us do not strive for a life without any constraints or commitments because such a life would probably not be very worthwhile. This is not to deny the value of freedom; it suggests that a conceptualization of freedom only in terms of the absence of constraints misses the point of just what is valuable about freedom. Conceptualizations might not only be untenable for such substantial reasons, they may also be inconsistent or incompatible with some of our other moral beliefs.

To make values operative in design, they do not only need to be conceptualized but also to be *specified*.[17] Specification is the translation of a general value or norm into more specific design requirements. The requirement can be more specific with respect to (a) scope of applicability of the norm, (b) the goals or aims strived for, and (c) actions or means to achieve these aims (cf. Richardson 1997, p. 73). An example is the specification of the value of safety into the following design requirement: "minimize the probability of fatal accidents (specification of the goal) when the chemical plant is operated appropriately (specification of the scope) by adding redundant safety valves (specification of the means)." In this case, the design requirement specifies the general norm in three dimensions, but specification may also be restricted to one or two dimensions.

A specification substantively qualifies the initial value or norm by adding information "describing what the action or end is or where, when, why, how, by what means, by whom, or to whom the action is to be done or the end is to be pursued" (Richardson 1997, p. 73). Obviously, different pieces of information may be added so that a general value or norm can be specified in a large multiplicity of ways. Not all specifications are adequate or tenable, however. In general, one would want to require that actions – or in our case, designs – that count as satisfying the specific design requirements also count as satisfying the general value or norm (cf. Richardson 1997, pp. 72–73). In the above example, "safety" is specified as "minimizing the probability of fatal accidents." This specification is adequate if in all cases in which the probability of fatal accidents is minimized, safety is maximized. Now arguably, safety encompasses not only avoiding or at least minimizing fatal accidents but also avoiding or minimizing accidents in which people get hurt

---

[17]This and the next paragraph draw from Van de Poel (2013).

but do not die. This does not make the specification necessarily inadequate, however. Maybe, it is known on the basis of statistical evidence, for example, that in this type of installation, there is a strict correlation between the probability of fatal accidents and the probability of accidents only leading to injuries, so that minimizing the one implies minimizing the other. In that case, the specification may still be adequate. In other situations, it may be inadequate and it might be necessary to add a design requirement related to minimizing nonfatal accidents.

Usually, more than one specification of a value will be tenable. This offers opportunities for dealing with value conflict in design. Value conflicts in design are in practice always conflicts between *specifications* of the values at stake because abstract values as such are too general and abstract to guide design or to choose between options. So if there is room for different possible specifications of the values at stake, it might be possible to choose that set of the specifications of the various values at stake that are not conflicting. Sometimes, it will only become apparent during the design process, when the different options have been developed and are compared that certain specifications of the values at stake are conflicting. In such cases, it may sometimes be possible to respecify the values at play so as to avoid the value conflict.

An interesting example of respecification took place in the refrigerant example 3.[18] In the first instance, the industry preferred HFC134a as alternative to CFC 12, basically following the satisficing reasoning as explained in the previous section (see also Table 6). However, environmental groups were against this alternative as they viewed the threshold for environmental sustainability (at least one H atom) too lenient, especially because it resulted in much higher GWPs (global warming potentials), than if a flammable coolant was chosen. At some point, Greenpeace succeeded in convincing a former East German refrigerator producer of using a flammable coolant in its new design. The refrigerator was also able to acquire the safety approval of the German certification institute TUV. Following the success of this refrigerator, German and later other European refrigerator producers also switched to flammable coolants like propane and isobutane. Such coolants were seen as acceptable despite their flammability because a new specification of safety was developed. Where safety was first specified as nonflammability of coolants, it now came to be specified as a low explosion risk of the whole refrigerator. It turned out to be possible to achieve a low explosion risk even with flammable coolants.

Although it might be possible to solve a value conflict in design through respecification, this will not always be possible. Even in cases in which it is possible, it may not always be desirable. It may especially not be desirable if respecification leads to a serious weakening of one of the values compared to the original specification (Hansson 1998). Still, solving a value conflict through respecification does not necessarily or always imply a weakening of one of the values (Van de Poel forthcoming).

