

# **Persuasive Technology and Moral Responsibility**

## *Toward an ethical framework for persuasive technologies*

**Paper for Persuasive06,  
Eindhoven University of Technology**

*Peter-Paul Verbeek,  
University of Twente, The Netherlands  
p.p.c.c.verbeek@utwente.nl*

### **1. Introduction**

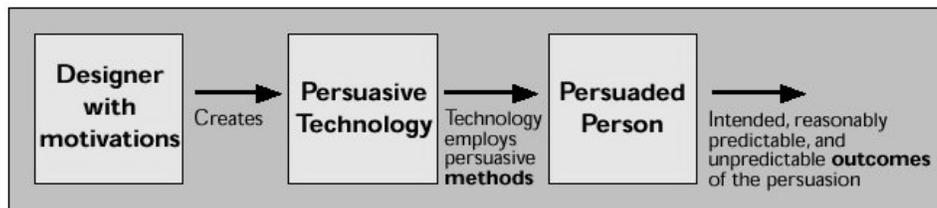
How to deal with ethical issues surrounding persuasive technologies? Despite the obvious good intentions behind technologies like the FoodPhone (helping to fight obesity), the HygieneGuard (motivating people to wash their hands after using the toilet), and the EconoMeter (helping people to drive their cars more economically), ethical questions abound. Persuasive technologies can have undesirable consequences, employ problematic methods of persuasion, or persuade users to do things which cannot be morally justified. Moreover, the very practice of technological persuasion can raise moral questions in itself, because it might be seen as a threat to human autonomy, a source of moral laziness, or an anti-democratic force in society which lets designers rather than representatives of the people steer our behavior.

This paper will address the ethical aspects of persuasive technologies and their designs in three steps. First, it will localize the most important points of application for moral reflection during the design of persuasive technologies. By integrating the concept of 'persuasive technology' with the concept of 'technological mediation', as developed in theoretical frameworks from the field of persuasive technology and from the philosophy of technology, it will be possible to identify three such points of application. Second, this paper will deal with the process of moral reflection itself. A method will be elaborated to facilitate moral decision-making during the design of persuasive technologies, using the points of application identified in section 2. Third, the paper will discuss a number of ethical issues which are a bit more external to the practice of designing persuasive technologies, and which mainly concern the moral acceptability of specific persuasive technologies and of the phenomenon of persuasive technology as such.

### **2. Points of Application for Moral Reflection**

In 1999, Daniel Berdichewsky and Erik Neuenschwander published a highly relevant and helpful framework for evaluating the ethical aspects of persuasive technologies. Central to

their framework is the interaction of persuader, persuasive technology, and persuaded person. All elements in this interaction embody a specific point of application for moral reflection: the motives of the designer, the methods of persuasion employed in the technology, and the intended and unintended outcomes of the persuasion. The framework is represented in figure 1 (reproduced from Berdichewsky and Neuenschwander, 1999, p. 54).



**Figure 1:** Berdichewsky&Neuenschwander's framework

In order to cover all relevant ethical aspects of persuasive technology, this framework could benefit from some modification and augmentation, though. First of all, from the point of view of contemporary philosophy of technology, technological persuasion should be seen as part of the more encompassing phenomenon of *technological mediation*. This concept, as developed by several theorists in the field (Ihde 1990; Latour 1994; see also Verbeek 2005) can contribute to a better understanding of the role of persuasive technologies in their use contexts. The concept of mediation indicates the ways in which technologies help organize relations between users and their environment, by helping to shape both human perceptions and human actions. Augmenting Berdichewsky & Neuenschwander's framework with insights from the theory of technological mediation we can, therefore, will result in a broader understanding of the effects of persuasive technologies than “intended persuasions” versus “other, unintended outcomes”.

Secondly, after having expanded ‘persuasion’ to ‘mediation’ more needs to be done to explicitly include the unintended effects of technologies in ethical reflection and decision-making. In Berdichewsky & Neuenschwander's model, moral responsibility of designers for such unintended outcomes remains underexposed. Yet, these outcomes are highly important, since the theory of mediation shows that they occur always and inevitably. Integrating the conceptual framework of mediation theory with Berdichewsky & Neuenschwander's model will therefore make it possible to better understand and predict unintended outcomes, and to incorporate them in moral decision-making during the design process.

