VALUES IN ENGINEERING DESIGN

Ibo van de Poel

1 INTRODUCTION

Value is at the heart of engineering design. Design creates value for companies, users and, ultimately, for society. Few would disagree with such statements but what exactly do they imply? What is value? What types of values do technological artifacts have or contribute to? How are value considerations inherent to design choices? Is engineering design plagued by plural and conflicting values and, if so, how do, could and should engineers deal with such value conflicts? These are the types of questions that are explored in this contribution.

The aim of this contribution is to philosophically explore the role of values in engineering design. Although, where relevant, I will draw on empirical evidence on the role of values in design, my aim will not be to merely empirically describe how values come to play a part in engineering design. Instead I shall aim to clarify, from a philosophical point of view, the role that values do, can and — according to some — should play in engineering design. I will not defend any specific approach to integrating values into design though I will discuss various approaches, especially in relation to conflicting design values together with the pros and cons of these approaches.

The focus in this contribution will be on the values that are created through technical artifacts and, especially, on how the prospect of such value is integrated into the process of engineering design. I am not therefore focusing on the values, or virtues, of designing engineers (for this, see Pritchard's chapter in this Volume, Part V) or on the value of engineering design as an activity (for engineers design may be an inherently valuable activity).

This contribution starts with a brief overview of relevant notions from philosophical literature, like the notion of value. I shall then discuss the value of technological artifacts. In Section 4, I will discuss the role of values in the engineering design process. Sections 5 and 6 will deal with value conflicts in engineering design and approaches to dealing with them; Section 5 examines optimizing approaches and Section 6 looks at non-optimizing approaches. In the final section, I shall draw my conclusions.

Handbook of the Philosophy of Science. Volume 9: Philosophy of Technology and Engineering Sciences.
Volume editor: Anthonie Meijers. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

2 VALUES

2.1 Value statements

Value statements are statements about whether certain things or state of affairs are good, i.e. valuable, or bad in a certain respect. If things or states of affairs are bad, they often not only lack value but also even have a negative value. I will use the term *disvalue* to refer to such negative value.

Value statements are to be distinguished from statements of preference, i.e. statements about what individuals prefer. Establishing that something is a value or professing it to be valuable means not only claiming that it is valuable to me but also that it is or should be of value to others.

Most things or states of affairs are not just plain good but good in a certain sense. A hammer may be good in the sense of being a useful instrument for driving nails into a piece of wood but it may simultaneously be aesthetically ugly. We might then say that it has a utility value with respect to driving nails into a piece of wood but that it has no, or merely a negative, aesthetic value. So conceived, values are varieties of goodness. Things or states of affairs may thus have — express, instantiate — different values at the same time.

Statements about the value of things or state of affairs are evaluative statements: they evaluate something or a state of affairs in terms of a value. Value statements are therefore to be distinguished from descriptive and prescriptive statements.

Descriptive statements are statements about how things are. We can, for example, describe a hammer in terms of its shape and the materials of which it is made; such descriptions do not attribute value to the object described. Still, the descriptive, non-evaluative, features of a hammer are relevant to answering the question of whether the hammer is a good hammer, in the sense of being useful for driving nails into a piece of wood. The value of a hammer as a useful instrument thus depends on the descriptive features of that same hammer. It can be expressed by saying that the value of a thing or state of affairs is "supervenient" on certain non-evaluative features of that thing or state of affairs [Zimmerman, 2004].

Value statements are also to be distinguished from prescriptive statements, i.e. statements about how to act (see, e.g. [Stocker, 1990; Dancy, 1993]). The statement "this is a good hammer for driving nails into a piece a wood" does not entail the statement "you should use that hammer to drive a nail into a piece of wood" (for a detailed discussion of the normative aspect of statements like these, see Franssen's chapter in this Volume, Part V). This is not to deny that prescriptive statements may sometimes be derived from evaluative statements. In general, however, one should distinguish evaluative from prescriptive statements. In moral theory, there is a parallel distinction between "goodness" and "rightness." States of affairs and in particular consequences of actions can be evaluated in terms of "goodness," while actions themselves are evaluated in terms of "rightness." Consequentialism is the doctrine which states that the goodness of consequences of actions determines the rightness of actions. However, consequentialism is certainly not the only theory in moral philosophy on the relation between goodness and rightness. If one takes a Kantian approach, for example, the rightness of an action in a certain sense determines or guarantees the goodness of the consequences of that action (cf. [Korsgaard, 1983, p. 183]).

2.2 Intrinsic versus instrumental value

Often a distinction is made between intrinsic and instrumental values. Intrinsic values are those that are good in themselves or for their own sake, while instrumental values are valuable because they help to achieve other values. It should be noted that in this respect an object can be instrumentally valuable and intrinsically valuable at the same time. A hammer may, for example, be instrumentally valuable as a useful tool for driving nails into wood while at the same time being intrinsically valuable as a beautiful object.

Although the distinction between instrumental and intrinsic value may seem straightforward, it is not. Various philosophers have pointed out a number of terminological and substantive issues with respect to the distinction [for a discussion, see Zimmerman, 2004]. One issue is that the notion of intrinsic value is ambiguous. The notion is usually understood to refer to objects or states of affairs that are valuable in themselves. Intrinsic value is then value of a non-derivate kind. Intrinsic value may, however, also refer to things that are valuable due to their intrinsic natural, i.e. descriptive, properties. As Christine Korsgaard has pointed out things that are valuable due to their intrinsic properties are unconditionally good [Korsgaard, 1983]. Their goodness does not depend on the relation with other objects or with people; otherwise their value would not be intrinsic to the object. However, according to Korsgaard, some things may be good in a non-derivate sense, even if they are not unconditionally good. An example is human happiness in a Kantian respect. According to Kant, human happiness is non-derivate goodness. Happiness is good in itself, and not because it is a means to another end or contributes to another value. Nevertheless, according to Kant, happiness is only conditionally good; it is only good insofar as it corresponds to good will, i.e. respect for the moral law. To avoid the ambiguity to which Korsgaard refers, I will use the notion of 'intrinsic value' for non-derivate value and the notion of 'value intrinsic to an object or state of affairs' to describe value that only depends on the non-relational properties of an object or state of affairs.

The notion of instrumental value is also more complex than it seems. It might refer to things that are useful for achieving some end, whether that end is valuable or not. Frankena [1973, p. 66] refers to such instrumental values as utility values. He puts forward the notion of extrinsic value to refer to "things that are good because they are a means to what is good" [Frankena, 1973, p. 66]. However, the term 'extrinsic value' is somewhat confusing because it is often used for all derivate values, i.e. all non-intrinsic values. Being a means to an end is, however, not the only way in which something can be valuable in a derivate way (see, e.g.

[Zimmerman, 2004]).[1] Things can, for example, also be valuable because they *enable* a good life, just as privacy or health enable people to live a valuable life or because they *contribute* to a good life as do the virtues in an Aristotelian account of good life. I will therefore use the notion of *instrumental value* for the value of being a means to achieving a good end, i.e. another positive value.

2.3 Sources of value

What of the origins of value, i.e. what is it that brings value into the world? Here it is relevant to distinguish between intrinsic and extrinsic value. Extrinsically valuable objects or states of affairs derive their value from their contribution — for example as a means to an end — to intrinsic value. But where does intrinsic value come from? Three main doctrines can be distinguished here. Subjectivists believe that human desires, or more generally psychological states, constitute the source of value. Objectivists believe that values reside in the world outside us. Rationalists see human rationality as the ultimate source of value. Subjectivism does justice to the connection between values and human desires and interests. It runs, however, the risk of confusing value with preference. Not everything that is desired or preferred by people is valuable. Objectivism does justice to the fact that statements about value are not statements about preferences but rather about how the world is or ought to be from a normative point of view. Objectivists often hold that intrinsic value is value intrinsic to the valuable object (e.g. [Moore, 1912]). Since value that is intrinsic to an object cannot depend on the relational properties of that object human desires or interests can never be the, or even a, source of value. Rationalism can be seen as an in-between position. It restores the connection between human desires and values, which is lost in objectivism but strives to avoid confusing value with preference by claiming that things are valuable not just because people prefer them but because rational beings have sufficient practical reason to pursue them.[2]

With most of the issues discussed in this contribution it will not be necessary to take a stance in the debate about the source of value. The stance that I will take, where relevant, might be described as mildly rationalist. It is rationalist in the sense that both extreme subjectivist and extreme objectivist positions are avoided. I believe that values should be distinguished from preferences but not completely divorced from human desire, interest, interpretation and meaning-giving. I will call the position mildly rationalist because I will not assume one specific theory of practical rationality, nor will I make substantive assumptions about the ultimate sources of value in the world.

---

[1] The point is not that the instrumental value may be insufficient to cause the end but that extrinsic values may contribute to intrinsic values in non-causal ways. They may, for example, be an indication of the achievement of an intrinsic value or they may be conceptually part of the intrinsic value (e.g. health and the good life).

[2] Nevertheless, rationalism may make it difficult to express the fact that part of the reason for valuing an object may lie in the object itself.

2.4 Value conflicts and incommensurable values

A final general issue that is relevant to the discussion on the role of values in engineering is the notion of value conflict and value incommensurability. I speak of there being value conflict if two or more values conflict. Usually values will not generally conflict although that can happen but only in specific circumstances. Two or more values conflict in a specific situation if, when considered in isolation, they evaluate different options as best. For example, a governmental policy that is best from the viewpoint of individual freedom might not be best in terms of social justice. As we shall see, value conflict situations are common in engineering, especially in the engineering design process.

Two or more values are incommensurable if they cannot be expressed or measured on a common scale or in terms of a common value measure. This can be an ordinal, an interval or a ratio scale. There is no agreement in the philosophical literature on the implications of the incommensurability of values for the comparability of objects or options (for a discussion, see [Chang, 1997; Hsieh, 2007]). It should be noted that the incommensurability of values does not as such entail the incomparability of options. For example, if all relevant values evaluate the same option as best in the absence of value conflict then incommensurability will not entail incomparability. In situations in which a lexicographical ordering of values exists, that is in situations where any small amount of one value is worth more than any large amount of another value, incommensurability will not lead to incomparability. We can then 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.

