CHAPTER 7

Ethics and Unintended Consequences of Technology

In Chapter 5, we have seen that technical artefacts are often part of larger sociotechnical systems and that those systems also contribute to determining the consequences of the use of such artefacts. We ended Chapter 6 with the conclusion that technology can result in unintended consequences. These two observations lead one to wonder to what extent engineers are able to predict the consequences of what they design during the actual design process. Those consequences provided an important basis for the ethical questions accompanying the design processes that were dealt with in Chapter 3. What happens if these consequences are not known or are perhaps not foreseeable? What kind of ethical questions does that then precipitate? How can and must an engineer cope with such matters?

7.1 UNINTENDED CONSEQUENCES AND RISKS

Technologies have unintended effects. Aeroplanes crash with fatal consequences. Chemical plants pollute the environment. Security cameras diminish privacy. In other words, such technologies can be dangerous for humans, the natural environment and animals.

The various dangers and other unintended consequences are often not known beforehand, but that does not always mean to say that they are impossible to know. One strategy towards the unintended consequences and dangers of technologies, and addressing them in time, is therefore to try to predict the consequences as good as possible and to expressing the dangers in terms of risks. What are the possibilities and constraints of such a strategy?

'Risk' is usually defined as *the probability of an undesired event* times *the consequences of that event*. If we want to express dangers of a technology in terms of risks, then we need to have reliable knowledge about what precisely the consequences of the technology can be, and we also need to know what the probabilities are that the consequences might materialise. For technologies that have been in use for a long time and where a number of incidents and accidents have occurred, we tend to have reliable information about both factors in the form of accident statistics. We are, for instance, able to estimate the probability of a passenger dying in an air crash in terms of the probability of being killed per kilometre flown. Much the same applies to most other forms of transport.

In the case of new innovations, such reliable information for calculating risk levels is usually lacking. We are often not aware of all the unintended consequences and dangers of a certain new technology. And even if we are aware of the possible undesired consequences, we do not always have enough knowledge about the failure modes: the possible ways in which a new kind of technology

P. Vermaas et al., *A Philosophy of Technology*

can fail. Apart from anything else, in the case of new and innovative technologies, we do not have accident statistics for calculating failure probabilities for the simple reason that no accidents have yet occurred. In such cases, engineers often employ fault trees or event trees in order to estimate the probability of failure. An event tree sketches possible sequences of events that can follow some kind of potential technical failure, like the failure of a plane's landing gear to properly operate. A fault tree sketches the possible series of events that can lead to an accident such as, for instance, the crashing of an aeroplane. By attributing probabilities to individual events in an event or fault tree, the probability of certain accidents can be calculated. Generally, speaking, predicting risks in such a way tends to be less reliable than making use of real accident statistics.

Let us now presume that despite these constraints, we are able to generate more or less reliable estimates of the risks of a certain technology. What then is the use of such risk assessments? A possible way in which we could use such risk assessments is to judge their acceptability. Indeed, the smaller the risk, the sooner we will be inclined to decide that the advantages outweigh the risks of the technology in question.