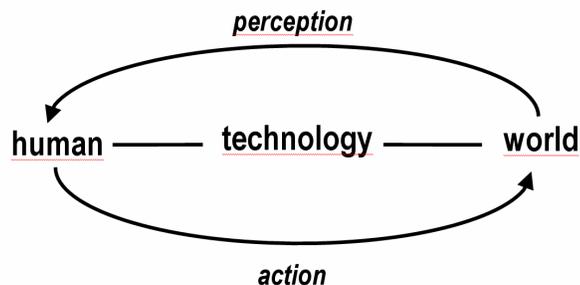
### 2.1 Persuasion and mediation

A first step to be taken for analyzing the moral aspects of persuasive technologies, is to conceptualize their impact on human beings. This impact can entail more than only the behavior into which users were to be persuaded according to the intentions of the designers. Many, if not all, persuasive technologies have unintended effects too. A helpful

framework for getting a better understanding of these unintended effects is provided by contemporary philosophy of technology. Technological persuasion can be seen as a specific manifestation of the more encompassing phenomenon of *technological mediation*. This concept indicates the ways in which technologies inevitably and often implicitly help to shape human actions and perceptions by establishing relations between users and their environment.

The central idea in the theory of mediation, as articulated in Verbeek (2005, 2006b), is that technologies should not be understood as functional *instruments*, but as active *mediators* in relations between humans and reality. This mediating role is made possible by the specific ways in which technologies in use are present to their users: such technologies are not the *terminus* of human perception and action, but rather withdraw from our attention, making possible specific experiences and practices. When driving a car, e.g., our attention is not with the car itself, but with the road, the surroundings, and the traffic signs. Only when we still have to learn to drive, or when something goes wrong, our attention is drawn to the car itself. But as soon as we know how to drive and everything works properly, we *embody* the car, to use a concept of Don Ihde (1990). Rather than interacting with it, we form an association with it, and interact *through* the car with the world around us.

This implies that, rather than being mere instruments for realizing human goals, technologies-in-use help to shape specific relations between humans and reality. Technological mediation is precisely this capacity of technology: technologies can mediate between humans and reality, by establishing specific relations between both. This phenomenon of technological mediation has two dimensions, each of them pertaining to one aspect of the relations between humans and reality. First technologies help to shape how reality can be present for human beings, by mediating human *perception* and *interpretation*; second, technologies help to shape how humans are present in reality, by mediating human *action* and *practices*. The first dimension can be called *hermeneutic*, since it concerns meaning and interpretation; the second is *pragmatic*, since it concerns human activities.



**Figure 2:** Technological Mediation

The mediatory capacities regarding human action are often indicated with the concept of ‘*scripts*’, as developed by Madeleine Akrich (1992) and Bruno Latour (1992). Just like the script of a theater play or a movie tells the actors what to do at what moment, technologies can prescribe their users how to act when they are used. Classic examples are the speed bump which prescribes car drivers to slow down, and the turnstiles at metro stations prescribing travelers to buy a ticket before entering the train. Mediation of action usually takes shape in encouraging or inviting specific forms of action, while discouraging or inhibiting other actions. The mediating role of technologies in human experience is often indicated as ‘*technological intentionality*’: technologies are ‘directed’ at specific aspects of reality and help to shape our perceptions and interpretations accordingly. Such intentionalities usually take shape in amplifying specific perceptions while reducing others, thus providing a specific basis for interpreting what is perceived.

Technological persuasion can be seen as a specific manifestation of this phenomenon of technological mediation. Most persuasive technologies actually perform a hermeneutic form of mediation, by shaping experiences and interpretations which inform behavior, like the FoodPhone which helps to develop new interpretations of food which inform people’s eating practice. Not all ways in which technological mediation helps to shape human actions needs to take shape as *persuasion*, however. Elsewhere (Verbeek 2004a) I have argued that three behavior-steering forms of mediation can be distinguished. First of all, technologies can *force* people to behave in specific ways, like a speed bump which leaves car drivers hardly any choice regarding their speed. Second, technologies can try to *persuade* users into specific actions, like the EconoMeter in cars which gives feedback about the energy-efficiency of a driver’s driving style. And thirdly, technologies can *seduce* users into a specific form of behavior, like specific road designs (curves, markings) which make it more attractive to drive with a specific speed at specific places.

There is a second connection between persuasive technology and mediation, though, which is at least equally important. Not only can technological persuasion be seen as a specific form of technological mediation, but the persuasive function of technologies can also have a mediating effect *itself*. The FoodPhone, to stick to the example, may persuade its users to develop more healthy eating habits, which can be seen as a hermeneutic form of mediation, shaping human interpretations of what they are eating. This persuasive effect, however, can also play a mediating role itself in the relation between humans, their food, and their social environment. Beside having the desired effect of stimulating a more healthy eating pattern, the FoodPhone can e.g. make eating something stressful; it can stimulate humans to interpret their health exclusively in terms of their eating pattern while neglecting the importance of other factors like having enough exercise; and taking pictures of all food consumed will definitely organize social relations at the table. Such mediation effects of technological persuasion itself also need to be taken into account in moral decision-making regarding persuasive technologies.