Even in cases of conflicting incommensurable values, between which no lexicographical ordering exists, there is no agreement on whether this is enough to entail the incomparability of options. In such cases the strongest argument for incomparability is probably the argument from small improvements [Raz, 1986; Chang, 1997] as illustrated in the following example. Suppose that one needs to choose between a career in engineering and a career in philosophy. Suppose too that there is good reason to judge that neither the philosophy nor the engineering career is better (i.e. more valuable overall) than the other. It is possible that we will still find that the philosophy career is not better than the engineering career, even if the former is slightly improved. Conversely, the engineering career may not be better than the philosophy career even if it is slightly improved. If this were the case then the two careers could not be said to be equally good because if that were so a slight improvement in one of the two would make it better than the other. Therefore the two options — a career in engineering and a career in philosophy — may in this case be said to be incomparable, because neither is better than the other and they are not equally good.

The argument from small improvements can be interpreted in at least two, not necessarily mutually exclusive, ways. One solution is to interpret the incomparability as vagueness [Broome, 1997]. With this interpretation, on the scale of overall goodness, options can be surrounded by a zone of mild indeterminacy; in this zone, the options are neither better nor worse than each other nor are they equally good. Another interpretation is to argue that the two careers are “roughly equal” or “on a par.” The latter has been proposed as a fourth comparative relation in addition to “better than,” “worse than,” and “equally good” [Chang, 1997]. With both interpretations, it is debatable whether they can erase incomparability between options. But even if they do not, both interpretations limit the extent of incomparability. Even if a specific career in engineering is not comparable with one in philosophy, a particularly good career in engineering may still be considered to be better than a particularly bad career in philosophy.[3]

In the argument given above, incommensurability has mainly been seen as a formal feature, as the impossibility to express or measure two values on a common scale. Some authors put forward a more substantive notion of incommensurability. Raz [1986], for example, has suggested that resistance to certain trade-offs is constitutive of certain values or goods. 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.

It has also been suggested that some values may resist trade-offs because they are ‘protected’ or ‘sacred’ [Baron and Spranca, 1997; Tetlock, 2003]. This seems especially true of moral values and values that regulate the relations between, and the identities of, people. One way to understand this phenomenon is by considering the notion of moral residue. Moral dilemmas — i.e. situations in which an agent is confronted with conflicting moral obligations and in which at least one moral obligation is not met whatever option one chooses — leave the agent with a moral residue or with a sense of guilt arising from the obligation not met [Williams, 1973; Marcus, 1980]. Similarly, trade-offs between protected values may create an irreducible loss because a gain in one value may not always 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. The point here is not that the value of friendship is lexicographically ordered over that of money or having a career (probably it is not). Even if a lexicographical ordering is absent, the nature of the values is so different that one cannot compensate for the other.

Some philosophers have denied the existence of value incommensurability. They believe that all values can ultimately be expressed in terms of one overarching or super value. Utilitarianism often attributes such a role to the value of human happiness, but a similar role may be played by the value of the ‘good will’ in Kantianism or the value of ‘contemplation’ in Aristotelian ethics [Korsgaard, 1986].[4]

---

[3] Such comparisons are known as nominable/notable comparisons.

[4] The way in which such super values operate may, however, be quite different in different approaches. In Kantianism, it is ‘good will’ that solves value conflicts by means of (practical) reasoning, while in utilitarianism value conflict is solved by expressing all values in terms of the super value ‘utility’.

The notion that there is ultimately only one value that is the source of all other values is known as value monism. Value monists do not necessarily deny the existence of more than one value but they do consider other values to be only conditionally good. A value monist believes that value conflicts can essentially be solved by having recourse to the super value. Nevertheless, value monists might recognize that there are practical or epistemological limitations to solving value conflicts. A utilitarian might, for example, believe that all values are only good in so far as they contribute to human happiness but she might recognize that in practice we cannot always exactly determine how much a specific value contributes to human happiness, so that some options might have an indeterminate ranking.

3 TECHNOLOGY AND VALUES

Sometimes the thesis of technology being value-neutral is defended [Florman, 1987; Pitt, 2000]. The main argument usually given for this thesis is that technology is just a neutral means to an end which can be put to good or bad use. Value is thus created during use and is not located in technology. This also means that the objectionable effects of technology are to be blamed on the users and not on technological artifacts, or their designers. As the American National Rifle Association has expressed it: "Guns do not kill people, people kill people".

What does claiming that technology is value-neutral exactly entail? One interpretation would be to say that it means that the value of technological artifacts *only* depends on their extrinsic properties. In this interpretation, the thesis that technology is value-neutral is clearly false. It can be seen as follows. Technological artifacts have a physical or material component, in other words they are also physical objects, even if they are not mere physical objects. The value of physical objects as a means to an end depends — partly at least — on their intrinsic properties. A stone can be used to split a nut thanks to its intrinsic physical properties. A tree leaf would have a much smaller or no utility value when it comes to splitting nuts. Since it is implausible that the utility value of physical objects merely depends on their extrinsic properties, the same may be said of technologies.[5] So the value of technological artifacts does not only depend on their extrinsic properties.

The thesis that value is not intrinsic to technology may also be interpreted as implying that such value also *partly* depends on the extrinsic properties of a technology. To judge the plausibility of such a claim, it is crucial to define technology or technological artifacts because to a large extent that is what will determine what we consider to be the intrinsic and extrinsic properties of technological artifacts. Radder (this Volume, Part V), for example, presents a definition of technology in which the actual realizability of the function of a technological artifact is part of what technology actually is. This seems to make the actions of users internal to

---

[5] It might be remarked that there is also such a thing as non-physical technologies, like software programs. It seems obvious that the utility value of such non-physical objects should also partly depend on the intrinsic properties of such objects, e.g. the intrinsic properties of the software program in question.

technology which, in turn, makes it more likely that at least some values will only depend on the intrinsic properties of technologies. In fact if we define technology sufficiently broadly, we can always make values internal to technology, it seems. But what happens if we start off with a minimal definition of technology? I think that any plausible minimal account of technology needs to refer to the notion of function, and/or comparable notions like ends, purposes and intentions. The fact technologies have a function implies that they have utility value, i.e. that they can be used for some end. On a minimal definition of technology, then, technology at least has utility value. This does not mean that such utility value is intrinsic to technological artifacts in the sense that it only depends on the intrinsic properties of technological artifacts. That, indeed, is not usually the case: the particular utility value of a particular hammer for driving nails into a piece of wood also depends, for example, on the physical abilities of users and such abilities are extrinsic to the hammer. So even if having utility value is part of what it means to be a technical artifact, that same utility value is not necessarily intrinsic to the technological artifact. We might express this by saying that having utility value is conceptually *inherent* to the notion of a technological artifact. The presence of such inherent values in technology seems to be a good reason for rejecting the thesis that technology is value-neutral. A further question is what types of value are precisely inherent to technology. I will answer this question below.

3.1 Instrumental value

Whereas utility value is the value of being a means to any end, instrumental value is the value of being a means to a *good* end, of being a means to a positive value. A popular thesis is the one that holds that the instrumental value of technological artifacts, so defined, depends on the goals for which the artifact is *used.* Users may implement technology for good or bad. A macabre example of use for bad ends is the gas chambers designed by German engineers that were used to effect the planned extermination of the Jews held in the concentration camps during the Second World War. In this case the gas chambers clearly contributed to the morally objectionable goal of eradicating the Jews. So conceived, the gas chambers were a source of *dis*value. They had a negative instrumental value but at the same time a positive utility value: they were useful for Nazi purposes. One might even argue that the larger the utility value of the gas chambers was, the greater their negative instrumental value would be.[6]

The distinction between utility value and instrumental value leads one to question how technological artifacts can achieve instrumental value or disvalue. One possibility is that they are thus invested during use: the gas chambers had a negative instrumental value due to the way in which they were used. After all, it is conceivable that this same technology could alternatively have been used to achieve morally good or neutral ends. In other words, technological artifacts only

---

[6] This possibility is in fact a main reason why some philosophers deny that utility value is valuable at all.

initially have utility value. They acquire instrumental value or disvalue depending on how they are used. However, this thesis is problematic. In most cases technological artifacts are developed with certain uses in mind. It is hard to believe that the engineers who designed the gas chambers were unaware that they were intended for killing Jews. In view of these specific historical circumstances, it makes little sense to assert that the engineers designed the gas chambers and that subsequently, completely *independent* of the design efforts, it was decided to use these gas chambers to kill Jews. This point seems to have more general validity. Houkes and Vermaas argue that in the design process designers do not just design an artifact but also a use plan, that is, a plan in which the artifact functions as a means to achieve certain user ends [Houkes and Vermaas, 2004]. As a matter of fact, the gas chambers that were designed during the Second World War were part of an overall use plan aimed at eradicating the Jews. If it is true, as Vermaas and Houkes argue, that designing always involves designing a use plan, then user goals are not just added to the artifact later on but are intrinsic at least to the intentional history of a technological artifact.

It is not so difficult to think of design processes in which an attempt is made to design an artifact that is particularly fit for achieving a certain positive value or good end. An example is the speed bump, which is intended to force drivers to slow down in built-up areas in order to increase safety. Speeds bumps do not literally force car drivers to drive more slowly but they certainly make it unattractive to drive fast. Speed bumps can thus be seen as an attempt to prescribe a certain moral maxim — do not drive too fast in a built-up area — for people, in order to uphold a positive moral value, in this case human safety. Speed bumps are thus not only designed to have utility value but also instrumental value.

