University of Zurichᵁᶻᴴ

Towards an Effective and Ethical Design of Persuasive Systems for Sustainability

BACHELOR'S THESIS

AUTHOR

PATRICK LOOSER

18-728-675

SUPERVISOR

PROF. DR. LORENZ HILTY

INSTITUTE FOR INFORMATICS

UNIVERSITY OF ZURICH

DATE OF SUBMISSION: 21.07.2021

Contents

List of Figures

List of Tables

List of Abbreviations

CO$_2$ Carbon Dioxide

PD   Participatory Design

PSS   Persuasive Systems for Sustainability

PT   Persuasive Technology

UTAUT   Unified Theory of Acceptance and Use of Technology

VSD   Value Sensitive Design

Abstract

The increasing threat of environmental problems urges for solutions. At the same time, several ethical concerns about the misuse of technology, such as privacy violations or manipulation, have emerged. Even though there is a substantial amount of research regarding persuasive technology (PT), this work aims to find potential use cases of PT to address environmental problems. This work states that persuasive systems for sustainability (PSS) can address environmental problems in an ethical fashion by revealing the many unused potentials of ethical design for PSS. The review of theoretical and practical literature results in a broad set of ethical design requirements for PSS design. instead of only providing further constraints to the already quite difficult area of PSS, this works also reveals the unused potential of PSS by applying ethical requirements to well-known practical PT design principles and strategies.

Zusammenfassung

Die zunehmende Bedrohung durch Umweltprobleme drängt nach Lösungen. Gleichzeitig sind etliche ethische Bedenken über den Missbrauch von Technologie, wie z. B. Verletzungen der Privatsphäre oder Manipulation entstanden. Obwohl es eine beträchtliche Menge an Forschung in Bezug auf persuasive Technologie (PT) gibt, zielt diese Arbeit darauf ab, potenzielle Anwendungsfälle von PT zur Lösung von Umweltproblemen zu finden. Diese Arbeit argumentiert, dass nachhaltigkeitsorientierte persuasive Systeme (PSS) Umweltprobleme auf ethische Art und Weise angehen können, indem sie die vielen ungenutzten Potenziale von ethischem PSS-Design aufzeigen. Die Überprüfung sowohl der theoretischen als auch der praktischen Literatur führt zu einigen ethischen Designanforderungen für die Entwicklung von PSS. Anstatt nur weitere Einschränkungen für den bereits sehr schwierigen Bereich von PSS Design zu liefern, zeigt diese Arbeit auch das ungenutzte Potenzial von PSS, indem ethische Anforderungen auf bekannte, praxisnahe PT-Designprinzipien und -Strategien angewendet werden.

1 Introduction

Persuasive technology (PT) can help to raise the awareness of future or distant issues by pointing out cause-and-effect relationships [59]. Especially because information systems are increasingly ubiquitous and tend to overwhelm us with information, there is a growing importance of having the relevant information present at the right time. The countless possibilities of modern technology allow for different methods of PT design to change an individual's attitude or behavior. In contrast to a human persuader, computers are tirelessly persistent and influence us without being affected themselves [53].

There are many types of guidelines focusing on different aspects of PT. For instance, the framework proposed by Oinas-Kukkonen and Harjumaa [66] focuses on system content and functionality. Persuasive systems for sustainability (PSS) face several unique challenges to achieve effective persuasion. For example, users have little incentive to change their habits based on waste considerations when they don't have enough knowledge of the ecological and social impacts resulting from their consumption [17]. To achieve sustainable development, the change of consumer behavior has been identified as a field of research [41].

Ethical concerns regarding PT have been identified as an important area of future research [86]. There is some research that considers ethical aspects of PSS (e.g., [8], [28] and [93]), but one problem of the existing guidelines is their expectation of PT-designers to anticipate future consequences of persuasion [98]. Even though there are ethical frameworks about PT, there seems to be less research about guidelines that combine ethical considerations with findings of areas focusing on system efficiency and effectiveness – especially in the field of PSS, i.e., there is a lack of knowledge of practically applicable, useful and ethical design guidelines. Based on that, the following research questions can be formulated:

Q1: What generic requirements should be considered by designers of ethical PSS?

Q2: How can these generic requirements be applied to well-known practical design principles and strategies?

Q3: What opportunities and risks come with the design of ethical PSS?

A general discussion about the legitimacy of PSS in general will be part of the Discussion Section. The focus will be on the following research question:

Q4: How should we value sustainability in contrast to an individual's freedom and welfare?

Even though it is not feasible to create an all-encompassing framework for such a large field of research, this work aims at contributing towards guiding the discussion of ethical PSS in the desired direction by addressing practical, user-centered and ethical considerations for PSS.

Designers require a great understanding of human capabilities when they want to design an information system that provides effective and efficient use. If we further want the system to change the user's behavior or attitude towards a certain target behavior, things can get more complicated. For instance, we might want to include attitudinal theories from social psychology [66].