## **2.2 Expanding the responsibility of designers**

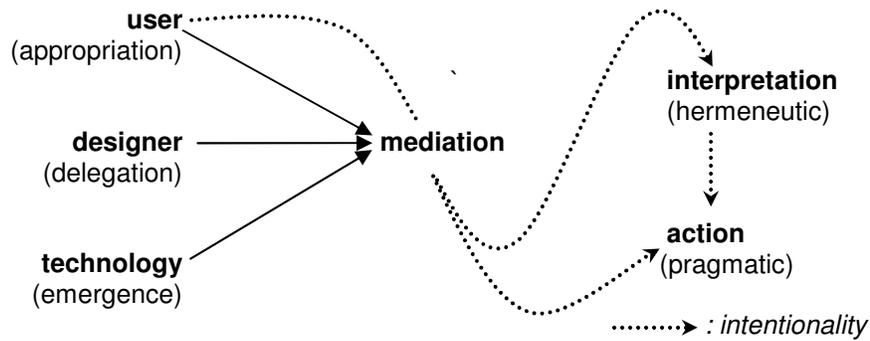
Having analyzed the concept of technological persuasion as a specific manifestation of

technological mediation, a second step needs to be taken to augment Berdichewsky and Neuenschwander's model for the ethics of persuasive technology. This second step concerns the uncertainty surrounding the eventual effects of persuasive technologies. The main reason for this uncertainty is that persuasion and mediation cannot be seen as intrinsic properties of technologies. Mediation theory, with its backgrounds in both phenomenology (Ihde) and actor-network theory (Latour) approaches technologies as entities that have no fixed identities, but that rather possess *interpretive flexibility* (Bijker) or *multistability* (Ihde). The central idea is that technologies need to be interpreted and appropriated by users in order to be used. Only in specific relations between humans, technologies, and reality can technologies be constituted as objects that are to be used specifically for something. The typewriter, for instance, was originally designed not as a piece of office equipment but as a device for helping visually impaired people to write. Scripts and intentionalities, therefore, come into being only within specific contexts of use, rather than having a predefined existence.

This implies that there is no direct relation between the activities of designers and the mediating role of the technologies they are designing (cf. Verbeek 2004b). The multistability of technologies makes it difficult to predict the ways in which they will influence human actions, and therefore also to evaluate this influence in ethical terms. First of all, technologies can be used in unforeseen ways, and therefore have an unforeseen influence on human actions. The energy-saving light bulb is a good example here, having actually resulted in an increased energy consumption since such bulbs often appear to be used in places previously left unlit, such as in the garden or on the façade, thereby canceling out their economizing effect (Steg 1999; Weegink 1996). Second, unintentional and unexpected forms of mediation can arise also when technologies do get used in the way their designers intended, but meet unforeseen use practices. A good example is the revolving door which keeps out not only cold air but also wheelchair users.

In short, designers play a seminal role in the coming about of technological mediations (and persuasions), but not the only role. Users, with their interpretations and specific forms of appropriation, also have a part to play; and so do technologies themselves, giving rise to unintended and unanticipated forms of mediation. The mediating role of technologies is not only the result of the activities of the designers, who inscribe scripts or delegate responsibilities, but also depends on the users, who interpret and appropriate technologies, and on the technologies themselves, which can evoke 'emergent' forms of mediation.

This state of affairs has important implications for the ways in which designers can take responsibility for their designs. It means that the mediating role of technologies in the interpretations and actions of users, which together constitute their behavior, cannot be reduced to explicit design specifications, but also depends on specific user interpretations and characteristics of the technology designed. Figure 3 illustrates these complicated relations in the mediation of actions and interpretations.



**Figure 3:** Sources of mediation and user intentionality

This figure makes clear that in all human actions and all interpretations informing decisions of how to act, three forms of agency are at work: (1) the agency of the human being performing the action or making the moral decision (in interaction with the technology), but also appropriating the technological artifact in a specific way; (2) the ‘agency’ of the artifact mediating these actions and decisions, sometimes in unforeseen ways; and (3) the agency of the designer who – either implicitly or in explicit delegations – gives a specific shape to the artifact used, and thus helps to shape the eventual mediating role of the artifact. Taking responsibility for technological mediation, therefore, comes down to entering into an interaction with two other forms of agency – the agency of future users and the agency of the artifact-in-design – rather than acting as a ‘prime mover’ (cf. Smith 2003).