---

[18]Based on van de Poel (2001). See also the chapter "▶ Design for Values and the Definition, Specification, and Operationalization of Values."

Innovation

The previous approach treats the occurrence of value conflict merely as a philosophical problem to be solved by philosophical analysis and argument. However, in engineering design value conflicts may also be solved by technical means. That is to say, in engineering it might be possible to develop new, not yet existing, options that solve or at least ease the value conflict. In a sense, solving value conflicts by means of new technologies is what lies at the heart of engineering design and technological innovation. Engineering design is able to play this part because most values do not conflict as such, but only in the light of certain technical possibilities and engineering design may be able to change these possibilities. An interesting example is the design of the storm surge barrier in the Eastern Scheldt estuary in the Netherlands (example 2) that eased the value conflict between safety and ecology.

The reason why technical innovation can ease value conflicts is that it enlarges the feasibility set. This is a clear advantage of this approach. Often, however, technical innovation will not lead to *one* option that is clearly better than all others, so that choices between conflicting values still have to be made. In this respect, innovation only presents a partial solution to value conflicts in design.

According to van den Hoven et al., "technical innovation results in moral progress in those cases in which it means an improvement in *all* relevant value dimensions" (Van den Hoven et al. 2012, p. 152). Of course, not all technical innovations imply an improvement in all relevant value dimensions. Sometimes, a gain in one value dimension comes at the cost of a loss in another value dimension. Sometimes, the technical innovation creates new problems or side effects, which require new value dimensions to be taken into account. Sometimes, the technical innovation only addresses the initial problem in as far as it is amendable to a technological solution. It might also be that the values themselves change due to technical development; an often mentioned example is the change in sexual morality due to the development of anticonceptives. Technical innovation may also create new choices and dilemmas, as in the case of prenatal diagnosis, that we do not want to have.

Pointing at technical innovation as a way to deal with value conflicts does not yet make clear how to develop the kind of innovations that actually eases value conflicts. One approach here may be to translate the values into more specific design requirements that can guide design (Van de Poel 2013). Another interesting approach is that of value dams and value flows that have been proposed in VSD (see chapter "▶ Value Sensitive Design: Applications, Adaptations, and Critiques"). A value dam is a technical feature that is (strongly) opposed by one or more stakeholders because it conflicts with important values; a value flow is a technical feature that is for value reasons supported by various stakeholders. So, a value dam tells where not to go in innovation, while a value flow suggests technical features that should be included. In the case of the Eastern Scheldt, a design feature like "complete closure" can be associated with a value dam given the strong opposition from environmental groups, but also the design feature "no dam" met strong opposition from the government agency Rijkswaterstaat and therefore also

constituted a value dam. The design features "half-open" and "flexibly open/closed" on the other hand constituted value flows as they allowed meeting both the values of safety and ecology.

Comparison of Methods and Conclusion

Above, I discussed six methods for dealing with value conflict in design. We saw that each method has its pros and cons; this has been summarized in Table 7. As the table shows, no method in isolation provides a complete solution to the problem of value conflict in design. It might, however, be possible to combine methods and so achieve an acceptable procedure for dealing with value conflict in design.

In particular, different methods may be required for value conflicts that amount to moral dilemmas than for value conflicts that do not entail a moral dilemma. As we have seen in section "Value Conflict in Engineering Design," not all value conflicts entail moral dilemmas but some do. Value conflicts amount to a moral dilemma if there are values at stake that correspond to moral obligations that cannot be simultaneously met. I have above suggested that such moral obligations may be characterized as a minimal threshold that should be met on each (or at least some) of the relevant value dimensions. Figure 3 represents this idea. For each of the values, a minimal threshold has to be met to meet moral obligations. We can now define the moral opportunity set as the set of options that is feasible and meets all minimal thresholds. If the moral opportunity set is empty, we are confronted with a moral dilemma.