Although the magnitude of its risk says something about the acceptability of a technology, this magnitude does not tell the whole story. Also other considerations are relevant. To see this, suppose that you want to compare the acceptability of two very different types of technologies. Some engineers and scientists maintain that if according to given risk assessments the risks of these two technologies are roughly the same, then those technologies are also morally equally acceptable. This would, for instance, mean that if you accept one technology, then you would be obliged to also accept the other one with approximately the same risk level. However, this line of reasoning is incorrect for the following reasons. In the first place, not all risk assessments are equally reliable. It is therefore not always appropriate to compare the outcomes of different risk assessments with each other. In the second place, risk assessments are often multi-dimensional whilst the risk comparisons are often one-dimensional, or at least confined to just a few aspects. A very commonly used risk measure is that of the number of fatalities per unit of time. In reality, though, two technologies that are equally dangerous in terms of the number of fatalities per unit of time, can be very different in terms of the danger they pose to human health, the damage to the ecosystem, their economic threat, et cetera. In the third place, the acceptability of risk does not only depend on how big a risk is but also on the question of whether the risk is voluntary and whether people have agreed to take the risk. The risks posed to people that take part in traffic thus tend to be more voluntary accepted than, for instance, the risks posed by a chemical plant near where they live, especially in view of the fact that chemical plants are frequently just built without first consulting the local inhabitants. On the other hand, the risks attached to skiing, for example, are more voluntary than those attached to participating in the road traffic system. On the whole, voluntarily accepted risks are seen as being more morally acceptable than those that are forced upon people. In the fourth place, risks are not just inherently acceptable but are acceptable because they bring with them some kinds of advantages. That means that if the advantages are great, people have reason to accept higher levels of risk. In the fifth place, the acceptability of risks also depends on the availability of alternatives. If there is a possible alternative that carries fewer risks and has no other major disadvantages, then the very existence of that alternative might be grounds enough for viewing the risks of an existing type of technology as undesirable. Finally, the acceptability of a given risk depends on how fairly the risks and benefits are distributed. If there is a certain group in society that only has the disadvantages of the risks without being able to enjoy any of the benefits attached to such risky activities, then that can also make the risk in question morally unacceptable.

In the discussion provided above, there are four important considerations that come to the fore in judging risk acceptability. The first is the question of whether the advantages of the risk-bearing activity outweigh its disadvantages. The second is whether there are alternatives that would, in that respect, score better. The third consideration is that of whether the risk is voluntarily taken and whether or not those involved have agreed to the risk. In ethics, this agreement is often described as giving informed consent. The main idea is that a risk is only acceptable if individuals agree to it after having first been fully informed of the risks, advantages and technological alternatives. A fourth matter of consideration is whether all the possible advantages and disadvantages of the activity-bearing risks are fairly or justly distributed. The value at stake here is that of distributive justice.

Risk assessment may thus be useful when assessing the magnitude of risks. A second reason for possibly wanting to make use of risk assessment might be to limit the risk magnitude at the actual design stage. This is often viewed as an important moral responsibility of engineers. Many professional codes of conduct emphasise that engineers are responsible for designing safe installations. As is stated in the code of conduct of the NSPE, the National Society of Professional Engineers in the USA:[34]

Engineers shall hold paramount the safety, health, and welfare of the public.

Although it would generally seem desirable to minimise the risks during the design process, it is not always possible or, for that matter, desirable to do so. It is not always possible because there is no such thing as artefacts or processes that are absolutely safe. It is not always desirable because reducing a certain risk often brings with it costs or other kinds of disadvantages. For example, a safe car or aeroplane is usually more expensive. If planes are made safer by being made heavier, that means that they will use more fuel which, in turn, will make flying more expensive and will further encumber the environment. In these sorts of cases the pros and cons have to be morally weighed up in order to determine whether it is worth reducing the risks. That is where the above-mentioned considerations play an important part.

If it is decided that the risks should be limited during the design process, then there are different strategies that can be adopted. Following Sven Ove [Hansson, S., 2007], we shall discuss four of those possible strategies.

A first strategy is *inherently safe designing*. This strategy aims at taking away dangers rather than managing them. That can, for instance, be done by substituting various dangerous substances, mechanisms or chemical reactions with less dangerous ones. In order to do that, one needs to have

[34]NSPE Code of ethics, from `http://www.nspe.org/Ethics/CodeofEthics/index.html` Accessed October, 19 2009.

knowledge of the kinds of substances, mechanisms or reactions that constitute a danger. It is not, however, necessary to have details concerning the exact probabilities of the dangers arising. To follow this strategy you need to have some but not full knowledge about the risks of a given technology.

A second strategy that can be adopted is that of building in *safety factors*. This means creating a structure so that its expected strength exceeds the expected load with a certain factor, the safety factor. The risk of failure of that particular construction is thus reduced. It is also a way of coping with the possible uncertainties attached to the expected loads and the predicted strength of the construction. Indeed, there is a tendency to think more in terms of the expected probability of failure of a construction or of components – that may not exceed certain thresholds – than in terms of the safety factors since the former results in less over-dimensioning and thus cheaper constructions. This, however, requires reliable knowledge of both the expected loads and the actual strength of the construction, and it eliminates uncertainty margins which are incorporated in a design if it is based on safety factors.