One reason why technological artifacts do not only have utility value but also instrumental value or disvalue is because the goals for which they are intended — and hence also the *intended* instrumental value or disvalue of an artifact — are not extrinsic to the artifact but are part of the intentional history of the artifact in question. Without such an intentional history, an object is not a technical artifact but rather a mere physical object [A similar position is defended by Franssen, this Volume, Part V, especially in Section 1]. Therefore intended instrumental value or disvalue is conceptually inherent to technology. This is not to deny that the *actual* instrumental value of a technological artifact may be different from the one intended in the design process due to the way the artifact is used. It remains true, after all, that artifacts may be used for ends not intended or foreseen by the designer. For this reason, instrumental value is conceptual inherent to technology even if it is not instrinsic to technology.

3.2 Economic value

The economic value of technological artifacts is often a main reason for developing them in a market economy. In economics, economic value is often viewed as the price (in monetary units) that people are willing to pay for tradable goods.

This value may be greater than the actual market price (so creating an economic surplus) but it is not usually less: otherwise, people would not buy the product. So conceived, the economic value of a product depends on the preferences of consumers. People might, of course, use it to create another positive value, but that contribution is not what gives the object its economic value. Therefore, economic value, in terms of the price people are willing to pay for goods, is a kind of utility value.

Economic value is not entirely intrinsic to technological artifacts because it depends on the preferences of consumers which are external to the technological artifacts. Still, economic values may be said to be inherent to technological artifacts insofar as such artifacts are *useful* artifacts that serve a function. This usefulness seems to give technological artifacts an economic value, even if they are not actually traded.

Technologies not only possess economic value as tradable goods but also as means of production. Many technologies facilitate the production of other economic goods and can therefore serve to generate income or profits. Technological innovation often makes it possible to create the same goods, or similar goods with the same function, for a lower price. If it is assumed that the economic value of the goods (i.e. the price that consumers are willing to pay) produced by the new technology is the same as previously, this means that the producer can increase profits and/or the product becomes cheaper and, hence, more affordable for customers. For this reason, it might be argued that technology as a means of production does not only enable the preferences of producers and consumers to be fulfilled but it also has an overall positive value for human welfare. In as far as human welfare is considered a positive value, as it often is, the economic value of technology as a means of production is therefore not just a type of utility value but also something that is instrumentally valuable.

Something similar applies to the economic value of infrastructural technologies. These are technologies such as roads, transportation, and information and communication technologies like the internet. Such technologies are not usually traded nor are they a *direct* means of production for other economic goods. They enable economic activities or make such activities easier or cheaper than they used to be. They may, for example, lower transaction costs or coordination costs. Again, this kind of economic value is a utility value from the point of view that it depends on the preferences of people, but it might be argued that it also contributes overall to other values such as human welfare, and so is instrumentally valuable.

3.3 Moral value

Technologies have utility or instrumental value because they have a consciously designed function. However, technological artifacts do not simply fulfill their function but in passing they also produce all kinds of valuable and harmful side-effects. Chemical plants produce useful substances but may also explode and kill people. Anti-conceptives are not only instrumental in preventing pregnancy, but they also influence sexual morality and have an effect on emancipation. Automatic train ticket dispensing machines may be problematic for the elderly and may discourage them from traveling.

Technologies thus have all kinds of effects, desirable and undesirable, beyond the goals for which they are designed or used. Through these side-effects, technologies create value or disvalue. New technologies can therefore be evaluated not only in terms of the value or disvalue they create through their (intended) use but also in terms of the value or disvalue created through their side-effects (see also the chapters by Hansson and by Grunwald in this Volume, Part V). Typical values that are relevant in this respect include safety, sustainability, human health, welfare, human freedom or autonomy, user-friendliness and privacy. All the referred to values are moral values since they are valuable for moral reasons. This does not mean that they are only valuable for moral reasons or that they are only strived for for moral reasons. In fact, it may be the case that an unsafe technological artifact is less useful than a safe one or is more difficult to sell. Most of the values in question will be enabling or contributory values, in other words values that enable people to live a good life or to contribute to making life good.

Side-effects are often, but not always, unintentional. The designers and users of cars do not intend car accidents to happen nor do they — we may assume — intend to pollute the environment. Nevertheless, we know that the design and use of cars will result in car accidents and in environmental pollution. What is especially relevant here is the fact that different car designs and different modes of use will result in different numbers of car accidents and different degrees of pollution. This point can be generalized: given a desirable technological function or a certain user end, there are usually alternative ways to achieve that function or end. Usually these alternatives not only differ with respect to how effectively and efficiently they meet the formulated end or function, but also with respect to their side-effects, and hence with respect to the values upon which our evaluation of these side-effects are based.

Sometimes side-effects are intentionally incorporated into artifacts. A famous but contested example is that of the low-hanging overpasses at Long Island in the USA [Winner, 1980]. According to Winner, Robert Moses — the architect who designed these overpasses — intentionally designed the road infrastructure in such a way as to prevent black people from reaching the beaches at Long Island. The idea was that black people could not afford private cars and would therefore have to use public transportation, i.e. buses that could not pass under the low-hanging overpasses (for a full discussion, see Radder's chapter in this Volume, Part V, especially Section 4.1).

It might be objected that the above analysis does not indicate that side-effects are either intrinsic or inherent to technology. Side-effects are not entirely intrinsic to technical artifacts because they partly depend on the ways and circumstances of use, even if they also partly depend on the intrinsic properties of technical artifacts. But are side-effects inherent to technological artifacts? If a technological artifact had no side-effects we would not stop calling it a technological artifact. Having side-effects is, in other words, not a defining characteristic of technology and it is not therefore conceptually inherent to technology. I think, however, that it can be argued that side-effects are inherent to technology in a weaker sense. This weaker sense has to do with what I take to be a fact of the world in which we live: technological artifacts are never perfect in the sense that they always also produce effects other than their primary function or aim. Given this fact, side-effects — and the (dis)value they create — are inherent to technology not in a conceptual but in a practical or empirical sense.

3.4 Cultural and aesthetic values

A beautiful building or a carefully designed car is not only instrumentally valuable as a means for living or for transportation but it is also intrinsically valuable as a beautiful object. Technological artifacts may also acquire cultural value. They can be the bearers of meaning: a 2CV has a different cultural meaning than a Mercedes Benz. These two types of cars are not just different means to achieve more or less the same end (of going from A to B); they also embody different kinds of cultural values. As culturally valuable objects, cars are worthwhile in themselves, not just as means of transportation.

Like economic and moral values, cultural and aesthetic values are not intrinsic to technological artifacts. The cultural value attached to an artifact is open to interpretation and may change in the course of time. Even ideals concerning what constitutes beauty may change in different historical periods and vary from culture to culture. If cultural and aesthetic values do indeed depend on interpretation, as is often argued, then such values will not be intrinsic to the valuable object but are rather relational properties of that object.

The fact that the cultural and aesthetic value of technological artifacts depends on their relational properties should not be taken to mean that the attribution of such value to technological artifacts is arbitrary. Just as technological artifacts are often designed with a certain use plan in mind, so they are also often, but certainly not always, designed for a certain cultural or aesthetic value. Car design and styling is a good example of this. Given current user practices, cultural traditions, and social trends, one can to some extent predict what cultural or aesthetic value artifacts will acquire. This value does not only depend on the extrinsic, relational properties of the artifacts but also partly on their material, intrinsic properties. The intrinsic properties partly determine the aesthetic and cultural values of any artifact.

Does this mean that cultural and aesthetic values are inherent to technology? It seems that technological artifacts can always be evaluated from an aesthetic or cultural point of view. However, this does not imply that they have cultural and aesthetic value as technological artifacts. A beautiful letter opener is instrumentally valuable as a technological artifact because it is intentionally designed and can be used to open letters. At the same time it might be a beautiful object and, as such, have intrinsic aesthetic value. Often, the aesthetic value of a technological artifact will not be completely distinguishable from its utility or instrumental value. A beautiful letter opener might not just be a beautiful object, but it may also be beautiful as a *letter opener*, its beauty may be based on the specific way its function is translated into a certain form. It is not just beautiful but it is beautiful because it fulfils its function in a certain way. So, without understanding the function of the letter opener, its beauty may be unexplainable or even be absent. As Roger Scruton says about architecture: " ... our sense of the beauty in architectural forms cannot be divorced from our conception of buildings and of the functions they fulfil" [Scruton, 1979, p. 10]. The same is true of cultural values. A 2CV or Ferrari might amount to a certain expression of a way of living, and as such have cultural value, but this cultural value is often difficult, if not impossible, to understand or even absent altogether if one does not realize that it is a means of transport.

4 VALUES IN THE ENGINEERING DESIGN PROCESS

Design is aimed at the creation of useful things: it aims to create or achieve utility value. For the user, such utility value is achieved by embedding the artifact-to-be-designed in a use plan. This use plan defines the functional requirements that the artifact should fulfill. These functional requirements, the function of the artifact, are achieved through a certain physical structure or make-up [Kroes, 2002]. During the design process, a translation is made from the desired function of the artifact to a physical structure. Put simply, the functional requirements are translated into technical specifications, which are embodied in a certain physical structure.

Engineering design is thus the process by which certain functions are translated into a blueprint for an artifact, system, or service that can fulfill these said functions. Engineering design is usually a systematic process that makes use of technical and scientific knowledge. The design process is an iterative process that can be divided into different stages, like (see, e.g. [Pahl and Beitz, 1984]):

1. Problem analysis and formulation, including the formulation of design requirements and planning the design and development of the product, system, or service.

2. Conceptual design, including the creation of alternative conceptual solutions to the design problem and a possible reformulation of the problem.

3. Choosing one conceptual solution from a set of possible solutions.

4. Embodiment design. The chosen solution is worked out in structural and material terms.

5. Detail design. The design is further detailed, ending up with a design that can function as a blueprint for the production process.