It should be noted that satisficing might help to ensure the choice of an option within the moral opportunity set if that set is nonempty, although it cannot choose between options within the moral opportunity set. If the moral opportunity set is empty, innovation is a particularly attractive option because, as suggested by Van den Hoven et al. (2012), it may make the moral opportunity set nonempty.

On the basis of these ideas, I want to end with a particular suggestion of a stepwise approach that combines the methods for cases of conflicting *moral* values:

1. *Satisficing with moral obligations*. The goal of this step is to rule out morally unacceptable options. To do so, one looks for moral obligations that correspond to the relevant moral values and judges whether these correspond with (minimal) thresholds to be met by those values.

   This step requires *judgment* in order to identify the moral obligations and to define the corresponding thresholds. It is important to focus on moral obligations rather than on other external constraints that may (also) set threshold values because these other constraints may not be morally desirable. It is also advisable to focus on clear and uncontroversial moral obligations as this step is meant to rule out clearly morally unacceptable options.

   For setting the thresholds, it is also advisable to define tangible attributes for the values that can be assessed in design. This will require a specification of the relevant values.

Table 7 Overview of methods for dealing with value conflicts in design

Method	How are the values weighted?	Main advantages	Main disadvantages
Cost-benefit analysis	All values expressed in monetary terms	Values are made commensurable on a common scale so that the best option can be chosen	Requires ratio measurability and value commensurability to avoid Arrow's theorem
			Various fundamental and methodological problems with expressing values in money
Direct trade-offs	Trade-offs between the different values	"Best" option chosen ("optimal design")	Requires unit commensurability to avoid Arrow's theorem
			Making trade-offs may not always be (morally) acceptable
Maximin	Comparison between scores on various values	Selects a design that scores best on worst criterion ("robust design")	Requires level commensurability to avoid Arrow's theorem
			May sometimes lead to (seemingly) irrational choices
Satisficing	A threshold is set for each value	The selected alternatives meet the thresholds	How to set thresholds in a rational and acceptable way?
		Can help avoid morally unacceptable alternatives	Subject to Arrow's theorem: criterion of collective rationality not met
		No direct trade-off between the criteria	May sometimes lead to (seemingly) irrational choices
Judgment	Conceptualization and specification of the various values	Might solve value conflict by judgment and (re) specification	Not all value conflicts can be solved in this way
			Respecification may lead to an unacceptable weakening of moral obligations
Innovation	Not applicable	Can lead to alternatives that are clearly better than all of the present alternatives	Does not solve the choice problem in many cases

After satisficing with the so-defined thresholds, the moral opportunity set may either be empty or not. If it is empty, innovation may be required to look for solutions that make the moral opportunity set nonempty. If the moral opportunity set is nonempty, innovation is still advisable because it might be possible to find better options than the currently available. In both cases, the next step therefore is:

Fig. 3 The moral opportunity set. V_a and V_b are the values at stake, and t_a and t_b the minimal thresholds corresponding to moral obligations for these values. The blank area left of the feasibility frontier is the feasibility set. Note that depending where the feasibility frontier is, the moral opportunity set may be empty, in which case, we are confronted with a moral dilemma (The figure is based on Van den Hoven et al. (2012))
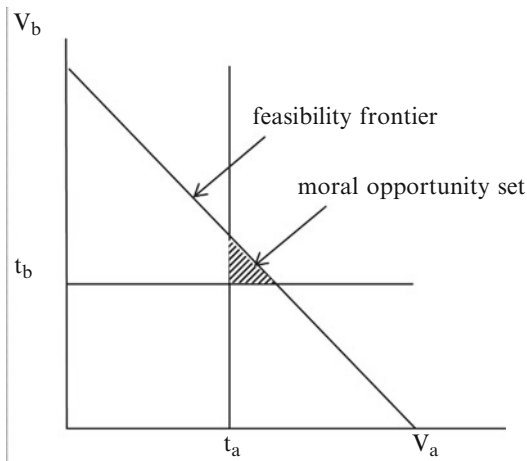
2. *Innovation.* The goal of this step is to develop new options that better meet the relevant values. Doing so might require a further *specification* of the relevant values in order to be better able to develop new options that better meet those values. Also, other VSD tools like value dams and value flows might be helpful here.