A third strategy is that of incorporating *negative feedback*, which involves designing installations in such a way that if an installation fails or if a operator loses control, a mechanism will be activated that will automatically switch off the installation. A good example of this is the driver's safety device (DSD) in trains which ensures that the train comes to a standstill if the driver falls asleep or loses consciousness. Note that up to a certain point, this strategy can cover partially unknown risks: a mechanism can be triggered if, for instance, a certain poisonous substance is in danger of escaping from a chemical plant. One does not need to know beforehand what the probabilities of such an incident occurring are in order to be able to design an effective negative feedback mechanism.

A final strategy can be to design *multiple independent safety barriers*. In that way, one creates a chain of different independent safety barriers to ensure that if the first barrier fails, there will be others to back it up. In this way, the probability of danger arising is reduced. With this strategy, it is once again the case that one does not need to know beforehand what precisely the probabilities of an incident occurring are or what the exact causes might be. In order to be able to design safety barriers that are actually independent, one really needs to have insight into the possible failure modes; otherwise, there is always the danger of a failure mode simultaneously undermining different safety barriers.

This brief overview of strategies shows that during designing, one can take measures to reduce possible dangers, even if one is not exactly aware of the dangers and one is not able to precisely quantify them in terms of risks. All the same, one must have some idea of what can go wrong and what unintended effects can occur if one is to successfully apply the referred to strategies. In other words, it is more or less impossible, during the design process, to account for unintended consequences which cannot possibly be foreseen beforehand. The unintended consequences variant of technological determinism discussed in Section 6.2 posits that such unintended and unforeseen consequences cannot be avoided in technological development. We concluded in the previous chapter that this assumption regarding the unintended consequences variant within technological determinism is indeed plausible. The conclusion must therefore be that during the design process, one cannot take into account *all* the possible unintended consequences of a technology because some of the consequences do not reveal themselves beforehand.

7.2 SOCIOTECHNICAL SYSTEMS AND RESPONSIBILITY

One reason why one cannot account for all the consequences of technology beforehand is because there is a limit to the amount of time that can be spent on finding out about those consequences. Quite where one places that limit is also an ethical question because the decision to finishing a design process quickly, may have the consequence that the probability increases that users and others are burdened with the negative consequences of technology.

There is, however, another reason why one cannot always predict the unintended consequences of technology and that is this: during the design phase the consequences are indeterminate because they also partly depend on the actions of other *actors* besides the designers, like the users. This especially applies when we take sociotechnical systems into account, which – as we saw in Chapter 5 – depend for their proper functioning on all sorts of actors, such as operators.

If undesired consequences arise in a sociotechnical system it is, in many cases, not possible to simply trace the cause of those consequences back to one actor who might have been able to foresee and prevent such consequences and who can, thus, be held responsible for them. It is much more likely that the consequences will depend on the actions of a number of actors and on the constellation of the sociotechnical system as a whole. This was a point that clearly emerged from the discussion in Chapter 5 on the mid-air collision above Überlingen (see Section 5.5). The final upshot would appear to be that one often cannot indicate who is responsible for certain undesired consequences. This is also sometimes known as *the problem of many hands*: because there are so many people who play a small (and difficult to pinpoint) part in the chain of the events that it is difficult to establish who is responsible for what.[35]

The problem of many hands is not just a problem because, in retrospect, no one can be held responsible but also because, apparently, no one feels the obligation to endeavour to prevent such consequences occurring, and so we do not learn from our mistakes. All the same, there are times when it is possible to hold one or more individuals responsible for an accident. To illustrate this point, we shall now examine an aeroplane accident that took place in 2007, in Brazil.