In this work, the discussion revolves around two highly debated topics: the role of technology for persuasion and persuasion in general. Both of these areas offer a wide range of possible good and bad practices. This work reconsiders existing ethical observations with a focus on PSS. Instead of discussing the general problems of persuasion and technology, the upcoming argumentation focuses on problems that are relevant for common practices in the area of PSS. To narrow down the problems, this work takes upon existing literature on PT and finds problems that are of special relevance concerning sustainability-related persuasion.

To create guidelines for PSS, basic research in the area of PT and information systems in general will be conducted. Insights from ethical and practical design frameworks for PT are considered as well. Based on the findings from the existing literature, guidelines are presented according to ethical justifications by putting PSS in an ethical and user-centered context.

Oinas-Kukkonen and Harjumaa [66] have provided a detailed framework for PT design (see Section 3.1.1). However, the design principles mentioned in their work are very generic and theoretical but often refer to specific practices that are already applied in PT. This work aims to further support some of these design principles through practical reasoning and by finding validating use cases.

In a first step, literature will be reviewed that serves as a foundation for the elaboration of this work and further supports the relevancy of the chosen approach. For this, a widely used framework in PT research is introduced. After that, the relation between the framework and the UTAUT 2.0 Model from Davis [18] is shown, which is also a well-known theory among

designers of information systems in general. This creates a link between persuasive design principles and relevant factors for technology acceptance. The literature review also includes some fundamentals about the psychology of behavior change that is relevant for PSS design.

Section 4 contains important considerations relevant for the design of ethical PSS. Based on those findings, a set of generic requirements for ethical PSS design is established. To get a clearer understanding of the chances and risks of those generic requirements, they will be applied to a set of widely used design principles (see Section 5). The urgency of behavior change and the role of PSS is then discussed in Section 7.

2 Terminology

2.1 Persuasive Technology (PT)

Persuasive technology (PT) is interactive information technology designed to change users' attitudes or behavior [29]. According to the definition of Ijsselsteijn et al. [38], PT does not include persuasion by force or misinformation. However, this work will still make efforts in arguing why coercion and deception should not be part of PT. When talking about PT, we can aim for three potential outcomes, namely reinforcement, change or shaping of attitudes and/or behaviors [67].

2.2 Persuasive System for Sustainability (PSS)

There are several terms used to describe PT that intends to persuade its users towards a more sustainable behavior or action. For example, Brauer and Ebermann [13] use the term "persuasive environmental sustainable systems" (PESS). The problem with this term is the ambiguity of whether sustainability refers to the system itself (e.g., being energy-efficient) or to the target behavior. Nyström and Mustaquim [64] use the term that will be used throughout this work called "persuasive system for sustainability" (PSS), which is in fact very common and a lot more precise because it indicates that the persuasive system is meant to *promote* sustainability, not *be* sustainable.

2.3 Ethical Dilemma

Kidder [51] defines the term "ethical dilemma" as right-versus-right situations where two core moral values conflict with each other. In the context of PSS, an ethical dilemma is to find an appropriate balance between individual and collective goals, i.e., asking how much we should put back the individual's welfare in favor of collective goals such as sustainability.

2.4   User-Centered Design

The guidelines proposed in this work include the consideration of effectiveness and efficiency, as well as ethical considerations regarding the design of PSS. Designing such a system can lead to certain conflicts due to the nature of PSS and sustainability in general. One of the main problems revolves around the prisoner's dilemma which explains why rational individuals might not cooperate to achieve the holistically best outcome. This puts the designers of PSS in a difficult spot because they have to decide whether to focus on the user's values and needs or the worldwide necessity to promote individual sustainable behavior. The design guidelines proposed in this work aim to account for – amongst other factors – the preferences and values of the user. In this paper, User-Centered Design includes not only the preferences but also the values of the user.

3 Literature Review

3.1 Acceptance of PSS

User acceptance is one of the fundamental topics that should not be neglected when designing systems for practical use. No matter how good the system is, its potential won't come into effect if it is not used. Because this work aims to create implications for practical use, user acceptance considerations should always play a role when reasoning about the applicability of certain design principles. By pointing out the interrelationships between Oinas-Kukkonen and Harjumaa's [66] systematic framework and the user acceptance theory proposed by Venkatesh [92], design potentials for PSS are revealed.

3.1.1 Conceptual Design of PT

Oinas-Kukkonen and Harjumaa [66] suggest a classification of design principles for PT that represent the state of the art in persuasive design research [86]. Their framework is based on the work of Fogg [29] and contains the following four categories:

- Primary Task Support

- Dialog Support

- System Credibility

- Social Support

Each of the four categories contains a set of design principles for persuasive system design.