   After this step, there are roughly three possibilities: (1) the moral opportunity set is (still) empty and one should choose whether to design a product or not, (2) the moral opportunity set is nonempty and contains exactly one option, and (3) the moral opportunity set is nonempty and contains more than one option. In the second case, no further step is required. In the first and third cases, the next step is a choice, but as the nature of the choice is somewhat different, I differentiate between two versions of the next step.

3a. *Choice (if the moral opportunity set is empty).* As there is no design option that meets all relevant moral obligations, one should wonder whether to design a product or not. Depending on the specific case, it might be possible to solve the moral dilemma through respecification, i.e., in such a way that there is a design that meets all minimal moral thresholds. However, one should take care not to unacceptably play down moral obligations or values in respecification.

3b. *Choice (if the moral opportunity set contains more than one option).* If the moral opportunity set contains more than one option, a choice has to be made between these options. The most appropriate approaches for doing so are cost-benefit analysis, direct trade-offs, and maximin, as the other three approaches do usually not narrow down the choice to one option. Of these three approaches, direct trade-offs often seem to me the most desirable. The reasons for this are as follows. In our case, the options among which a choice needs to be made all meet minimal thresholds set by moral obligations (as a consequence of step 1); this makes, as earlier suggested, trade-offs between values in most cases acceptable. Compared to cost-benefit analysis, direct trade-offs have the advantage of being less informationally demanding, and it lacks the

disadvantages that come with expressing values in monetary units. If we compare maximin with direct trade-offs, one might think that maximin is less informationally demanding; on the other hand, as we have seen, the level commensurability that is required for maximin requires quite complicated judgments. Moreover, maximin may occasionally lead to (seemingly) irrational choices.

I do not want to suggest that the above approach is the only way to combine the various methods for dealing with conflicting values in a reasonable way. There may be other ways of doing it. My proposal also specifically is meant for conflicts between *moral* values, and value conflicts between nonmoral values (or between moral and nonmoral values) may be better treated differently. Moreover, I have been assuming in this contribution that the designer takes a moral point of view. This assumption may not always be realistic, and even from a moral point of view, the designer may not always be required to do what is morally best, as it may be good enough to choose an option that is morally acceptable but perhaps not morally best.

Cross-References

▶ Design for the Value of Safety
▶ Design for the Value of Sustainability
▶ Design for Values and the Definition, Specification, and Operationalization of Values
▶ Value Sensitive Design: Applications, Adaptations, and Critiques

References

Arrow KJ (1950) A difficulty in the concept of social welfare. J Polit Econ 58:328–346
Arrow KJ, Raynaud H (1986) Social choice and multicriterion decision-making. MIT Press, Cambridge, MA
ASHRAE (2013) Designation and safety classification of refrigerants, ANSI/ASHRAE standard, 34-2013. ASHRAE, Atlanta
Ball LJ, Evans JSBT, Dennis I (1994) Cognitive processes in engineering design: a longitudinal study. Ergonomics 37(11):1753–1786
Baron J, Spranca M (1997) Protected values. Organ Behav Hum Decis Process 70(1):1–16
Boyce RR, Brown TC, McClelland GH, Peterson GL, Schulze WD (1992) An experimental examination of intrinsic values as a source of the WTA-WTP disparity. Am Econ Rev 82(5):1366–1373. doi:10.2307/2117484
Franssen M (2005) Arrow's theorem, multi-criteria decision problems and multi-attribute preferences in engineering design. Res Eng Des 16:42–56
Hansson SO (1998) Should we avoid moral dilemmas? J Value Inq 32(3):407–416. doi:10.1023/a:1004329011239
Hansson SO (2007) Philosophical problems in cost–benefit analysis. Econ Philos 23:163–183
Horowitz JK, McConnell KE (2002) A review of WTA/WTP studies. J Environ Econ Manag 44(3):426–447. doi:10.1006/jeem.2001.1215