On 17th July 2007, an Airbus A320 overshot the end of the runway when landing at Congonhas International Airport in São Paulo. It ploughed across a motorway and finally came to rest, next to a petrol station, in a warehouse belonging to the Brazilian airline company TAM. It then exploded. Some 199 people were killed in the incident, 12 of whom were on the ground.[36]

The initial conclusion was that the accident was attributable to the relatively short runway and to wet weather. The runway in question had been resurfaced not long before, which meant that gullies had not yet been created to deal with excess water on the runway and thus prevent *hydroplaning*. Already airport safety there had previously been questioned, especially in wet weather.

[35] Bovens, M. [1998].
[36] `http://en.wikipedia.org/wiki/TAM_Linhas_A%C3%A9reas_Flight_3054`, last visited on 7th December 2007.

In February 2007, a Brazilian judge had imposed a flying ban at Congonhas for a number of aeroplane types, following complaints lodged by pilots about rainwater on the runways, which had negatively influenced the planes' breaking performance. The airport, the Brazilian airline company TAM and the Brazilian Civil Aviation Authority all contested the ban, and so within a day it was dropped following a higher court overruling. The A320 did not fall under the original ban because according to Airbus, its braking distance was shorter than that of the banned aircraft types.

When it became clear that during the landing of the fatal flight water accumulation on the runway was limited, the investigation focused on other possible causes. What soon emerged was that the thrust reverser of the right engine had not operated during the landing procedure. Thrust reversal is deployed during touchdown to aid deceleration. The authorities knew about the problem, but according to Airbus and TAM, the airline company, it was not unsafe to land with a defective thrust reverser. According to reports – that were vehemently repudiated by TAM – there had been landing problems with the same plane only a day before.

From analyses of the *Flight Data Controller* and the cockpit conversations, it became evident that the pilots had been aware of the problem with the inoperative thrust reverser in the right engine. During the landing roll, they had switched the right engine to 'climb', probably to prevent the defective reverser from starting to operate. Because of that, the spoilers on the wing did not work. During landing, those spoilers are used to increase the plane's air resistance or drag while at the same time reducing the lift factor so that the plane is forced down onto the runway and can brake more easily. With the Airbus A320, the *spoilers* only work properly if the engines are in the 'idle' position, but as the right engine was in the climb mode during that landing procedure, the spoilers did not come into play. Most probably, that was the cause of the accident or, at least, one of the major contributing factors. Interestingly, after the disaster, Airbus ordered that with that particular series of aircraft, all engines should be switched to the idle mode in readiness for touchdown.

Who, then, was responsible for the Brazilian crash? Before answering this question, it is useful to first think about the conditions under which someone can be held responsible for something. Over the course of time, philosophers have given various answers to this very question. In many philosophical discussions, there are a number of conditions that recurrently arise. The following conditions are the ones that are most frequently mentioned: [37]

1. The action taken by somebody for which (s)he is held responsible must have been consciously undertaken;
2. There must be a causal connection between the action taken and the ensuing consequences for which someone is held responsible;
3. The person must have foreseen or at least have been able to foresee the consequences;
4. The person could have taken a different course of action;
5. The action taken was wrong or otherwise blameworthy, and that mistake or blameworthiness contributed to the negative consequences.

[37] See for example Bovens, M. [1998] and Fischer and Ravizza [1993].