Values can be relevant at all stages of the design process (cf. [van de Poel, 2005]). They are, however, most prominent and explicit during the first phase when the design requirements are formulated, and during the third phase when a specific design solution is chosen and trade-offs between different design requirements often have to be made. In Sections 4.1 and 4.2, these two phases will be analyzed in more detail.

4.1 Values and design requirements

In the first phase of any engineering design process, requirements are formulated on the basis of the intended use of the artifact and the wishes of the client or user. In addition, economic constraints, legal requirements, technical codes and standards, as well as moral considerations will play their part. Below I will distinguish between functional design requirements and additional design requirements. A number of value issues will be discussed with respect to each of these requirements.

4.1.1 Functional design requirements

Functional design requirements are an indication of what the artifact-to-be-designed is supposed to do; they are expressed in functional language. For a pencil, for example, the possible functional requirements are ‘easy to hold,’ ‘does not smear,’ ‘point lasts’ and ‘does not roll’ (cf. [Wasserman, 1993]). During the design process, such functional design requirements are translated into technical specifications. In the case of the pencil, technical specifications may be expressed in terms of the length of the pencil, or the degree of hexagonality of the pencil.

The functional design requirements are an expression of the intended utility value of the artifact-to-be-designed. They may be formulated by the client or the intended user; often, however, the designers will play an important role in translating the rather vague wishes and ideas of clients and prospective users into more concrete design requirements. The designers themselves may also formulate functional design requirements on the basis of a use plan.

It should be noted that the formulation of functional design requirements is value-laden in itself. One issue is: for whom is one designing? As we have seen, the instrumental value of technological artifacts, i.e. their value as means to good ends, depends on how these artifacts are used and can be used. The instrumental value of a Kalashnikov is different from that of a cheap medicine to relieve AIDS in Africa. Of course, not all cases are so clear-cut, but choices concerning users and use plans are clearly value-laden.

4.1.2 Additional design requirements

Not all design requirements are based on the intended use or function of the artifact-to-be-designed. One source of additional design requirements is the possible different stakeholders who are affected by a technological artifact and who have different desires, needs and interests. If one designs, for example, a production machine for paper clips, the client might be the company producing the paper clips. The direct users are the people who operate the production apparatus. In this case, there is also a range of indirect users and stakeholders including the paper clip production facility managers, the safety officer at the production facility, the paper clip users but also labor unions and environmental groups — the apparatus might cause environmental pollution. These stakeholders will place different requirements on the design, requirements that may conflict with each other or with those of clients or users.

A second source of additional requirements is the larger socio-technical system within which the artifact that is designed will be embedded (see Bauer & Herder's chapter in this Volume, Part III). For example, electrical apparatus has to be compatible with the electricity from the grid. The embedding of artifacts in larger socio-technical systems may also give rise to value issues. When designing a car, should one see the existing infrastructure for the distribution of gasoline as a fixed constraint and therefore design a car that runs on gasoline or should one — for example in conjunction with sustainability considerations — opt for a car that uses hydrogen or electricity and lobby for a infrastructure that fits the use of such cars?

A third source of additional design requirements is moral considerations. Engineering codes of ethics, for example, suggest that engineers should hold "paramount the safety, health and welfare of the public" (for more on engineering codes of ethics, see Pritchard's chapter in this Volume, Part V). This suggests that moral values should play an explicit role in the formulation of design requirements. A range of relevant moral values can be mentioned: safety, human health, human well-being, human welfare, privacy, autonomy, justice, sustainability, environmental care, animal health, animal well-being (cf. [Friedman *et al.*, 2006]).

Values are often too broad and vague to be used directly in the design process: they first have to be translated into more tangible design requirements. Often, different translations are possible. In the design of a chemical plant, one can look at the safety of employees and of people living close to the plant. Ethically, it would not be acceptable to limit safety to just the employees. Obviously, people in the direct vicinity of such a plant will experience the consequences of the design choices made without having the opportunity to agree or disagree or to benefit directly from the plant.

4.2 Trade-offs in design and value conflicts

While some design requirements are formulated as requirements that can be met or not — e.g. this electrical appliance should be compatible with 220V — others are formulated in terms of goals or values that can never be fully met. An example of the latter is safety. There is no such thing as an absolutely safe car: cars can only be safe to greater or lesser degrees.

If design requirements are formulated as strivings, they are more likely to conflict with each other. Cars that are made lighter in order to be more sustainable (less fuel consumption) are, for example, less safe [van de Poel and van Gorp, 2006]. The refrigerator coolant that replaced CFC 12 in most European household refrigerators — after the ban on CFCs — is more environmentally sustainable but flammable and therefore somewhat less safe than CFC 12 and other alternatives [van de Poel, 2001]. Most designs involve trade-offs between different design requirements. If the design requirements are motivated by different values these conflicts amount to value conflicts. Below I sketch two examples of value conflicts in engineering design.

4.2.1 Safety belts

A first example is the automatic seatbelt. A car with automatic seatbelts will not start if the automatic seatbelts are not put on. This forces the user to wear the automatic seatbelt. One could say that the value of driver safety is built into the technology of automatic seatbelts. This comes at a cost, however: the user has less freedom. Interestingly, there are various seatbelt designs which exist that would imply that hre are different trade-offs in terms of safety and user freedom. The traditional seatbelt, for example, does not enforce its use, but there are various systems that give a warning signal if the seatbelt is not being worn. This does not enforce seatbelt use, although it does encourage the driver to wear his seatbelt.

4.2.2 Refrigerants for household refrigerators

As a consequence of the ban on CFCs in the 1990s, an alternative had to be found to CFC 12 as a refrigerant in household refrigerators. Apart from utility value, three moral values played an explicit role in the formulation of design requirements for alternative coolants: safety, health and environmental sustainability. In the design process, safety was mainly understood as non-flammability, and health as non-toxicity. Environmental sustainability was equated with low ODP (Ozone Depletion Potential) and low GWP (Global Warming Potential). Both ODP and GDP mainly depend on the atmospheric lifetime of refrigerants. In the design process, a conflict arose between those three values. This value conflict can be illustrated with the help of Figure 1, which derives from a publication in the *ASHRAE Journal* of December 1987 by two engineers, McLinden and Didion, both employed by the National Bureau of Standards in the USA [McLinden and Didion, 1987].

For thermodynamic reasons, the most attractive coolants are hydrocarbons or CFC based on such hydrocarbons. Figure 1 is a graphic representation of CFCs based on a particular hydrocarbon. At the top, there is methane or ethane, or another hydrocarbon. If one moves to the bottom, the hydrogen atoms are replaced either by chlorine atoms (if one goes to the left) or fluorine atoms (if one goes to the right). In this way, all the CFCs based on a particular hydrocarbon are represented. The figure shows how the properties of flammability (safety), toxicity

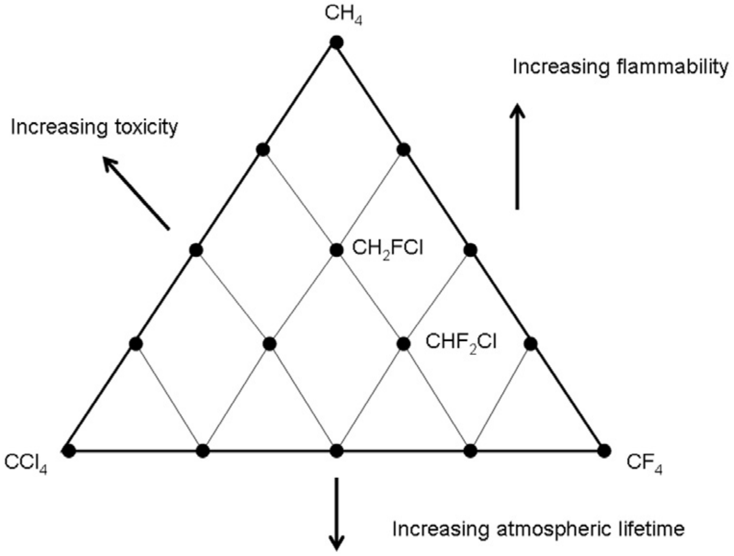

Figure 1. Properties of refrigerants

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

5 DEALING WITH VALUE CONFLICTS IN ENGINEERING DESIGN: OPTIMAL DESIGN

When dealing with trade-offs, engineers are inclined to look for the best or optimal design solution. Below, I will investigate various approaches to optimal design: efficiency and effectiveness (Section 5.1), cost-benefit analysis (Section 5.2) and multiple criteria design analysis (Section 5.3).

5.1 Efficiency and effectiveness

A first-order approach to optimal design is to consider design to be optimal if it results in an artifact that optimally fulfills the desired function. But how do we know — or measure — whether a design optimally fulfills its function? Two measures come to the fore: effectiveness and efficiency. Effectiveness can be defined as the degree to which an artifact fulfills its function. Efficiency could be defined as the ratio between the degree to which an artifact fulfills its function and the effort required to achieve that effect. As is pointed out in by Alexander (this Volume, Part V), efficiency in the modern sense is usually construed as an output/input ratio. The energetic efficiency of a coal plant may thus be defined as the ratio between the energy contained in the power produced and the thermal energy contained in the unburned coal.

Historically, there seems to be a close connection between the rise of the modern notion of efficiency and optimal design. Frederick Taylor, for example, believed that production should be based on "the one best way" which, according to him, was simply the most efficient way [Taylor, 1911].