Primary Task Support. The first category, primary task support, contains design principles that enable users to reach their goals [13]. The main focus of this category lies in providing meaningful content for the user in contrast to only supporting a process [66]. Simulation will also be referred to as a "cause-and-effect relationship" during this work. The primary task support category consists of the following design principles:

1. Reduction
2. Tunneling
3. Tailoring
4. Personalization

5. Self-Monitoring
6. Simulation
7. Rehearsal

Dialogue Support. Dialogue support design principles describe the support of computer-human dialogue to help users to achieve their target behavior [66]. The dialogue support category consists of the following design principles:

1. Praise
2. Rewards
3. Reminders
4. Suggestion

5. Similarity
6. Liking
7. Social Role

System Credibility. The system credibility design principles address how a system should be designed so that it is more credible, which leads to better persuasiveness [66]. The system credibility category consists of the following design principles:

1. Trustworthiness
2. Expertise
3. Surface Credibility
4. Real-World Feel

5. Authority
6. Third-Party Endorsements
7. Verifiability

Social Support. Based on Fogg's [29] principles on mobility and connectivity, Oinas-Kukkonen and Harjumaa [66] proposed social support as the last category of design principles, which identifies ways to design the system to motivate users through social influence. The social support category consists of the following design principles:

1. Social Learning
2. Social Comparison
3. Normative Influence
4. Social Facilitation

5. Cooperation
6. Competition
7. Recognition

3.1.2   UTAUT 2.0

As a major step towards the development of effective PT, PSS design needs to make sure that the system is well received and accepted by the user. To address this, Davis [18] has
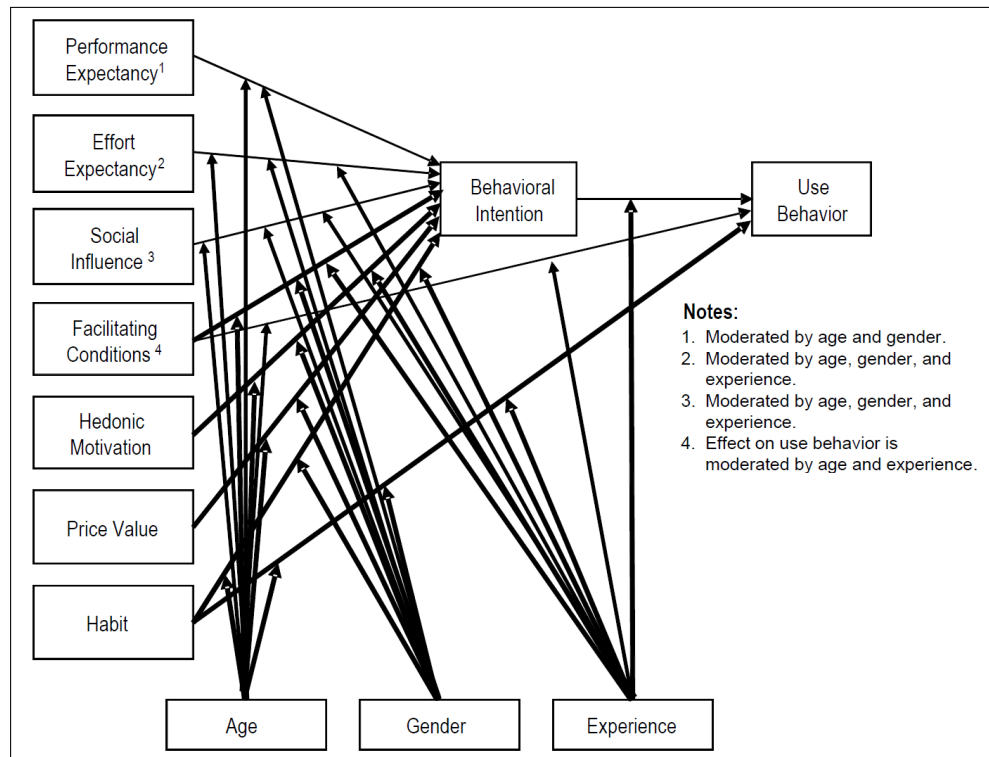
Figure 1 – UTAUT 2.0 from [92]. Age, gender and experience have lines pointing towards factors moderated by it. All other lines can be read as "... is predictor of ...".

proposed the Technology Acceptance Model (TAM) to predict the intention and actual use of the system. According to the TAM, the system should be (1) adequate to fulfill a certain task (perceived usefulness) and (2) have a high perceived ease of use to be accepted by the user [13, 18]. However, to increase the explainability of technology acceptance behavior, Venkatesh [92] soon extended the TAM into a new theory called "Unified Theory of Acceptance and use of Technology" (UTAUT), which was later extended into UTAUT 2.0, that has proven to help in understanding the influencing factors of acceptance for new technologies [13].

UTAUT 2.0 proposes seven factors that influence the user's behavioral intention to use technology (see Figure 1) [92]. Performance expectancy and effort expectancy refer to the perceived usefulness and perceived ease of use as we have seen in the TAM. Social influence occurs when users perceive that important others think that they should use a certain technology [92]. Facilitating conditions are the degree to which the consumers perceive the availability of resources and support to perform a behavior [91, 92]. All the factors mentioned until now were also included in the previous model, UTAUT (1.0). In UTAUT 2.0, hedonic motivation, price value and habit were added to the theory. Hedonic motivation is the fun or pleasure perceived by the user through the use of technology and price value refers to the difference

between the perceived benefit of a system and its monetary cost [92]. Habit is the degree to which users are already familiar with the system [92]. Age, gender and experience moderate the effects of certain factors [92].