In the light of this information, look once again at the actions of the pilots. The relevant actions of the pilots involved consciously adjusting the right engine to the 'climb' mode or, more precisely, leaving it in that position. In so doing, they fulfil the first condition. Thanks to this course of action, though, the spoilers did not function in the way that they should have done, so the plane failed to decelerate sufficiently, flew off the runway, crashed and went up in flames. Condition 2, concerning the existence of a causal connection, is also satisfied. Whether or not the pilots could have foreseen the disaster, which is the third condition, is less clear. One presumes that the pilots were not able to anticipate the consequences of their actions: they did not deliberately crash the plane. The real question is whether they could have or should have been able to anticipate the consequences. One argument that can be levelled is that according to the flying instructions for that particular type of Airbus, the engines should be in the 'idle' mode when landing; the onboard computer also backed up those instructions, but the pilots chose not to follow them. It could be argued that pilots should know that failing to follow the flying instructions or to observe the directions of the onboard computer can lead to a disaster. The question as to whether the pilots could have done any differently, the fourth condition, is disputable. At first sight, it would seem obvious that the pilots could have done something differently: after all, they could have allowed the right engine to idle instead of climb. Nevertheless, one could argue that the pilots were forced to make a quick decision in what was not an everyday situation (with a malfunctioning thrust reverser in the right engine), and that it was perhaps not obvious that it would make a difference in that situation if the right engine was in the climb mode. On the other hand, one could argue that it is part of the pilot's job to be able to react quickly and appropriately in such situations. As far as the fifth condition is concerned, the pilots clearly breached a norm, which means that their actions could be labelled wrong. They failed to abide by the flying instructions given in the manual (see also Figure 7.1). On top of everything else, they ignored the directions given by the onboard computer. Still, one can argue that this was an unusual situation in which the pilots had to think and react fast, so they could not be blamed for what they did wrong.

Apart from looking at the pilots, one can also look at other actors that could possibly have borne some responsibility in this case such as the air traffic controllers, the aviation authorities, the plane's maintenance technicians and the engineers, who designed the aeroplane and drew up the flying instructions. Just think for a moment of the engineers and go once again down the list of the five conditions.

One may presume that the engineers did make the relevant technical decisions in a conscious fashion; they knew what they were doing or should at least have been aware of what they were doing. They therefore satisfied the first condition of responsibility. To what extent the second condition was also met – that of a causal connection between the actions of the engineers and the disaster – is not so clear. If the design had been different, the disaster may possibly have been averted. In addition, the manual could have stated that the aeroplane should only be allowed to take off if both thrust

[38] *Flights Control, Flight Crew Operating Manual, Airbus A 320*. Available at `http://www.smartcockpit.com/data/pdfs/plane/airbus/A320/systems/A320-Flight_Controls.pdf`, last visited on 23rd June 2008.

AIRBUS TRAINING A320 SIMULATOR FLIGHT CREW OPERATING MANUAL	FLIGHT CONTROLS	1.27.10	P 12
	DESCRIPTION	SEQ 001	REV 37

GROUND SPOILER CONTROL

Spoilers 1 to 5 act as ground spoilers.
When a ground spoiler surface on one wing fails, the symmetric one on the other wing is inhibited.

Arming
The pilot arms the ground spoilers by pulling the speedbrake control lever up into the armed position.

Full extension
The ground spoilers automatically extend during rejected takeoff, at a speed greater than 72 knots, or at landing when both main landing gears have touched down, when :
R · Ground spoilers are armed and all thrust levers are at or near idle, or
· Reverse is selected on at least one engine (other thrust lever at or near idle), if ground
R spoilers were not armed.

Figure 7.1: A section of the '*flight crew operating manual*' for the A320.[38]

reversers are operating properly. That would probably have been sufficient to prevent the accident. To conclude, it would therefore seem that the engineers' manuals did have a causal connection with the accident. To what extent the engineers could have foreseen this – the third condition – is once more not so clear. When designing the aircraft, the engineers probably did not take into consideration a scenario like that which occurred during the Brazilian air disaster. One can really question whether the engineers should perhaps have borne in mind the possibility of something like that happening.

In this particular case, it would seem that the fourth condition was met: the engineers probably could have dealt with matters differently. There are, at any rate, no indications that they were forced to adhere to this design or were somehow pressured to do so by their superiors. The mention, in the flight manual, that aeroplanes can take off if one of the thrust reversers is out of order is probably dominated by commercial considerations because it means that a plane can put in more flying hours before requiring maintenance.