Hylland A (1980) Aggregation procedure for cardinal preferences: a comment. Econometrica 48(2):539–542. doi:10.2307/1911117

Jacobs JF, Van de Poel I, Osseweijer P (2014) Clarifying the debate on selection methods for engineering: Arrow's impossibility theorem, design performances, and information basis. Res Eng Des 25(1):3–10

Keeney RL (2002) Common mistakes in making value trade-offs. Oper Res 50(6):935–945. doi:10.2307/3088614

Kroes P, Franssen M, Bucciarelli L (2009) Rationality in design. In: Meijers A (ed) Philosophy of technology and engineering sciences, vol 9, Handbook of the philosophy of science. Elsevier, Oxford, pp 565–600

Levi I (1986) Hard choices. Decision making under unresolved conflict. Cambridge University Press, Cambridge

Mas-Colell A, Sonnenschein H (1972) General possibility theorems for group decisions. Rev Econ Stud 39(2):185–192. doi:10.2307/2296870

May KO (1954) Intransitivity, utility, and the aggregation of preference patterns. Econometrica 22(1):1–13

McLinden MO, Didion DA (1987) Quest for alternatives. ASHRAE J 29(11):32–42

Raz J (1986) Value incommensurability: some preliminaries. Proc Aristot Soc 86:117–134

Richardson HS (1997) Practical reasoning about final ends. Cambridge University Press, Cambridge

Roberts KWS (1980) Interpersonal comparability and social choice theory. Rev Econ Stud 47(2):421–439

Sen AK (1970) Collective choice and social welfare. Oliver & Boyd, Edinburg/London

Simon HA (1955) A behavioral model of rational choice. Q J Econ 69:99–118

Simon HA (1956) Rational choice and the structure of the environment. Psychol Rev 63:129–138

Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds) (2007) Climate change 2007: the physical science basis: contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge/New York

Tetlock PE (2003) Thinking the unthinkable: sacred values and taboo cognitions. Trends Cogn Sci 7(7):320–324

Tinbergen J (1959) Sociaal-economische aspecten van het Deltaplan. In: Deltacomissie (ed) Rapport Deltacommissie, Bijdrage VI, Onderzoekingen van belang voor het ontwerpen van dijken en dammen. Staatsdrukkerij- en Uitgeverijbedrijf, Den Haag, pp 61–74

Tsui K-Y, Weymark JA (1997) Social welfare orderings for ratio-scale measurable utilities. Econ Theory 10(2):241–256. doi:10.1007/s001990050156

van de Poel I (1998) Changing technologies. A comparative study of eight processes of transformation Of technological regimes. University of Twente, Enschede

van de Poel I (2001) Investigating ethical issues in engineering design. Sci Eng Ethics 7(3):429–446

Van de Poel I (2009) Values in engineering design. In: Meijers A (ed) Philosophy of technology and engineering sciences, vol 9, Handbook of the philosophy of science. Elsevier, Oxford, pp 973–1006

Van de Poel I (2013a) Translating values into design requirements. In: Mitchfelder D, McCarty N, Goldberg DE (eds) Philosophy and engineering: reflections on practice, principles and process. Springer, Dordrecht, pp 253–266

Van de Poel I (forthcoming) Dealing with moral dilemmas through design. In: Van den Hoven J, Miller S, Pogge T (eds) The design turn in applied ethics. Cambridge University Press, Cambridge

Van de Poel I, Royakkers L (2011) Ethics, technology and engineering. Wiley-Blackwell, Oxford

Van den Hoven J, Lokhorst G-J, Van de Poel I (2012) Engineering and the problem of moral overload. Sci Eng Ethics 18(1):143–155. doi:10.1007/s11948-011-9277-z

Williams B (1973) Problems of the self. Philosophical papers 1956–1972. Cambridge University Press, Cambridge