3.1.3   Design for User Acceptance

By conducting surveys, Brauer et al. [13] showed that all categories mentioned in Section 3.1.1 positively impact certain factors leading to behavioral intention to use the technology as proposed by UTAUT 2.0 (see Figure 2). The primary task support category has a positive impact on performance expectancy, effort expectancy and facilitating conditions, dialogue support showed a positive impact on performance expectancy, facilitating conditions and hedonic motivation, system credibility has a positive impact on performance expectancy and social support positively impacts performance expectancy, social influence and hedonic motivation [13]. This means, that features supporting the persuasiveness of a system indirectly support the users' acceptance of the system. This knowledge lets us focus on the persuasive effect of the system without constantly being concerned about whether the systems will be accepted by the users. Of course, any additional information about the actual use of the system through practical evaluation will still be valuable.
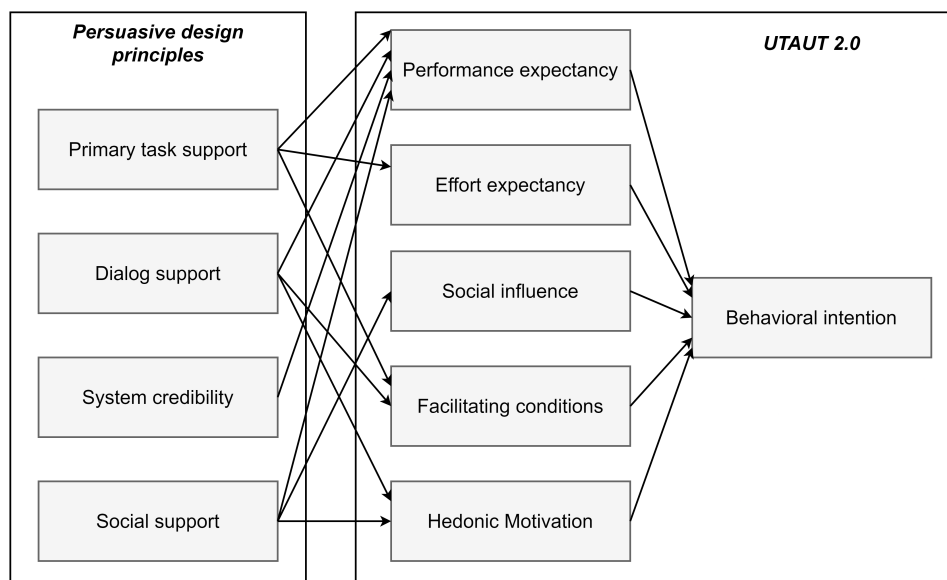

Figure 2 – Simplified Structural Model based on the findings of [13]. The arrows show which factor positively impacts another.

3.2   Cognitive Dissonance

Cognitive dissonance occurs when a person's thought or behavior is not consistent with another thought, which is felt as a pressure to get rid of the inconsistency [26]. The person experiencing cognitive dissonance has several options to get rid of it [26]. Consider a person who claims to be aware of environmental problems caused by CO_2 emissions that wants to live an eco-friendly life. He knows that airplanes cause a lot of CO_2 emissions. At the same time, he knows that he is flying very often and already plans his next flight. The fact that he wants to fly often while trying to live an eco-friendly life causes an inconsistency. This inconsistency can be felt as a form of stress for the person, which is referred to as cognitive dissonance [26]. He will do everything within his power to get rid of this inconsistency [26]. There are four methods to get rid of cognitive dissonance: (1) justifying the behavior by changing one of the conflicting parts, (2) justifying the behavior by adding another behavior, (3) denying one of the conflicting parts or by (4) changing the behavior [26]. A PSS-specific example is provided for each method in Table 1.

	Method	Example
1	Justifying the behavior by changing one of the conflicting parts	"Flying four times a year is still eco-friendly."
2	Justifying the behavior by adding another behavior	"If I take my bike to work from now on, I can still fly four times a year."
3	Denying one of the conflicting parts	"Flying four times a week is not causing environmental problems."
4	Changing the behavior	"I'll fly less often."

Table 1 – Four methods to get rid of cognitive dissonance as proposed by Festinger [26].

This theory gives us a much clearer picture of how PSS can make use of cognitive dissonance to achieve behavior change. If PSS wants to make use of cognitive dissonance to persuade its users into a sustainable target behavior, designers need to make sure that they are aware of (1) behavioral goals that lead to more sustainability and (2) make them aware of the conflicts between the users' current behavior and those behavioral goals. By explicating this inconsistency, PSS can motivate the users to reduce their discomfort, preferably through method 4 (see Table 1). This can be achieved by increasing the chances that the users choose to change their behavior instead of changing or denying conflicting parts, which can be done

directly, by making the behavior change more likely, or indirectly by decreasing the chance of the users to opt into any of the other three methods.