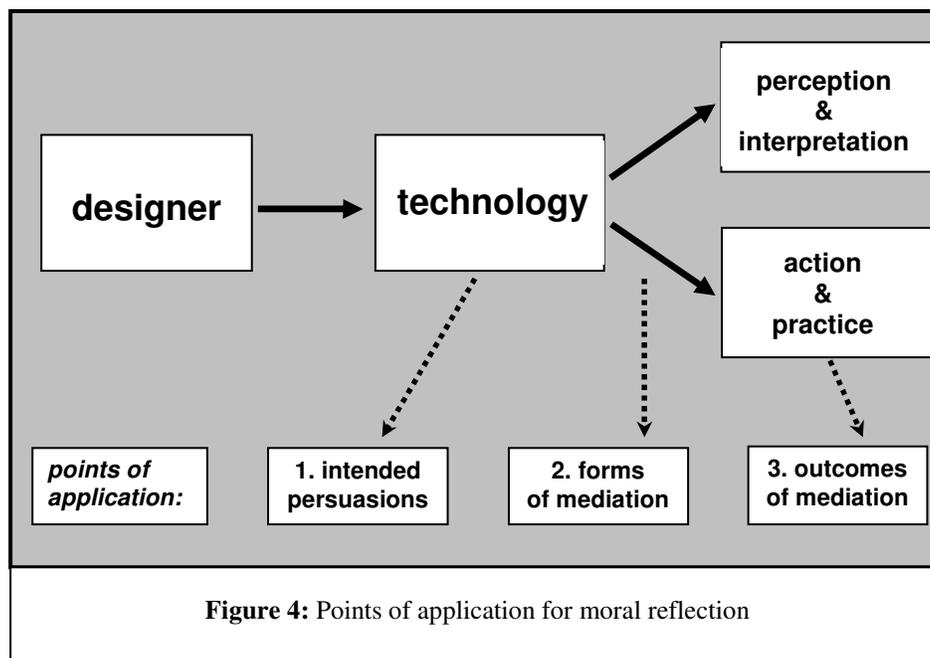
The theory of mediation makes it possible to take this insight beyond the conclusion that persuasive technologies have merely ‘intended’ and ‘unintended’ effects. Performing a mediation analysis (cf. Verbeek 2006a, 2006b) can be a good basis for making an *informed prediction* of the future mediating role of a technology. In this way, the moral responsibility of designers can be expanded to also cover the unintended outcomes of technologies, at least to the extent to which they can reasonably be foreseen. The concept of mediation makes it possible to evaluate new technologies not only in terms of the quality of their functioning (including risks and unintended consequences), but also in terms of the ways in which they help to shape new practices and experiences – including technological persuasion, but not limited to it. Explicit reflection on the possible mediating roles of a technology-in-design should, therefore, be part of the moral responsibility of designers. Moreover, the organizational context in which they function should leave enough room for such reflection and responsibility (cf. Coeckelbergh 2006).

### **2.3 An expanded framework for evaluating the ethics of persuasive technologies**

This elaboration of the connections between persuasive technology and technological mediation makes it possible to expand Berdichewsky&Neuenschwander's framework, and to modify it slightly. It has become clear that persuasive technologies do not only

*persuade* people into specific behavior or attitude-changes, but also *mediate* their behavior in multiple ways. This implies that not only the outcomes of *persuasion* are relevant here, but the outcomes of *all mediations* that arise in the use of the technology. Rather than linking the persuasive technology, via persuasive methods, with a persuaded person, therefore I would like to link it, via mediation, with *behavior*.

As argued above, human behavior results from technological mediation in complex ways, which cannot entirely be reduced to the intentions of designers, but which nevertheless need explicit moral reflection of designers. Three points of applications for moral reflection can be identified: the *intended persuasions* as built into the technology; the *forms of mediation* that occur on the basis of this, including the employed method of persuasion; and the *outcomes* of the technological mediations, including both the intended persuasion effects and the concomitant mediation effects. Figure 4 illustrates this adaptation of Berdichewsky&Neuenschwander's framework.



This figure shows that moral reflection in the design of persuasive technologies should primarily be directed at:

- *the intended persuasions of the technology-in-design:*  
can these persuasions be morally justified?

- *the methods of persuasion used, and the emerging forms of mediation:*  
are these methods morally acceptable, and are the implicit forms of mediation morally acceptable?
- *the outcomes of the mediation:*  
can the consequences of the persuasive and otherwise mediating role of the technology be morally justified?

The next section will take this inventory of points of application for moral reflection as a starting point, and elaborate a method of moral decision-making which can inform moral reflection about these aspects of persuasive technologies.