Two things need to be noted with respect to effectiveness and efficiency. Firstly, effectiveness and efficiency are different values that may well conflict. The design that most effectively fulfills its intended function may not necessarily be the most efficient one. A very effective vacuum cleaner that removes more dust than a less effective one may nevertheless be less energy-efficient, that is to say, it may use more energy per unit of dust removed than the less effective vacuum cleaner. So, we may be faced with a conflict between effectiveness and efficiency. A well-defined notion of optimal design requires a solution to this potential conflict. Secondly, effectiveness and efficiency are often very difficult to measure. Although this is partly a practical problem, this difficulty is often based on the more fundamental problem that often neither the function of an artifact (i.e. its output) nor the input can be uniformly formulated. This is witnessed, for example, by the fact that the desirable function of an artifact is often expressed in terms of a range of functional requirements, which may conflict. The following quote from Petroski about the design of paper clips illustrates this point:

> Among the imperfect things about the Gem [the classic paper clip, IvdP] that many a recent inventor has discovered or rediscovered when reflecting upon how the "perfected" paper clip is used to clip papers together are the following:
>
> *1. It goes only one way.* Half the time, the user has to turn the clip around before applying.
>
> *2. It does not just slip on.* The user first has to spread the loops apart.
>
> *3. It does not always stay on.* The clip gets snagged on papers or other objects and gets pulled off.
>
> *4. It tears the papers.* The sharp ends of the clip dig into the papers when it is removed.
>
> *5. It does not hold many papers well.* The clip either twists badly out of shape or flies off the pile.
>
> *6. It bulks up stacks of papers.* A lot of file space can be taken up by paper clips.

> When a design removes one of the annoyances, it more likely than not fails to address some others or adds a new one. . . . All design involves conflicting objectives and hence compromise and the best designs will always be those that come up with the best compromise. Finding a way to bend a piece of wire into a form that satisfies each and every objective of a paper clip is no easy task, but that does not mean that people do not try. [Petroski, 1996, p. 29-30]

This quote illustrates two points. One is that the ideal of optimal design or what Petroski calls “perfected” design is an important source of inspiration for designers. As long as the perfect paper clip does not exist, people will try to design it. The other is that in practice this ideal will probably never be achieved: the best is always the best compromise. The crucial question then is how to determine what the best compromise is. This requires trade-offs between the different requirements and it is unclear how we can make these trade-offs in a justified way (see also Kroes *et al.*, this Volume, Part III).

The actual situation is, however, even worse. Up until now, we have conceived of optimal design as design that optimally fulfills its intended function, or — put differently — as that which maximizes the (expected) utility value of the design. However, as argued in Section 3, the value of technological artifacts is not restricted to their utility value. The question that then arises is: what would it mean to try to maximize the overall value of technological artifacts during design, what would optimal design in such a broader sense amount to? Engineers have, in fact, dealt with this problem and have developed a number of approaches to the issue. The following two such approaches will be briefly discussed below: cost-benefit-analysis and multiple criteria design analysis.

5.2 Cost-benefit analysis

Cost-benefit analysis is a general method that is often used in engineering. What is typical of cost-benefit analysis is that all considerations that are relevant for the choice between different options are eventually expressed in one common unit, usually a monetary unit, like dollars or euros.

Cost-benefit analysis may be an appropriate tool if one wants to optimize the expected economic value of a design. Still, even in such cases, some additional value-laden assumptions and choices need to be made. One issue is how to discount future benefits against current costs (or vice versa). The choice of discount rate may have a major impact on the outcome of the analysis. One might also employ different choice criteria once the cost-benefit analysis has been carried out. Sometimes all of the options in which the benefits are greater 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.

Cost-benefit analysis is more controversial if non-economic values are also relevant. Still, the use of monetary units does not mean that only economic values can be taken into account in cost-benefit analysis. In fact, approaches like willingness to pay (WTP) have been developed to express values like safety or sustainability in monetary units (for some more details on WTP, see Hansson's chapter in this Volume, Part V, Section 4.5 and Grunwald's chapter in this Volume, Part V, Section 3.2.4). These approaches are often questioned but it would be premature to conclude that cost-benefit analysis necessarily neglects non-monetary or non-economic values. Moreover, when employing cost-benefit analysis, different ethical criteria might be used to choose between the options [Kneese *et al.*, 1983; Shrader-Frechette, 1985]. One might, for example, choose an option with which nobody is worse off. By selecting a specific choice criterion, ethical considerations beyond considering which options bring the largest net benefits might be taken into account.

In terms of values, cost-benefit analysis might be understood to be the maximization of one overarching or super value. Such a value could be an economic value like company profits, or the value of the product to users but it could also be a moral value like human happiness. If the latter is chosen, cost-benefit analysis is related to the ethical theory of utilitarianism. With Bentham's classical variant of utilitarianism, for example, the assumption is that all relevant moral values can eventually be expressed in terms of the moral value of human happiness. One might question this assumption, however. One issue is that it is often difficult to indicate to what extent values like safety, health, sustainability, and aesthetics contribute to the value of human happiness, and to furthermore express this in monetary terms. A second, more fundamental issue, is that such an approach treats all these values as extrinsic values, whose worth should ultimately be measured on the basis of their contribution to the intrinsic value of human welfare. One might wonder whether values like human health, sustainability and aesthetics do indeed have only extrinsic value or are worthwhile in themselves. This potential objection to cost-benefit analysis amounts to an objection to the method being based on value monism.

However, it might be argued that the above account of cost-benefit analysis is too substantive, from the point of view that it presupposes that the method is all about maximizing a specific value. Some proponents of cost-benefit analysis would probably maintain that it is merely a technical way of comparing alternatives in the light of heterogeneous considerations or values. The use of a common measure, they may admit, presupposes a common value but this value is merely formal, in other words, it is merely a means of comparison, rather than a substantial value. The claim, made above, that cost-benefit analysis presupposes value monism might thus be misconceived. I think two points need to be made in reply to such criticism. Firstly, even if cost-benefit analysis were merely a technical approach, interpretations of what this approach amounts to — even by most proponents of the approach – would often suppose a kind of value monism. Secondly, as a merely technical approach cost-benefit analysis might not indeed suppose value monism, but it does suppose value commensurability because it presumes that all values can be measured on a common scale. This may be a problematic assumption (see Section 2.4 and Hansson's chapter in this Volume, Part V, especially Section 4.5).

5.3 Multiple criteria design analysis

In multiple criteria analysis, different options are compared with each other in the light of several criteria. I focus here on one specific approach that is often used in engineering design: the method of weighted objectives. With this method, the relative importance of the criteria is first determined, because usually not all criteria are equally important. Next, each option is weighed for all the criteria and a numeric value is awarded, for example on a scale from 1 to 5. Finally, the value for each option is calculated according to the following formula: $w_j = \sum f_i * v_{ij}$ over I, where w_j is the value of the j^{th} option, f_i is the relative weight of the i^{th} criterion, and v_{ij} is the score of the j^{th} option on the i^{th} criterion. The option with the highest value is then selected.

The method can be illustrated using the case of coolants discussed in Section 4.2. I mentioned there three relevant values in the designing of coolants: safety, health and environmental sustainability. As we have seen, these values conflict in terms of the choice to be made between various coolants. How would the method of weighted objectives proceed in a case like this? The most simple and straightforward approach is to conceive of each of these values as a decision criterion. Table 1 gives a hypothetical interpretation of the choice between three of the alternatives that were considered: the traditionally used CFC 12; HFC 134a, the main alternative proposed by the chemical industry; and isobutene, an alternative proposed by environmental groups because it contributes less to greenhouse warming than HFC 134a, but is flammable.

	Safety (flammability)	**Health (toxicity)**	**Environmental sustainability (atmospheric lifetime)**	**Total score**
Weight of criterion	**2**	**1**	**2**	
CFC 12	5	5	1	17
HFC 134a	4	4	3	18
Isobutane (HC 600a)	1	4	5	16

Table 1. Hypothetical application of the method of weighted objectives to the coolants case

Table 1 suggests that we should choose HFC 134a. But how sound is this advice? The first thing to note is the use of the scoring scale 1 — 5 for each decision criterion. The way this scale is used to calculate the overall worth of each alternative means that this scale is interpreted as a ratio scale for each of the design criteria.[7] This means that it should make sense to say that HFC 134a scores three times as well on the criterion of environmental sustainability as CFC 12. If seems beyond debate to claim that HFC134a scores better on environmental sustainability than CFC 12 — so we can order the alternatives on an ordinal scale — but can we say that it does three times as well according to this criterion? To do this we would require an operationalisation of environmental sustainability that we can measure. Such operationalisation still seems to be lacking. Even if we can compare the atmospheric lifetime of two substances, it is not obvious that changing the atmospheric lifetime of a substance by a factor of 2 would correspond to a similar change in the environmental sustainability of that substance.

The method of weighted objectives also suggests that the weights of the criteria can be measured on a ratio scale. This can, for example, be attained by asking the marginal rate of substitution question: "By how much should f_i be increased to compensate for a loss of one unit in f_1?," in which f_i refers to the weight of the i^{th} criterion and f_1 to the weight of the criterion that is selected as a metric case [Otto, 1995, p. 97]. Obviously, it is only possible to answer this question if the design criteria, and the underlying values, are commensurable.

The method of weighted objectives is not, of course, the only multiple criteria method that can be applied in engineering. Any of these methods must, however, deal one way or another with value trade-offs and that seems to presuppose some form of value commensurability. Moreover, Franssen [2005] has shown that with all multiple criteria methods it is very hard to make justifiable trade-offs; as Kroes, Franssen and Bucciarelli conclude in their contribution to this volume: "… there is no general rational procedure for making trade-offs in engineering design" (see Kroes *et al.*, this Volume, Part III).

5.4 The ideal of optimal design

As we have seen, optimizing approaches to value conflicts in engineering are likely to come up against formal and substantive problems. Philosophically, these problems are mainly attributable to the fact that it is often impossible to identify one overarching or super value (value pluralism) and to value incommensurability. We should not, however, conclude from this that the ideal of optimal design has no importance whatsoever.

Firstly, the ideal of optimal design often motivates and guides the design process. Since, at the start of the design process, engineers do not yet know what is technically feasible, the ideal of optimal design — and especially the more specific

---

[7] If it is interpreted as an ordinal or an interval scale, it can easily be shown that the method does nor result in one but in a number of potentially conflicting overall orderings (see also [Franssen, 2005]).

design requirements that are relevant to optimal design — is important in looking for new technical possibilities.