3.3   Motivation, Ability and Triggers
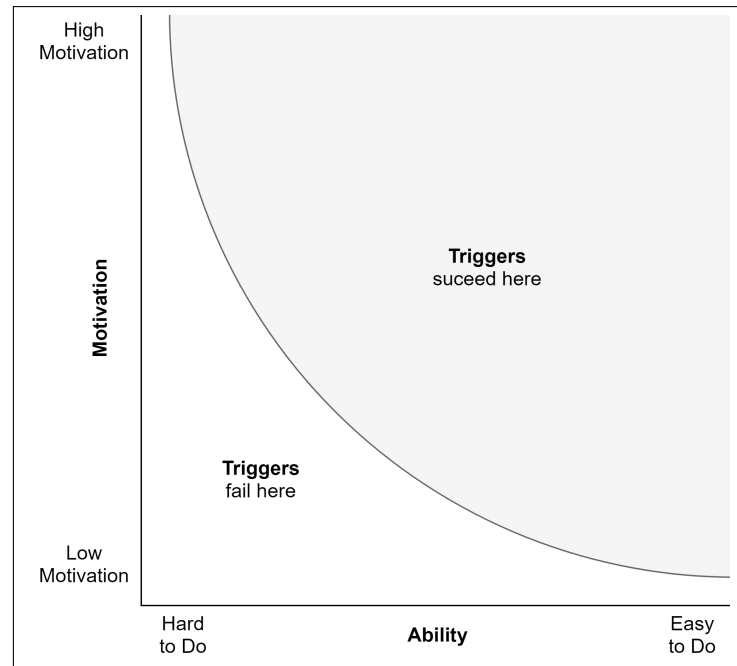
Figure 3 – Model based on the work of Fogg [30]

To increase the chances of the users to opt for method 4 of the Cognitive Dissonance Model as a way to get rid of his cognitive dissonance, Fogg's [30] Behavior Model can be applied. This model is fundamental yet very important to understand how PSS can cause a change in the users' behavior. According to Fogg's [30] Behavior Model, the persuadees need to be (1) able to perform the behavior, (2) have sufficient motivation for the behavior and (3) require a trigger that sparks the action (see Figure 3). An example of such a trigger could be concrete suggestions for action [97] (e.g., through just-in-time prompts shown in Section 5.2).

4 Ethical Considerations

Ethical violations in technology design are not uncommon and further research is necessary for ethics to catch up with the current practices of PT [52]. Even though some unethical practices are not explicitly used for a persuasive intent, the design of technology always has some persuasive implications [52, 75]. This section aims to address important considerations that help designers to acknowledge the responsibility that comes with designing PSS and provide a clearer picture of several ethical challenges and opportunities. The resulting generic requirements for ethical PSS design will be displayed in red boxes after each section.

4.1 Responsibility and Backfires

4.1.1 Intention-Outcome Matrix

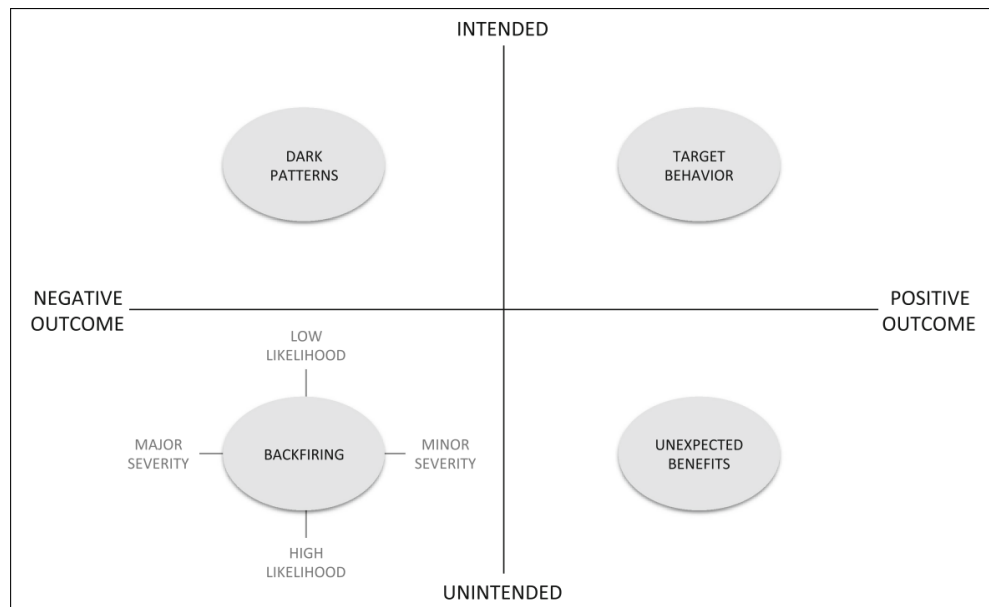**Figure 4** – Intention-Outcome Matrix [84]

When using PT, we can classify behavior based on two factors, namely the intention and the outcome, which is shown by Stibe and Cugelman [84] in the Intention-Outcome Matrix (see Figure 4). If the outcome is intended and positive, we get the "target behavior", which is the primary intended positive behavioral outcome [84]. "Unexpected benefits" happen if a

positive outcome occurred unintentional [84]. Negative intended outcomes[1] are called "dark patterns", which can be considered unethical practices, such as coercion, deception and fraud [84].[2] "backfires" occur if a negative outcome was not intended and can themselves be classified based on severity and likelihood [84].