The fifth condition would imply that the actions of the engineers might have been incorrect or blameworthy. In this connection, the following stipulation laid down in the NSPE code of conduct is of relevance[39]:

> *(II.1.b) Engineers shall approve only those engineering documents that are in conformity with applicable standards.*

As far as it is known, the aeroplane did satisfy all the safety regulations. It is possible, though, that the design endangered the well-being of society in a different way. In this light, the following

[39]NSPE. 2010. *NSPE Code of Ethics for Engineers*. National Society of Professional Engineers, USA 2007 [cited 10 September 2010]. Available from `http://www.nspe.org/Ethics/CodeofEthics/index.html`.

stipulations in the professional code of conduct of the Dutch association of engineers KIVI-NIRIA are relevant[40]:

> *(1.3.1) The engineer must carefully evaluate the safety and reliability of the systems that have been designed and for which he is responsible before going on to give his approval.*
>
> *(1.3.2) The engineer must provide manuals (containing the relevant standards and quality norms) so that the user is able to make safe and correct use of the products and systems for which the engineer is responsible.*

Although there are no clear indications that the engineers violated any of those stipulations, there are a number of questions that could be asked. One might, for instance, query whether the engineers had sufficiently evaluated the safety of the system. As has already been suggested, one could ask whether the engineers might not have been able to anticipate events of the kind that took place in Brazil in 2007. If they had, then they probably would have designed the system slightly different. Regarding the second directive, directions for use had been drawn up that would also have made it possible for the pilot to land safely in these circumstances. As was suggested above, one might ask whether those instructions for use were not too complex; the pilots had to react in a very short space of time. From the point of view of safety, the manual should perhaps also have mentioned that planes should only be allowed to take off if both thrust reversers are in working order. Finally, one might wonder to what extent the engineers were involved in or were able to influence this particular aspect of the instructions for use.

There are therefore arguments for holding responsible both the pilots and the engineers who designed the aeroplane. However, there are also arguments against holding these particular actors responsible. It has to be acknowledged that there were many other actors involved who, directly or indirectly, may have contributed to the disaster. For instance, the Brazilian Civil Aviation Authority had been held responsible for previous accidents; some even spoke of a safety crisis within the Brazilian aviation sector. There had also been complaints about lack of safety at the airport where the accident happened. One might even go as far as to assert that the cause of the crash needed to be sought in the sociotechnical system as a whole. This was definitely so in the case of the accident discussed in Chapter 5 where there was a mid-air collision. In these kinds of situations, the matter of responsibility is more diffused. As we also saw in Chapter 5, it is not usually one actor that designs a given sociotechnical system; the system is much more often the – partly unintended – result of the actions of many actors.

We may therefore conclude that if engineers design components or artefacts that are destined to operate within a wider sociotechnical system, then the social consequences of those components or artefacts are hard to anticipate in the design phase because they partly depend on the actions of other actors and the character of the sociotechnical system. It would, however, be too simple to conclude that, in this case, the responsibility for the possible negative consequences should solely be

[40] `http://www.kiviniria.net/CM/PAG000002804/Gedragscode_2006.html`, last visited on 23rd June 2008. Our translation.

placed on other actors. We have seen that in the case of the aeroplane incident in Brazil, it is certainly possible to argue that the engineers also bear a certain amount of responsibility. We also saw that in sociotechnical systems, in general, it is often not easy to identify who is responsible for what. That could lead one to surmise that no one is responsible. Yet that would seem to be a rather unattractive outcome, given the widely supported desire to develop technology in a socially responsible way, and in the light of the knowledge that if all concerned attune their actions to those of others, it may well be possible to prevent certain negative consequences. In the next section, we shall therefore consider how an engineer can contribute to responsible technological development provided that (s)he recognises that technology sometimes has unintended and undesired consequences that could not have been fully foreseen in the design phase and provided that (s)he recognises that those consequences also depend on other actors as well as the sociotechnical system within which the artefacts operate.

7.3 TECHNOLOGY AS A SOCIAL EXPERIMENT

In the light of the above-mentioned findings, one might wonder if it is a good idea to address ethical questions already during the design process in the way suggested in Chapter 3. Might it not be better to wait until the consequences become clearer? We would then, at least, know what we are talking about.