### 3. Methods of moral reflection in the design process

The theory of technological mediation reveals an inherent moral dimension in technology design. It shows that technologies always help to shape human actions and interpretations on the basis of which (moral) decisions are made. This has important implications for our understanding of the ethical roles of both technological artifacts and their designers. If ethics is about the question of how to act, this not only means that technologies help us to do ethics, and hence may even be said to embody some form of moral agency, but also that designing is ‘ethics by other means’. By helping to shape artifacts that will inevitably mediate human action and experience, designing comes down to materializing morality. Every technological artifact that is used will mediate human actions, and every act of design therefore helps to constitute specific human practices.

This implies that, in order to design persuasive technologies in a morally responsible way, designers should perform a moral assessment of the nature, method, and consequences of the persuasions and mediations they are designing into a technology. Performing a *mediation analysis* along the lines set out in section 2 is a first step to do this. The most important ‘instrument’ to be used for doing this, however trivial it may sound, is the designer’s *moral imagination*. By trying to anticipate in a structured way the future roles of the technology-in-design in human practices and experience, designers can establish a connection between the context of design and the future context of use.

There are two ways to take mediation analysis into the ethics of technology design. First of all, it can be used to develop moral *assessments* of technologies in terms of their mediating roles in human practices and experiences. In this approach, the main outcome is a ‘yes’ or a ‘no’ with regard to the question if this technology is morally acceptable or not. A second way to augment the ethics of technology with the approach of technological mediation is to not only assess mediations, but to also try to help *shape* them. Rather than working from an external standpoint *vis-à-vis* technology, aiming at rejecting or accepting new technologies, the ethics of technology then aims to *accompany* technological developments. This explicit designing of technologies in view of their mediations is, of course, an activity which is daily business for designers of persuasive technologies.

In this section, I will elaborate and augment a method for incorporating technological mediation in moral reflection on persuasive technology. This method is often used in applied ethics: the method of stakeholder analysis, which I will expand to also cover technological mediation.

### 3.1 Stakeholder Analysis

The first way to take mediation into ethics is closest to common practice in applied ethics: performing a stakeholder analysis and making an inventory of all relevant moral arguments in favor of and against a specific design. The gain with regard to both the Berdichewsky&Neuenschwander model and mainstream stakeholder analysis is to be found in the mediation analysis that supports it. Rather than following a purely instrumental approach to technology, or focusing on the motivations of designers, the persuasive methods employed, and the outcomes of the persuasion, this method focuses on *all* mediation effects. It therefore also includes effects that cannot be reduced to persuasion or to the (intended or unintended) outcomes of persuasion.

Starting point is always an inventory of all stakeholders involved: all persons who are likely to use the technology or to be somehow affected by other people using it. Moral reflection about the mediating roles of persuasive technologies can then take place along deontological or utilitarian lines.<sup>1</sup> Deontological arguments start from moral principles like beneficence, preventing harm, respect for autonomy, and justice. A deontological approach to the ethics of persuasive technology will therefore investigate to what extent the intended persuasions, the forms of mediation used, and the outcomes of the mediations are in accordance with such moral principles (see figure 4). Utilitarian arguments aim at balancing positive and negative outcomes of the specific design against each other. A

---

<sup>1</sup> There is, however, another approach to the ethics of technology, which is less practically oriented and which will therefore not be discussed in detail here. This approach can be called life-ethical or virtue-ethical, in contrast to the action-ethical approach discussed here. It focuses not so much on the *actions* resulting from specific design but on the quality of the *practices* that are introduced by mediating technologies, and their implications for the kind of life we are living. In such a life-ethical approach, it is not primarily the impact of mediation on the behavior of human beings which is important, but mainly the ways in which mediating technologies help to shape specific ways of 'dealing with things' or 'taking up with reality'. An example, which is a bit outside the realm of persuasive technologies, can illustrate this: the role of obstetrical ultrasound in antenatal diagnostics. Because of its ability to provide people with an indication of the possibility that an unborn child will suffer from a serious disease, this technology does not only influence behavior, it also installs a practice of dealing with unborn life which can be assessed on its own terms. Moral discussions around this technology, therefore, do not only focus on the safety of routine ultrasound scans and on the stress it causes to expecting parents, but also on the way of dealing with unborn life that comes along with it. A life-ethical approach here – which does not necessarily need to be associated with the pro-life movement – tries to assess the quality of the practices that arise around ultrasound scanning, in which the fetus and its expecting parents are constituted in specific ways, viz. as possible patients and decision-makers about the life of their unborn children.

utilitarian approach will make an inventory of the positive and negative consequences of the intended persuasions, the forms of mediation used, and the outcomes of the mediation. These consequences are then assessed in terms of their contribution to an intrinsic good like happiness (hedonistic utilitarianism), to a combination of intrinsic goods (pluralist utilitarianism), or to the majority of the preferences of all agents involved. A good stakeholder analysis aims to address all relevant arguments and develop a substantiated appraisal.