4.1.2   Decision Tree
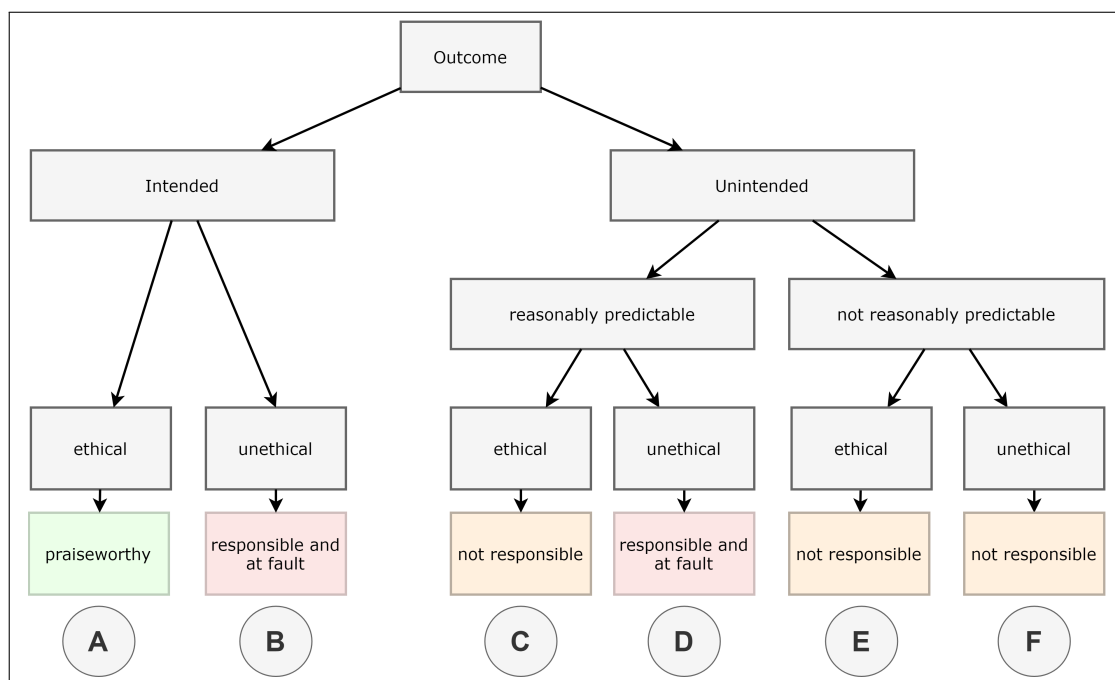
Figure 5 – Decision tree to classify outcomes according to their intention and predictability (based on [8]).

According to Berdichevsky and Neuenschwander [8], designers of PT can or cannot be held responsible based on certain outcomes. Like for many other technologies such as self-driving cars, the question about responsibility in case of a mishap arises. On one hacd, the designer of PSS needs to take certain precautions when deploying his system, but on the other hand, the user has to be aware of the risks and shortcomings of using the system. After all, not the computer, but its designers are responsible for the means and outcomes of persuasion [53].

---

[1]   Stibe and Cugelman [84] talk about *negative outcomes*, while Berdichevsky and Neuenschwander [8] use the term *unethical*. Even though both terms are somewhat vague, the term *negative outcome* could also be interpreted as a non-ethical negative outcome, such as a profit decline for the designer. However, since this work focuses on user-centered, ethical design and sustainability-oriented design guidelines, the focus of my argumentation will be on unethical outcomes.

[2]   These practices will make up parts of the discussion about paternalism (see Section 4.3)

To decide upon the responsibility of the designer for different outcomes, Berdichevsky and Neuenschwander [8] have created a decision tree that helps to classify the responsibility based on intention and predictability of the outcome (see Figure 5). This decision tree can support us in drawing special attention to design decisions that might lead to outcomes that fall in the scope of responsibility for the designer. In the first step of designing the system, the designer has to take precautions in clearing out any design choices resulting in intended, unethical outcomes (Figure 5, B) or reasonably predictable, unintended and unethical outcomes (Figure 5, D).

4.1.3   Backfires

The designers of PT have to consider several issues that might occur during PT design which are not trivial. Some scenarios result in the opposite of the target behavior or negative side effects [67]. If such outcomes occur unintentionally, they are called "backfires" [67]. Berdichevsky and Neuenschwander [8] would further separate these backfires according to their predictability, which corresponds to category D and F of Figure 5. This brings another factor to the initial concept of Stibe and Cugelman [84], which is responsibility. To identify concrete requirements for responsible PT design, persuasive backfires which can be credited towards the designer have to be identified and accounted for.