Letting the consequences of technology manifest themselves after the designing of a technology, also has a number of disadvantages. It is then not only the case that these undesired consequences will arise, but also the costs of preventing the consequences from arising again in the future will often be much higher than when matters are tackled in the design phase. Returning to the drawing board is not only expensive because something new has to be designed but also because – once the consequences manifest themselves – the relevant technology will already have become embedded in society. That means that users will already have become used to it and will have adapted their behaviour to that technology; it also means that regulations and other social institutions will have become adjusted to that particular technology. Breaking open such a level of embeddedness is often not only difficult (if indeed possible at all) but also expensive. Furthermore, as we saw in the last section, if negative effects are manifested in a sociotechnical system, it is often not very easy to determine who exactly is responsible for those effects. That, in turn, can lead to situations in which people avoid to address those undesired effects once they have manifested themselves.

When dealing with the possible unknown effects of technology, we are therefore confronted with a dilemma. On the one hand, we often do not know, in the design phase of a given type of technology, the possible future consequences that we have to take into account when designing. On the other hand, by the time those consequences become manifest, the technology in question is already operational and the costs of redesigning or making other adaptations within the sociotechnical system are often high and difficult to realise. This, in fact, is the Collingridge's control dilemma that was already described in Chapter 6.

Although there is no easy solution to the Collingridge's control dilemma, one might not conclude from the mere existence of the dilemma that responsible technological development during the design phase is impossible. In fact, in many cases one does, already in the design phase, really have some knowledge of how a certain kind of technology will probably be implemented and of the sort of consequences which that will have. The ideas discussed in Chapter 3, relating to designing responsibly, such as the notion of value sensitive designing, have not therefore suddenly become nonsensical. The analysis given in this chapter shows that those ideas are not enough to prevent all the possible negative consequences. For example, when implementing a certain technology, it might become apparent that a value sensitive design does not work, or that in practice it gives rise to a new value conflict. In a certain respect, we already noticed that in the example discussed in Chapter 3 on the development of aeroplane engines. Despite the fact that it has become easier to develop ever-more quiet aeroplane engines, as far as society is concerned the noise factor problem has only been resolved to a limited extent because flight movements are increasing. This increase is a development that has been partly made possible by the creation of quieter engines: it has meant that noise hindrance has become less of an issue and so the aviation sector and fleets have been able to expand. We thus see that technologies that seem, on the drawing board, to be capable of resolving certain problems do not always manage to do that in practice. This obviously does not mean that one should therefore stop designing in a value sensitive way; it means that one must simply be aware of the boundaries and limitations of such an approach.

To a certain extent, it is probably possible, during the design phase, to anticipate the fact that the consequences of technology will be partly unknown or undetermined. This can, for example, be achieved by endeavouring to come up with designs that are robust, flexible and transparent.[41]

A design may be said to be *robust* if it functions well in different and preferably also unforeseen circumstances. The strategies discussed in Section 7.1, for making designs more resilient to risks and dangers contribute to robust designing. There we saw that up to a certain point such strategies are also suitable for dealing with unknown factors. Another possible way of making a design robust is by building in redundancy, for instance, by creating certain vital components in duplicate so that if one fails the other can take over. A sociotechnical system can, for example, be made robust by ensuring that if one component fails, it does not immediately bring down the entire system.

A design can be *flexible* in different ways. It can be flexible in a physical sense, which means that it can relatively easily be adapted or rebuilt. In this respect, many modern appliances are either not very flexible, or they are not flexible at all; for instance, they often cannot be opened so that a component part can be replaced. A design can also be termed flexible in a functional sense: it can fulfil different functions. Likewise, sociotechnical systems can display greater or lesser degrees of flexibility; not only in the sense that they can fulfil different functions but also from the point of view that the various component parts can work together in different ways to arrive at a system that functions well. Flexibility makes it possible to react and adapt to unexpected eventualities. Finally, in a flexible system, it is generally less expensive to make adjustments to dispel the undesired consequences that only manifest themselves during use or implementation.

[41] For similar and other suggestions, see Collingridge, D. [1992].