Secondly, the ideal of optimal design often spurs on investigations into technical parameters that are relevant to design improvement. This helps to provide a better understanding of the design problem while improving the resulting design.

Thirdly, design problems can often be subdivided into smaller problems. Some of the smaller problems might be so well-defined that optimal solutions do exist.

Fourthly, the costs of ignoring the philosophical problems related to value pluralism and value incommensurability and, for example, to carry out a cost-benefit analysis may be less than the costs of selecting another choice mechanism or just picking one design (cf. [Sunstein, 2005, p. 371]). This is especially the case in choice situations where one type of values, for example economic values, is obviously more important than others. Even in cases where there are a number of conflicting values, the crucial question is whether there are other, non-optimizing approaches that fare better than optimizing approaches. In the next section, I will suggest a number of such alternative, non-optimizing, approaches and discuss some of their advantages and weaknesses.

6 NON-OPTIMIZING APPROACHES TO VALUE TRADE-OFFS IN ENGINEERING DESIGN

Non-optimizing approaches are not alien to engineering. In fact, a number of authors have argued that it is not possible to optimize in engineering design (e.g. [Simon, 1973; Cross, 1989; Schön, 1992; Simon, 1996]). A major argument for the impossibility of optimizing in engineering given by Herbert Simon and Nigel Cross has to do with the ill-defined nature of engineering design problems.

Simon lists a number of characteristics that problems should have if they are to be well-defined [Simon, 1973]. Three characteristics that are especially relevant in relation to engineering design are these:

1. A clear criterion needs to be available to judge possible solutions and this criterion can be applied uniformly;
2. A problem space can be defined in which the initial state, the desired state and all possible interim states — that can be considered or achieved during problem solving — can be represented;
3. Possible actions or solutions can be represented as transitions between different problem space states. Insofar as actions affect the real world — and are thus outside a formal language or play — the representation should match the natural laws of the external world.

Most design problems do not meet these criteria. A clear and uniformly applicable choice criterion is not usually available, as is clear from the discussion presented in Section 5. Moreover, the problem space is not usually well-defined.

In engineering design we frequently do not have an overview of the complete set op possible solutions, let alone a representation of this set in a well-defined problem space. Design problems are thus usually ill-defined.

The ill-defined nature of design problems makes optimizing difficult, if not impossible. There are, however, alternative, non-optimizing strategies in engineering design. One non-optimizing strategy that was first described by Simon and which engineers are reported to follow is satisficing involving the selection of an alternative that is not optimal but 'good enough.' Satisficing is also a possible strategy for dealing with conflicting values, as already noted by Simon [1955]. Other non-optimizing strategies are possible with respect to conflicting values in engineering pertaining to reasoning about values, technological innovation and the choice of a diversity of products. Each of these strategies will be discussed below.

6.1 Satisficing

In contrast to an optimizer, a satisficer 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, 1976]. 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. Setting threshold values does not only occur in the design process but also in legislation and the formulation of technical codes and standards.

An example of satisficing is to be found in the earlier-discussed case of the design of new refrigerants. On the basis of Figure 1, the engineers McLinden and Dion, drew more specific figure with respect to the properties of CFCs (Figure 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 (toxicity), safety (flammability) and the environment. These thresholds were partly based on technical codes and standards. For example, the threshold value for safety — non-flammability — was based on the then existing *ASHRAE Code for Mechanical Refrigeration* (ASHRAE Standard 15-1978), which prohibited the use of flammable coolants in equipment intended for household applications.

Satisficing can also be combined with optimizing. For example, a designer who has to trade off safety and cost considerations when designing a chemical installation may well choose to make a design that meets the legal requirements with respect to safety and is as cheap as possible. This can be interpreted as satisficing behavior with respect to the value of safety, while optimizing with respect to cost within the safety constraints.

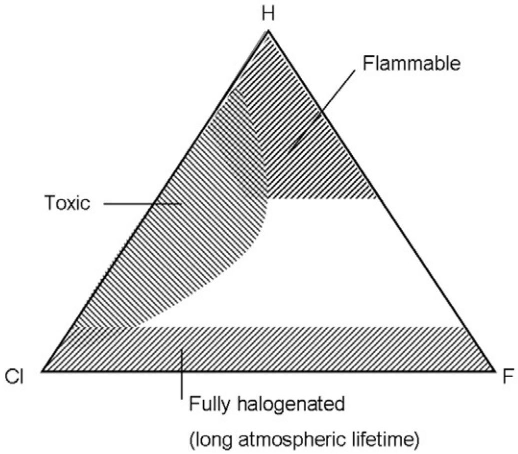

Figure 2. Properties of refrigerants. [Figure from McLinden and Didion, 1987]

Philosophically, the important question is whether, and if so when, satisficing is a morally and rationally permissible strategy. If someone satisfies he does not aim for the best, but for an option that is good enough from a certain point of view. Some ethicists have argued that satisficing with respect to moral values might be allowable: in many situations we are not required to do what is morally best, but we should at least do what is morally good enough (see, e.g. [Stocker, 1990; Dancy, 1993]). Risking one's life to save another person from a burning house might be morally praiseworthy, but that does not mean that it is morally required. So, not everything that is morally praiseworthy is also morally required because that might just demand too much of someone. This phenomenon is known as moral supererogation.

One argument for why satisficing is not only allowed but may even be advisable in the case of moral values might go as follows. Moral values sometimes resist trade-offs as we have seen in Section 2.4. One possible explanation for this is that they may often be understood as moral obligations [Baron and Spranca, 1997], as the obligation to meet a certain value to a certain minimal extent. If interpreted thus moral obligations 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. If this picture is right, it provides an additional argument for satisficing with respect to moral values: for each of the relevant moral values the threshold should be set in such a way that the corresponding moral obligation is at least met. If this is done, unacceptable trade-offs between moral values or between moral values and other values can be avoided.

Do the above arguments also apply to non-moral values? An important argument for the existence of moral supererogation is that we have other values (and reasons) besides moral values (and reasons), and that these may go against what is morally most praiseworthy. It is, for example, obvious that I have reasons *not* to risk my life to save somebody else from a burning house, even if I have good reason to believe that it would be best from a moral point of view to risk my life. It is the existence of such other reasons that can justify moral satisficing. The justification here is based on the — presumed — existence of a broader perspective that includes the moral perspective, for example the perspective of the entire life of a person.

Can satisficing also be justified at the broadest perspective level, like the perspective of an entire life? Some philosophers, preeminently Michael Slote, believe that it can [Slote, 1984; 1989]. Slote argues that it can be rationally justifiable to forego the best choice even if we know what the best choice is and even if it is readily available; he calls this 'rational supererogation.' Slote's argument is strongly contested, however (e.g. [Pettit, 1984; Schmidtz, 1995; Byron, 1998]). What makes it especially problematic is his claim that it is *rationally* allowed to choose some lesser option over an available better option even though we have no overriding reason for doing so. It seems that we either have an overriding reason to choose the lesser option, which makes it not the lesser option all things considered, or we are simply not rationally allowed to choose the lesser option.

The argument against Slote's position suggests that satisficing cannot be rational in the broadest perspective. It can only be rational with respect to a partial perspective; satisficing on such a sublevel can be rational because, seen from the wider perspective, it is the best way to achieve one's overall values or aims. Some philosophers have therefore suggested that at the highest level we are always optimizing if we are rational, at least implicitly and tacitly [Byron, 1998]. The idea of implicit and tacit optimization, however, seems problematic. It suggests that we optimize even if we are not aware of it. However, this makes it impossible to empirically distinguish optimizing from non-optimizing, so that the thesis that we optimize cannot be empirically falsified. Maybe, however, the thesis is not intended to refer to an empirical fact but rather to a conceptual truth: the conceptual nature of rationality leaves no other option than to optimize at the highest level if we want to be rational. This however, is a very bold assertion. To make it plausible we should at least argue how people *could*, if they would wish, optimize at the highest level, and that they could do so always if they would wish so. One reason for doubting whether such an argument can be given has to do with the existence of plural and conflicting values and value incommensurability. This places doubt on the possibility of optimizing as we have seen: the notion of a best option may simply not be well-defined.

Another relevant issue with respect to the acceptability of satisficing is that of whether we are considering a static or dynamic context [Simon, 1956; Schmidtz, 1995]. In static contexts, all the options are known, the consequences of the options are known with a certain probability and the options are readily available.

Slote sets out to argue that satisficing is rationally allowed in a static context. Such an argument is very hard — if not impossible — to substantiate, but what of the rational acceptability of satisficing in a dynamic context? In a dynamic context we are either not fully aware of all the options, or it requires effort to investigate the consequences of options or to make them available. In such a context, we are confronted with the following question: how much effort should be put into getting to know the solution space better? The effort may pay off because there is a chance that we will discover a better option than the options we already know. However, there are limits: in situations in which the solution space is not closed, we can go on endlessly searching for a better solution but at some point the results no longer justify the effort. In such a dynamic situation, satisficing may be a useful and rationally defensible stopping rule: look for a better option until you have found an option that meets all threshold values.[8]

What does this tell us about the acceptability of satisficing in engineering design? Firstly, it suggests that satisficing with respect to moral values – or more specifically morally motivated design requirements — can be permissible due to the phenomenon of moral supererogation. Secondly, it suggests that satisficing with respect to other values and design requirements can be rationally justified from a broader perspective. In the case of the design of a part or component, this broader perspective can be the design process of the entire artifact. In the case of a design process for an artifact the broader perspective can be the sociotechnical system in which the artifact is embedded. The broader perspective can also be the company that wants to make a profit with a certain design or it can be society at large that aims to sustain certain values through technology. Thirdly, satisficing can be rational in a dynamic context where the solution space is not closed. As we have seen this is often typically the case in engineering design. Satisficing can therefore provide a rationally defensible stopping rule for the search process that engineering design is.