Some persuasive practices can result in backfiring, but would exceed the designer's capabilities to influence the occurrence in many cases (Figure 5, F). For these kinds of outcomes, the designer can not be held responsible [67]. An example of such a backfire is the so-called "self-licensing", which occurs when a user carries out a good action, which makes them feel like they have a license to misbehave at another time [84]. For example, it has been shown that energy-reducing behavior might rebound and lead to more usage after all due to self-licensing [53]. Even though PT offers ways to address this backfire, the designer can usually not be expected to account for it.

Additionally, designers should not be blamed if a negative outcome occurs due to the system being used in a way that it was not supposed to [90]. Because the focus of this work lies on the ethical considerations concerning the designers of PT, those kinds of unforeseen backfires will be excluded from my work. However, informing the user about how and when to use the system still lies in the responsibility of the designer. The upcoming sections explain why backfires fall into category D of the decision tree (Figure 5) and thus, should definitively be

accounted for by designers of ethical PSS.

Overemphasizing. Overemphasizing occurs when you motivate people to do an action for a strongly emphasized benefit while omitting harmful other factors [84]. This has a lot to do with transparency as discussed in Section 4.6. A user-centered system should generally be transparent about the advantages and disadvantages for the user. For instance, if a public transportation app tries to convince the user to use the train instead of the car, he should point out the sustainability-related advantages of using the train. But on the other hand, the app should also provide information about possible disadvantages, such as the occupancy of the train, possible delays and so on. The designer should be held responsible for bad experiences caused by persuading the user into behavior that results in a backfire due to the distorted perception of cost and benefit of the user caused by the PSS. Since we are talking about PSS, the main goal is to persuade the user to choose the more sustainable action, but the system should not deceive the user to reach that goal.

Mistailoring. One backfire that can be avoided to a large extent by the designer is mistailoring. This occurs when a message contains information that can result in a negative outcome for some users because it is not considering the persuadee's current situation [84]. An example of this might be an app that shows the user how sustainable he is compared to others. If he is less sustainable than the shown average, he might successfully get persuaded into a more sustainable action, but if the user is already way above the average, he might be encouraged to stop improving his sustainable behavior due to the previously mentioned self-licensing effect.[3]

Mistargeting and Misdiagnosing. When the message is sent to a different audience than intended, misinterpretations can occur, which can lead to a negative outcome [84]. Misdiagnosing occurs when user behavior or psychological processes are not considered in the behavior change intervention [84]. In the case of PSS, this might correspond to a public transportation app suggesting a seat in the kid's compartment of a train to a businessman without kids. He might want to work during the train ride and then an offer in the business compartment would surely bring more value for the user. If he chooses to take the offer and ends up having a bad experience, he might consider using a car the next time, thus backfiring both as a negative experience for the user and the loss of a public transportation user. The counterpart to a one-size-fits-all approach would be personalization. As we will see in Section 4.5.2, personalizing does not come without drawbacks.

---

[3]   In this case, the self-licensing effect is evoked by the system, which is the result of bad design and should be accounted for by the designer.

Anti-Modeling. Anti-modeling occurs when you demonstrate negative behavior, which exposes the user to memory triggers of temptation to perform the negative behavior [84]. For instance, if you show a picture of an overcrowded train in an app that is supposed to motivate you to use public transport, you might motivate them to use the car instead of the train.

Generic Requirement 1: Preventing Backfires

To persuade the user into a more sustainable behavior, the designer of PSS's should account for persuasive backfires that can and should be avoided. Some of the effects to be considered include:

- Overemphasizing

- Mistailoring

- Mistargeting and Misdiagnosing

- Anti-Modeling

4.2   Rational Persuasion

"Rational persuasion", as proposed by Tsai [88], refers to the practice of persuading users through reasons, evidence or arguments and is often also contrasted with coercion and manipulation [88]. In the area of persuasive technologies addressing sustainable urban mobility, the display of estimated CO_2 emissions per trip has shown to yield great persuasion results [12]. Cause-and-effect simulation is one approach to rational persuasion which addresses the user's locus of control. Locus of control is the perceived control that the user has over the outcome [27]. Keeping track of the user's mobility behavior and giving personalized feedback on their CO_2 emissions caused has been used as an effective means to address the user's locus of control and sense of efficacy [11, 44]. This approach makes use of the rational persuasion principle by providing the user with facts about his mobility behavior to convince the user towards a certain target behavior.