### 3.2 Assessing persuasive technologies

For making a moral assessment of a persuasive technology, this stakeholder analysis should be done for all three points of application identified in the first section: the intended persuasion; the form of mediation (including the method of persuasion); and the outcomes of the mediation. For the last point of application, also a mediation analysis is needed. In each of these points of application, typical issues emerge:

1. *the intended persuasions of the technology-in-design:*  
From a utilitarian point of view, we need to balance here the desirability of the intended behavior and its cost for all people involved. From a deontological point of view, the intended persuasions need to be in accordance with moral principles. Most relevant here seem to be:
  - *no harm:* does the intended persuasion cause no harm to people using it or those affected by it being used?
  - *beneficence:* does the intended persuasion benefit people using the technology or those affected by it being used?
  - *justice:* is the intended persuasion fair, treating people in the same circumstances in the same ways?
  
2. *the methods of persuasion used, and the emerging forms of mediation:*  
From a utilitarian point of view, an analysis has to be made of the benefit of persuading people into a desirable form of behavior as compared to the cost of specific methods of persuasion and other forms of mediation. From a deontological point of view, the following moral principles have particular relevance:
  - *respect for autonomy:* do people know they are being persuaded?
  - *no harm:* is people's privacy respected?
  - *justice:* is the technology free from a bias toward specific social groups?
  
3. *the outcomes of the mediation:*  
Here, mediation analysis comes into play. With the help of moral imagination, an inventory has to be made of all possible mediating roles of the technology in both human experiences and human actions. After this, these mediations need to be assessed morally, both in utilitarian and in deontological terms.  
For making a moral analysis of the design of a product like the FoodPhone, for instance, a creative inventory has to be made of the possible mediating roles of the

phone in human actions and experiences. This implies that one has to imagine the product as functioning in as many realistic use contexts as possible, focusing explicitly on its role in the practices and experiences of its users and other stakeholders involved. As for the domain of *action*, the FoodPhone, for instance, requires its users to make pictures of all food they are eating. This might complicate social interaction. Taking of pictures of everything you eat explicitly reveals to other people that you are working on your diet – something you may not want to share with everybody (utilitarian). And, more importantly, it also reveals that you are judging a social meal prepared to you by someone else mainly in terms of its nutritional aspects, which may be felt as inappropriate (deontological). These mediation effects may in fact impede people to use the FoodPhone at all.

As for the domain of *experience*, the FoodPhone will make its users more aware of what they are actually eating, probably resulting in losing weight and having a more balanced diet (utilitarian). But receiving permanent feedback on what you are eating may also result in an unhealthy obsession with food, and in a form of stress which is unhealthy in itself (utilitarian). Moreover, the FoodPhone invites its users to be ‘observers’ of their eating behavior, thus detaching them from their immediate involvements in their environment (utilitarian).

In order to assess whether the FoodPhone is a morally acceptable technology, all these arguments need to be taken into account and played out against each other. The main question here will be whether the probability of a positive effect on people’s eating habits outweighs the possible negative effects on people’s social life and on their attitude toward eating in general.

#### **4. Social Issues**

Having identified points of application for moral reflection in the design of persuasive technologies, as well as two methods for actually putting this moral reflection into practice, it is good to separate out some specific moral issues that are a bit more external to the design process itself. These issues do not arise primarily from the perspective of designers, but from that of users of persuasive technologies. They include *trust* (to what extent can consumers trust the creators of persuasive technologies?); *responsibility* (who can be held responsible for the resulting behavior of users?); *reliability* (can we be sure the persuasion does not have undesirable effects?); and the *desirability* or even *legitimacy* of technological persuasion at all. I will take the first three issues together in section 4.1, since they all revolve around the same characteristic of technological persuasion: the difficulty to predict its outcomes. The issues of desirability and legitimacy are of a different nature, and will be discussed in section 4.2, where I will pay special attention to possible threats to human freedom and to the democratic principles organizing our society.

#### **4.1 Trust, Reliability, and the Distribution of Responsibility**

When introducing technologies that explicitly aim to influence human behavior on a large scale, such technologies need to meet specific requirements in order to be morally acceptable. One of the most important requirements is that users can *trust* the technology they are using. Trust in this context means that people can reasonably expect the technology to do what it is supposed to do and that the consequences of using the technology will not be harmful to them or are otherwise undesirable, unless they are adequately informed about this. This implies that trust implies both *reliability* of the persuasive technology and *responsibility* on the part of the designers.