---

[8] It should be noted that it is also possible to formulate an optimizing stopping rule. A good candidate for the optimizer's stopping rule is: stop looking for new options as soon as the expected value of finding a better option is lower than the expected costs of any further searching process. This stopping rule is different from that of the satisficer in that it requires calculations to decide whether to continue the search or not; such calculating takes time and might prove counterproductive. So the optimizer needs a stopping rule concerning the time spent on calculating whether it is worth continuing the search. If the optimizer chooses a satisficing stopping rule for the making of the calculation, he is not optimizing any longer; but, if he chooses an optimizing stopping rule he has to make even more calculations because the time to be spent on the other calculations needs to be accounted for (even though these calculations presumably take less time than if he had not applied the stopping rule). The point is this: at least in some circumstances, the time it takes the optimizer to make all these decisions is simply not worth the effort. The optimizer is usually, if not always, better off if, at some level, he chooses a satisficing stopping rule. He does not know, however, when it is the right time to choose the satisficing stopping rule and in that sense he is not optimizing even if he chooses the satisficing rule because he wants to optimize. Wanting to optimize is, after all, not the same as optimizing.

6.2 Reasoning about values

The approaches to dealing with conflicting values that have already been discussed — efficiency, cost-benefit analysis, multiple criteria analysis and satisficing — are all calculative approaches. They strive to operationalize and measure value in one way or another. Of these approaches, satisficing does not aim at calculating the overall worth of an option, but it does presuppose that the worth of an option can be measured for each of the individual design requirements. I will now look at approaches that do not share this calculative approach, but which emphasize judgment and reasoning about values.

In the philosophical literature on value incommensurability, certain authors maintain that the presence of value incommensurability does not impair our ability to compare options because we can exercise judgment (e.g. [Stocker, 1990]). Often precisely what such judgment implies and how it could lead to a justified choice in situations of value conflict remains unclear. Nevertheless, a number of things can be said about what such judgment could imply.

The first thing to do when one wants to exercise judgment in cases of value conflict is to gain a better understanding of the values at stake. 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 construed 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 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.

A second judgment step would be to argue for specific conceptualizations of the relevant values. 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.

A third step in judgment is to look for the common ground behind the various values that might help to solve the value conflict. Taebi and Kloosterman [2008], for example, have argued that various trade-offs in nuclear energy can be reduced to a trade-off between the present and the future, and can thus be best understood in terms of the notion of intergenerational justice. One might argue that looking for common ground between different values presupposes a form of value monism. I suppose this is true. It should be stressed, however, that this kind of value monism is different from that presupposed by cost-benefit analysis or utilitarianism. In the latter cases, the presupposition is that all values can be expressed in terms of an overarching value, like human welfare. Here the value monism is more a monism of reasons. It is the type of value monism that Kant seems committed to. According to Korsgaard, Kant recognizes only one unconditionally good value as the source of all other values and that is: 'good will' [Korsgaard, 1986]. It is likely that Kant would maintain that good will can solve all value conflicts, at least in principle. This seems too optimistic to me, but that does not reduce the need to look for common ground between values. Even if such common ground cannot always be found, it may be available in specific cases.

6.3 Innovation: value sensitive design

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 situation. In a sense, solving value conflicts by means of new technologies is what lies at the heart of engineering design and technological innovation. Engineering 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 and R&D may be able to change these possibilities.

An interesting example is the design of a storm surge barrier in the Eastern Scheldt estuary in the Netherlands. After a huge storm-flood hit the Netherlands in 1953, killing more than 1,800 people, the government decided to dam up the Eastern Scheldt as part of what came to be known as the Delta plan. The main value taken into account in the Delta plan was safety. The closure of the Eastern Scheldt was scheduled to start in the early seventies, as the final part of the Delta plan. However, by that time it had led to protests in conjunction with the ecological value of the Eastern Scheldt estuary, which would in that way be destroyed. Many felt that the ecology of the estuary should be considered. Eventually, a group of engineering students devised a creative solution that would meet both safety and ecological concerns. The idea was to construct a storm surge barrier that would be closed only in cases of storm floods. Eventually this solution was accepted as a creative, though more expensive, option that took into account both the values of safety and ecology.

One approach that takes into account the possibility to solve, or at least ease, value conflicts through engineering design is Value Sensitive Design. Value Sensitive Design is an approach that aims at integrating values of ethical importance in a systematic way into engineering design [Friedman, 1996; Friedman and Kahn, 2003; Friedman *et al.*, 2006]. The approach aims at integrating three kinds of investigations: conceptual, empirical and technical.

Conceptual investigations aim, for instance, at clarifying the values at stake, and at making trade-offs between the various values. Conceptual investigations in Value Sensitive Design are similar to the kind of investigations that I described in Section 6.2. What Value Sensitive Design adds to this are empirical and technical investigations.

Empirical investigations "involve social scientific research on the understanding, contexts, and experiences of the people affected by technological designs" [Friedman and Kahn, 2003, p. 1187]. It is not hard to see why this is relevant: people's experiences, contexts and understanding are certainly important when it comes to appreciating precisely what values are at stake and how these values are affected by different designs. What remains somewhat unclear is just how the proponents of Value Sensitive Design see the relationship between conceptual and empirical investigations. For example, is it important how people perceive a value or how it should be understood on conceptual grounds? I would argue that people's understanding of values is not irrelevant but that it should not be taken at face-value either, people might err after all. One could, for example, require people to justify their understanding of the values at stake in a broad reflective equilibrium [Daniels, 1979; 1996; Rawls, 1971/1999]]. This could also provide a model for integrating conceptual and empirical investigations; a model that still seems to be lacking in the literature on Value Sensitive Design.

Technical investigations "involve analyzing current technical mechanisms and designs to assess how well they support particular values, and, conversely, identifying values, and then identifying and/or developing technical mechanisms and designs that can support those values" [Friedman and Kahn, 2003, p. 1187]. The second part is especially interesting and relevant because it provides the opportunity to develop new technical options that more adequately meet the values of ethical importance than do current options. As the example of the Eastern Scheldt barrier shows, technical investigations may also ease value conflicts. Usually, however, technical innovation will not entirely solve value conflicts, so that choices between conflicting values still have to be made. In this respect, Value Sensitive Design only presents a partial solution to value conflicts in engineering design.

6.4 Diversity, genre and value holism

All approaches to value conflict discussed so far presuppose that only one option is to be chosen. This presupposition is indeed true of most specific product design processes. If we zoom out from this perspective, a somewhat different picture emerges though: engineering provides society with different products that have roughly the same function. A number of possible justifications can be given for this diversity.

One is that different people have different needs and preferences. A can opener, which is useful for the average citizen, may not be the most apt device for the elderly who usually have less power in their hands. Due to divergences in, for example, physical make-up, different people have different needs, a fact of life which may justify the design of different products with roughly the same function for different groups of users. People may also have different preferences. Some people will prefer a fast but expensive car to a less expensive but slower car, whereas others will have opposite preferences. Even if not all preferences are justified or to be fulfilled, as such there is nothing wrong with differences in preferences and a diversity of products with roughly the same function may be instrumental in fulfilling such diverse preferences. Given the differences between people in terms of their needs and preferences, the existence of a diversity of technological products with roughly the same function may be one way of increasing the utility or economic value of technological products. The total utility or economic value for society is probably larger if a range of products is provided rather than just one.

A second possible justification is the existence of cultural differences and different cultural and aesthetic traditions. As we have seen, technological artifacts have meaning and may express certain cultural or aesthetic ideals. This is clearly visible in architecture, where various traditions matched to different evaluation standards exist. In such cases evaluation tends to be genre-specific: we identify a work as an instance of a genre and judge it by the standards of the genre. As Joseph Raz expressed:

> We can admire a building, and judge it to be an excellent building for its flights of fancy, and for its inventiveness. We can admire another for its spare minimalism and rigorous adherence to a simple classical language. We judge both to be excellent. Do we contradict ourselves? Not necessarily, for each displays the virtues of a different architectural genre — let us say, romantic and classical. [Raz, 2003, p. 45]

Of course, the relationship between a work and its genre is not always straightforward, since works can also ironically or ambiguously relate to a genre or to more than one genre at the same time. Nevertheless, also in such cases, we tend to judge the work in terms of how it is allied to existing genres and to the standards of excellence inherent in those genres [Raz, 2003, p. 41-42].

Typically, different genres — for example in architecture — do not often differ merely in terms of the general relevant values, but also in terms of the preferred 'mix' or 'ideal' combination of values [Raz, 2003, p. 39]. In such cases, an optimizing approach that seeks to optimize one overarching value or each of the values in isolation may actually destroy this 'ideal' mix and create less value instead of more. A satisficing approach might not be appropriate either because what is valued is not a certain degree of each value but a specific combination of values. Such cases of value holism, where we cannot reasonably appreciate the values in isolation of each other, usually require educated judgment, which is genre-specific. Therefore, with respect to values that are genre-specific — especially some cultural and aesthetic values — it may be worthwhile having a number of technological artifacts with roughly the same function that each excel in their own genre.

One might wonder whether such diversity is also worthwhile in the case of moral values. The idea of genre-based evaluation seems much harder to defend with respect to moral values than with respect to cultural or aesthetic values. It might be defensible with respect to perfectionist moral values, in other words those that go beyond what is morally required. The value of a good life, conceived as a moral value, may also be genre-specific. However, genre-specific evaluation seems hard to defend with respect to minimum moral standards. Still, it might be argued that minimum moral standards are not universal but situation-specific (cf. e.g. [Dancy, 1993]). For example, the minimum moral standard for environmental sustainability or animal welfare may be higher in a society of abundance than in a society of scarcity, especially if meeting high standards in, for example, animal welfare would involve further deterioration in living conditions for humans. This suggests that the thresholds for moral values — if one employs a satisficing approach — cannot been established completely independently of each other. This entails some degree of value holism; we cannot appreciate the values in complete isolation of each other. However, it does not extend to the type of genre-specific evaluations Raz is thinking of.