Transparency means to say that a technical artefact or a sociotechnical system and the way in which it works, is clearly understandable to the users and, in the case of sociotechnical systems, also to the operators. The advantage of transparency is that if something goes wrong or if undesired effects arise, it is easier to trace things back to a certain component or mechanism. In that way, the undesired effect does, in many instances, become easier to combat. In a sociotechnical system, transparency also increases the insight that actors have into their contribution to certain effects. In that way, it becomes easier to establish who is responsible for what, and the actors will also feel more responsible because it will be clearer to them just how their actions contribute to certain effects.

Even though a designer can anticipate different unintended and, as yet, undetermined consequences, (s)he cannot possibly also prevent all those negative consequences from arising. From that point of view, technological development always has an experimental character: some social consequences of a given technology only emerge when that technology is implemented. This is partly down to the fact that society also often changes when that technology is embedded; technology and society codevelop as it is phrased.

In creating their own environment, human beings are thus, in essence, experimenting beings. There can be quite a lot at stake in these experiments though. One need only think in this connection of CO_2 emissions, which are contributing to what may turn out to be an irreversible climate change situation, the possible consequences of a nuclear war and the possible consequences of human enhancement by intervening with people's physical and psychological constitution.

From the ethical angle, one may question when such experiments are acceptable and when they cease to be acceptable. There are certain philosophers, like Hans Jonas, who put the case that extreme precautions should be exercised with all the kinds of technology that could conceivably threaten the survival of our species.[42] Another proposition is that such experiments are acceptable if those who possibly stand to suffer from their consequences agree to the experiments going ahead.[43] This notion is similar with the principle of 'informed consent' that is widely used in medical practice. It is the idea that whenever people have to undergo surgery or whenever they participate in a medical experiment, they are first informed as fully as possible of all the possible inherent risks and dangers so that they can decide if they want to undergo the surgery or participate in the experiment. In medical practice, informed consent tends to be implemented on an individual basis; in the case of technological development, that would seem to be more difficult because the relevant 'experiments' often involve large groups all at once. A question which then arises is whether all the members of the group in question have to agree or whether the consent of the majority is enough. Further problems arise in the case of technologies that have possible repercussions for future generations. People who have not yet been born cannot, of course, be asked for their informed consent.

[42] Jonas, H. [1984].
[43] See, for instance, Martin and Schinzinger [2005].

7.4 CONCLUSION

Technological development has an experimental character: unintended consequences cannot be entirely predicted or, for that matter, avoided. This is not just something that is caused by our limited knowledge capacity but also by the fact that the unintended consequences are often the result of the actions of many actors within a sociotechnical system. The implications of this observation are that responsibly developing technology is more complicated than was presumed in Chapter 3. This does not mean to say that one can forget all the suggestions made in Chapter 3; in many cases, they are still relevant, even if success is not guaranteed. In this chapter, we have also demonstrated that engineers can anticipate the occurrence of unintended effects by endeavouring to come up with designs that are robust, flexible and transparent. The experimental nature of technology, finally, gives rise to the ethical question of the conditions under which such experiments are morally acceptable.

7.5 A FEW MORE ISSUES

In this chapter, one of the things we have discussed is why the magnitude of a risk does not determine the acceptability of such a risk. This means, for instance, that if you accept a technology with a certain risk while another technology carries lower risk, you do not automatically have to accept the second technology. Now apply the reasons just mentioned to the comparison between the risks attached to car driving and those attached to nuclear energy. Imagine that the risks of car driving are larger than those posed by nuclear energy and that we find the risks attached to car driving morally acceptable. Why then are the risks of nuclear energy not also necessarily morally acceptable?

A second and more difficult question has to do with the acceptability of social experiments in technology. We have briefly referred to the principle of informed consent as one way of dealing with this particular matter. We also presented two practical objections one could have to applying this principle to technological development as a social experiment. You could also ask yourself if the principle is desirable from a moral point of view. Is it not a much too strict principle? If everyone first has to agree, will technological development still be possible? Will all innovation not then grind to a halt? Is there perhaps another moral point of departure for social experiments that may not have this disadvantage?