The degree to which both of these aspects of trust are actually realized depends on the degree to which the eventual consequences of using a technology can be linked to the activities and intentions of the designers and users. After all, this is the only way to predict the future impact of the technology. If we cannot link the impact of an EconoMeter in a car to what its designers intended it to accomplish and to how its users employ it, it is hard to call it ‘reliable’ and to find comfort in the idea that it was designed in a responsible way. But this predictability of impact is a serious problem. As I elaborated in section 2, the theory of mediation shows that causal responsibility for technological mediation needs to be distributed among designers, users, and the technologies themselves, and that there will always emerge unforeseen mediations, like energy-saving light bulbs actually causing an increase of energy use, and cell phones changing patterns of social interaction.

Yet, the absence of full predictability about the mediating roles of technology does not imply that trust, reliability, and responsibility are fully incompatible with persuasive technology. After all, the identification of several *causal* responsibilities for mediation forms a good basis for attributing *moral* responsibility to those agents who are capable to take this responsibility, i.e. designers and users of persuasive technologies. Designers need to anticipate the mediation effects of their designs as much as they can, by performing mediation analyses with the help of their moral imagination and using such analyses in moral decision-making processes. Users, in their turn, need to anticipate technological mediations as well, to the extent to which such mediations result from their specific appropriations and interpretations of the technology.

If both users and designers act in a morally responsible way, by not simply designing and using technologies as mere *instruments* but by using their moral imagination to approach them as *mediators*, there is reason enough to trust that technologies will actually do what they were designed for, and that as little unacceptable consequences as possible will result from using them. At the same time, we can only speak of *trust* here – not of certainty, because technologies are inevitably surrounded by uncertainties and risks, dependent as they are from the whimsicality of the relations they will develop with human beings.

#### **4.2 Legitimacy**

An entirely different set of questions pertains to the phenomenon of technological persuasion or behavior-influencing technology as such. Two main arguments can be

identified here. The first is, than human *freedom* might be threatened when human actions are explicitly and consciously steered with the help of technology. This reduction of human freedom can even be perceived as a threat to human dignity; if human actions are not a result from deliberate decisions but from steering technologies, people are deprived from what makes them human. Secondly, a plea for developing behavior-influencing technologies can be seen as an implicit propagation of technocracy. When moral issues are solved by the technological activities of designers instead of the democratic activities of politicians, not humans but technology will ultimately be in control.<sup>2</sup>

Yet, these arguments are not really adequate. First of all, human dignity is not necessarily threatened when our freedom is deliberately limited. The law, after all, also forms a major limitation of our freedom, after all, but can hardly be seen as a threat to human dignity – to the contrary, in most cases, I would say. Human behavior is determined in many ways, and human freedom is limited in equally many ways. Few people will protest against the legal prohibition of murder, so why protest to the material inhibition imposed by a speed bump to drive too fast at places where children are often playing on the pavement?

Secondly, the analysis of technological mediation in section 1 made clear that technologies *always* help to shape human actions. Because of the pervasive role of technologies in our daily lives, it is hard to find an activity or experience which is *not* technologically mediated. Rather than holding on to an interpretation of human freedom in terms of absolute autonomy and sovereignty from technology, therefore, it seems wise to reinterpret freedom as a person's *ability to relate to what determines and influences him or her*.

Human actions always take place in a stubborn reality, and for this reason, absolute freedom can only be attained by ignoring reality, and therefore by giving up the possibility to act at all. Freedom is not a lack of forces and constraints; it rather is the existential space human beings have to realize their existence. Humans have a relation to their own existence and to the ways in which this is co-shaped by the material culture in which it takes place. The material situatedness of human existence *creates* specific forms of freedom, rather than impeding them. Freedom consists in the possibilities that are opened up for human beings to have a relation to the environment in which they live and to which they are bound. Rather than rejecting the idea of behavior-influencing technology, therefore, we had better try and give this influence a desirable form.

These counterarguments against the fear of a loss of freedom, however, do not take away the anxiety that a technocracy would come about when technologies are explicitly moralized. It might be true that technologies do not differ from laws in limiting human freedom, but laws come about in a democratic way, and the moralization of technology does not. Yet, this does not justify the conclusion that it is better to refrain from paying explicit attention to technological mediation during the design process. If technologies

---

<sup>2</sup> Arguments like these have actually been brought into play against a proposal of the Dutch philosopher Hans Achterhuis to stop moralizing each other, and to start delegating morality to technology. See Achterhuis 1995; 1998, 28-31.

are not ‘moralized’ explicitly, after all, responsibility for technological mediation is left to the designers only. And precisely this would amount to form of technocracy. A better conclusion would be that it is important to find democratic ways to design persuasive (and otherwise mediating) technologies. The methodology of ‘constructive technology assessment’ could contribute to this (see Verbeek 2006a), because it organizes a domination-free discussion with all stakeholders to anticipate and assess the impact of technologies-in-design and to feed the outcomes back into the design process.