7 CONCLUSION

I began this contribution with the observation that it is the creation of value that lies at the heart of the engineering design process. We may now conclude that value conflict is in fact at the heart of the design process. In many cases, value conflict is the engine that fuels innovation and design, as is underlined by an approach such as the Value Sensitive Design approach. I have discussed various approaches to value conflict in engineering design. None of them is obviously superior to any of the others. The discussion suggests though that the approach that perhaps turns out to be most fruitful will depend, partly at least, on the kinds of values that are at stake. Cultural and aesthetic values are often genre-specific and will often consist of ideal combinations. Optimizing or satisficing approaches are probably of little help in such cases. Rather, one might adopt the approach that I have described under the heading of diversity. Most moral values, on the other hand, are not genre-specific. Here we might initially try to solve, or at least ease, a moral value conflict by employing Value Sensitive Design or by reasoning. A possible resultant value conflict might be dealt with by satisficing which can, as we have seen, amount to a justified approach to moral values. Optimizing might be a less desirable approach, especially if we are dealing with heterogeneous moral values that resist trade-offs. With economic values and other utility values, optimizing approaches might be fruitfully applied, even if such approaches still come up against a number of methodological problems. As we have seen, under certain conditions satisficing approaches and diversity might also be useful approaches for utility and economic value.

BIBLIOGRAPHY

[Ball *et al.*, 1994] L. J. Ball, J. S. B. T. Evans and I. Dennis. Cognitive processes in engineering design: A longitudinal study. *Ergonomics* 37, 1753-86, 1994.

[Baron and Spranca, 1997] J. Baron and M. Spranca. Protected values. *Organizational Behavior and Human Decision Processes*, 70, 1-16, 1997.

[Broome, 1997] J. Broome. Is incommensurability vagueness? In *Incommensurability, Incomparability and Practical Reason*, R. Chang ed. pp. 67-89, Harvard University Press, 1997.

[Byron, 1998] M. Byron. Satisficing and optimality. *Ethics* 109, 67-93, 1998.

[Chang, 1997] R. Chang, ed. *Incommensurability, Incomparability, and Practical Reasoning.* Harvard University Press, 1997.

[Cross, 1989] N. Cross. *Engineering Design Methods.* John Wiley & Sons, 1989.

[Dancy, 1993] J. Dancy. *Moral Reasons.* Blackwell Publishers, 1993.

[Daniels, 1979] N. Daniels. Wide reflective equilibrium and theory acceptance in ethics. *Journal of Philosophy* 76, 256-82, 1979.

[Daniels, 1996] N. Daniels. *Justice and Justification. Reflective Equilibrium in Theory and Practice.* Cambridge University Press, 1996.

[Florman, 1987] S. C. Florman. *The Civilized Engineer.* St. Martin's Press, 1987.

[Frankena, 1973] W. K. Frankena. *Ethics.* Prentice Hall, 1973.

[Franssen, 2005] M. Franssen. Arrow's theorem, multi-criteria decision problems and multi-attribute preferences in engineering design. *Research in Engineering Design* 16, 42-56, 2005.

[Friedman, 1996] B. Friedman. Value-sensitive design. *Interactions* 17-23, 1996.

[Friedman and Kahn, 2003] B. Friedman and P. H. J. Kahn. Human values, ethics and design. In *Handbook of Human-Computer Interaction*, J. Jacko and A. Sears eds., pp. 1177–201, Lawrence Erlbaum Associates, 2003.

[Friedman *et al.*, 2006] B. Friedman, P. H. J. Kahn and A. Borning. Value sensitive design and information systems. In *Human-Computer Interaction in Management Information Systems: Foundations*, P. Zhang and D. Galletta eds., pp. 348-72, M.E, Sharpe, 2006.

[Houkes and Vermaas, 2004] W. Houkes and P. E. Vermaas. Actions versus functions. A plea for an alternative metaphysics of artefacts. *Monist* 87, 52-71, 2004.

[Hsieh, 2007] N.-h. Hsieh. Incommensurable values. In *The Stanford Encyclopedia of Philosophy (fall 2007 edition)*, E. N. Zalta ed. http://plato.stanford.edu/archives/fall2007/entries/value-incommensurable/, 2007.

[Kneese *et al.*, 1983] A. V. Kneese, S. Ben-David and W. D. Schulze. The ethical foundations of benefit-cost analysis. In *Energy and the Future*, D. MacLean and P. G. Brown eds., pp. 59-74, Rowman and Littefield, 1983.

[Korsgaard, 1983] C. M. Korsgaard. Two distinctions in goodness. *Philosophical Review* 92, 169-95, 1983.

[Korsgaard, 1986] C. M. Korsgaard. Aristotle and Kant on the source of value. *Ethics* 486-505, 1986.

[Kroes, 2002] P. Kroes. Design methodology and the nature of technical artefacts. *Design Studies* 23, 287-302, 2002.

[Marcus, 1980] R. B. Marcus. Moral dilemmas and consistency. *Journal of Philosophy* 77, 121-36, 1980.

[McLinden and Didion, 1987] M. O. McLinden and D. A. Didion. Quest for alternatives. *ASHRAE Journal* 32-42, 1987.

[Moore, 1912] G. E. Moore. *Ethics.* Oxford University Press, 1912.

[Otto, 1995] K. N. Otto. Measurement methods for product evaluation. *Research in Engineering Design* 7, 86-101, 1995.

[Pahl and Beitz, 1984] G. Pahl and W. Beitz. *Engineering Design: A Systematic Approach.* Springer-Verlag, 1984.

[Petroski, 1996] H. Petroski. *Invention by Design.* Harvard University Press, 1996.

[Pettit, 1984] P. Pettit. Satisficing consequentialism. *Proceedings of the Aristotelian society* suppl. 58, 165-76, 1984.

[Pitt, 2000] J. C. Pitt. *Thinking About Technology: Foundations of the Philosophy of Technology.* Seven Bridges Press, 2000.

[Rawls, 1999] J. Rawls. *A Theory of Justice.* The Belknap Press of Harvard University Press, 1999 [1971].

[Raz, 1986] J. Raz. *The Morality of Freedom.* Oxford University Press, 1986.

[Raz, 2003] J. Raz. *The Practice of Value* (with commentaries by Christine Korsgaard, Robert Pippin, & Bernard Williams; edited and introduced by R.Jay Wallace). Oxford University Press, 2003.

[Schmidtz, 1995] D. Schmidtz. *Rational Choice and Moral Agency.* Princeton University Press, 1995.

[Schön, 1992] D. A. Schön. Designing as reflective conversation with the materials of a design situation. *Research in Engineering Design* 131-47, 1992.

[Scruton, 1979] R. Scruton. *The Aesthics of Architecture.* Methuen, 1979.

[Shrader-Frechette, 1985] K. S. Shrader-Frechette. *Science Policy, Ethics and Economic Methodology. Some Problems of Technology Assessment and Environmental-Impact Assessment.* Reidel, 1985.

[Simon, 1955] H. A. Simon. A behavioral model of rational choice. *Quarterly Journal of Economics* 69, 99-118, 1955.

[Simon, 1956] H. A. Simon. Rational choice and the structure of the environment. *Psychological Review* 63, 129-38, 1956.

[Simon, 1973] H. A. Simon. The structure of ill-structured problems. *Artificial Intelligence* 4, 181-201, 1973.

[Simon, 1976] H. A. Simon. *Administrative Behaviour. A Study of Decision-Making in Administrative Organization.* The Free Press, 1976.

[Simon, 1996] H. A. Simon. *The Sciences of the Artificial.* MIT, 1996.

[Slote, 1984] M. Slote. Satisficing consequentialism. *Proceedings of the Aristotelian Society* suppl. 58, 139-64, 1984.

[Slote, 1989] M. Slote. *Beyond Optimizing. A Study of Rational Choice.* Harvard University Press, 1989.

[Stocker, 1990] M. Stocker. *Plural and Conflicting Values.* Clarendon Press, 1990.

[Sunstein, 2005] C. R. Sunstein. Cost-benefit analysis and the environment. *Ethics* 115, 351-85, 2005.

[Taebi and Kloosterman, 2008] B. Taebi and J. L. Kloosterman. To recycle or not to recycle? An intergenerational approach to nuclear fuel cycles. *Science and Engineering Ethics* 14, 177–200, 2008.

[Taylor, 1911] F. W. Taylor. *Principles of Scientific Management.* 1911.

[Tetlock, 2003] P. E. Tetlock. Thinking the unthinkable: Sacred values and taboo cognitions. *Trends in Cognitive Sciences* 7, 320-24, 2003.

[van de Poel, 2001] I. van de Poel. Investigating ethical issues in engineering design. *Science and Engineering Ethics* 7, 429-46, 2001.

[van de Poel, 2005] I. van de Poel. Engineering design ethics. In *Encyclopedia of Science, Technology, and Ethics* C. Mitcham ed. pp. 622-25, Macmillan Reference USA, 2005.

[van de Poel and van Gorp, 2006] I. van de Poel and A. van Gorp. The need for ethical reflection in engineering design the relevance of type of design and design hierarchy. *Science, Technology & Human Values* 31, 333-60, 2006.

[Wasserman, 1993] G. S. Wasserman. On how to prioritize design requirements during the QFD planning process. *IIE Transactions* 25, 59-65, 1993.

[Williams, 1973] B. Williams. *Problems of the Self. Philosophical Papers 1956-1972.* Cambridge University Press, 1973.

[Winner, 1980] L. Winner. Do artifacts have politics? *Daedalus* 121-36, 1980.

[Zimmerman, 2004] M. J. Zimmerman. Intrinsic vs. extrinsic value. In *The Stanford Encyclopedia of philosophy (fall 2004 edition)*, E. N. Zalta ed. http://plato.stanford.edu/archives/fall2004/entries/value-intrinsic-extrinsic/. 2004.