## **5. Conclusion**

The ethics of persuasive technology needs to approach technological persuasion as a specific form of technological *mediation*. This makes it possible to include also unintended influences and outcomes of the technology in moral decision-making and in the responsibility of designers. Moral reflection in the designing process then has three points of application: the intended persuasion, the form of mediation (including the method of persuasion), and the outcomes of the mediation. An augmented version of the ethical method of stakeholder analysis, which includes a mediation analysis carried out with the help of the moral imagination of the designer, appears to be an adequate way to guide moral reflection. Moreover, an augmented version of Constructive Technology Assessment can shift the process of moral reflection from the individual designer to all relevant stakeholders. Apart from this, there are more external moral questions regarding persuasive technology which concern their reliability, trust in their workings and their designers, the attribution of responsibility in case anything goes wrong, and the legitimacy and desirability of persuasive technology at all because of the alleged threat it poses to human freedom and to democracy. The first three issues can be met by responsible decision-making of both users and designers, taking up the mediation effects of the technology-in-design. The last issue can be met by developing an understanding of human freedom which includes the technologically mediated character of our actions and experiences.

## References

- Achterhuis, H. (1995), 'De moralisering van de apparaten'. In: *Socialisme en Democratie* 52 no. 1, 3-12
- Achterhuis, H. (1998), *De erfenis van de utopie*. Amsterdam: Ambo
- Akrich, M. (1992), 'The De-scription of Technical Objects'. In: W.E. Bijker and J. Law, *Shaping Technology / Building Society*. Cambridge: MIT Press, 205-224
- Berdichewsky, D. and E. Neuenschwander (1999), 'Toward an ethics of persuasive technology', *Communications of the ACM*, v.42 n.5, p.51-58, May 1999
- Coeckelberg, M. (2006) Regulation or Responsibility? Autonomy, Moral Imagination, and Engineering. In: *Science, Technology & Human Values* Vol. 31 no. 3 (May 2006), pp. 237-260
- Ihde, D. (1990), *Technology and the Lifeworld*. Bloomington/Minneapolis: Indiana University Press
- Latour, B. (1992), 'Where are the Missing Masses? - The Sociology of a Few Mundane Artifacts'. In: W.E. Bijker and J. Law, *Shaping Technology / Building Society*. Cambridge: MIT Press
- Latour, B. (1994), 'On Technical Mediation -- Philosophy, Sociology, Genealogy'. In: *Common Knowledge* 3, 29-64
- Smith, A. (2003) 'Do You Believe in Ethics? Latour and Ihde in the Trenches of the Sciences Wars'. In: Don Ihde and Evan Selinger, eds. 2003. *Chasing Technoscience: Matrix for Materiality*. Bloomington & Indianapolis: Indiana University Press
- Steg, L. (1999), *Verspilde energie? Wat doen en laten Nederlanders voor het milieu*. The Hague: Sociaal en Cultureel Planbureau (SCP Cahier no. 156)
- Verbeek, P.P. (2004a), 'Stimuleer gedragsbeïnvloedende technologie'. In: *CDV*, Herfst 2004 ('Tussen Apocalyps en Utopie'), pp. 117-124, ISSN 0167-9155
- Verbeek, P.P. (2004b), 'Material Morality'. In: Ed van Hinte (ed.), *Time in Design*. Rotterdam: 010 Publishers, pp. 198-210
- Verbeek, P.P. (2005), *What Things Do – Philosophical Reflections on Technology, Agency, and Design*. Penn State: Penn State University Press
- Verbeek, P.P. (2006a), 'Materializing Morality – design ethics and technological mediation', in: *Science, Technology and Human Values - special issue on Ethics and Engineering Design*, Vol. 31 no. 3 (May 2006), pp. 361-380
- Verbeek, P.P. (2006b - forthcoming), 'Acting Artifacts - the Technological Mediation of Action', in: P.P. Verbeek and A. Slob (eds.), *User Behavior and Technology Development – Shaping Sustainable Relations between Consumers and Technologies*. Dordrecht: Kluwer
- Weegink, R.J. (1996). *Basisonderzoek elektriciteitsverbruik kleinverbruikers BEK'95*. Arnhem: EnergieNed