One problem that can be held against rational persuasion is the conflict between short-term self-interests and long-term collective interests. That means that the best behavior from a

collective perspective might not be experienced as best for the individual [87]. Regarding the collective concern of sustainability, an example of a desired individual target behavior would be sustainable mobility, such as taking the bike to work instead of the car. But this conflicts with a lot of individual concerns such as comfort or efficiency [87]. These conflicts occur not only in the domain of sustainability, but pose a general problem for society [87, 89]. It has been shown that self-interested orientation negatively impacts sustainable behavior [21, 33, 79]. This means that rational persuasion towards a collectively preferred target behavior will be difficult, as long as people prefer behavior based on their individual concerns. A possible solution to address these conflicts would be to bridge the two groups of concerns [87].

There are certain less ethical approaches towards achieving user persuasion towards a collectively preferred target behavior that is worth addressing. If we select target behaviors purely according to collective concerns without respecting the short-term self-interests of the users, we are already at risk of interfering with the user's liberty in a way they don't agree with. This is especially the case if the user is not informed about this prioritization of concerns. Further deliberation of the user might occur through the usage of any form of manipulation that leads the user into falsely thinking that the collectively beneficial target goal is also most beneficial for his individual concerns.[4]

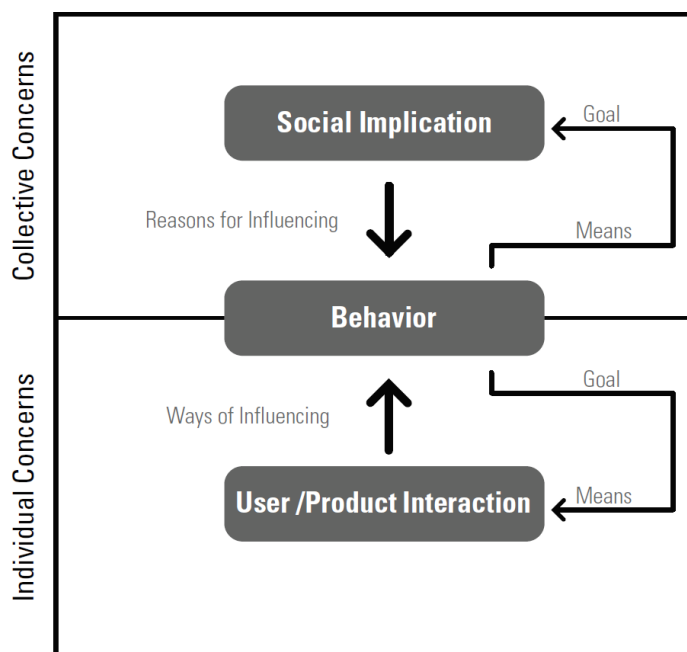**Figure 6** – Framework of Individual and Collective Concerns from [87]

---

[4]   Even if the collective goal can be seen as a benefit for the user, in the long run, we would still patronize the user by deciding what's best for them (see Section 4.3).

If PSS should use rational persuasion to achieve an overlap of individually preferred and collectively preferred target behavior, we need to find out what collective concerns coincide with individual concerns. Self-interested individuals have been shown to behave more sustainably when the behavior results in a personal benefit [21, 33]. Design can set incentives regarding individual concerns to reach for a behavior that is also desirable from a social perspective (see Figure 6) [87]. In the field of sustainability, there are several such overlaps. Individual sustainable behavior is often not based on sustainability-related incentives, but on non-environmental goals such as health improvement or money-saving [33]. Another overlap of sustainability and health concerns is found in the area of pesticides, which are known to not only cause environmental issues but also health issues [63]. This is exactly where PSS can provide support by showing the benefits for the user, not based upon distortion of the facts or manipulation, but based upon scientifically substantiated facts and rationale. Instead, we should focus on providing the user with persuasion approaches – such as rational persuasion – that don't interfere with the user's autonomy. The facts about the benefits of sustainable behavior exist, so why not use them to convince the users?

Through an example, the benefit of rational persuasion should be made clear. For that, consider a user group that usually opts to take the car to work instead of the bike. The PSS could try to convince them to take the bike to work by referring to the environmental benefits implied by it. However, these long-term collective benefits won't provide much of a reason for them to change their behavior [21]. Instead, PSS can argue for factors that support their individual well-being. If health plays an important role in their life, taking the bike to work might very well help them improve their fitness and health. PSS could show the users the relation between sustainable mobility behavior and benefits for their health to persuade them [42]. Not only that, but physical activity is also beneficial for mental health [73]. In addition, the users might also see financial benefits, such as saving gas money.

> **Generic Requirement 2: Rational Persuasion**
>
> To achieve the persuasion goal, PSS should avoid interfering with the persuadee's liberty as much as possible. A good alternative persuasion approach is rational persuasion, which aims to persuade users through facts, simulation and by addressing their locus of control.
>
> Instead of interfering with the user's liberty, rational persuasion can focus on finding means for persuasion based on the user's personal benefit.

4.3   Paternalistic Approaches

4.3.1   Paternalism

One concept that is associated with persuasion is called "paternalism". The term is not clearly defined, but Dworkin [22] defines paternalism as "[...] the interference with a person's liberty of action justified by reasons referring exclusively to the welfare, good, happiness, needs, interests or values of the person being coerced" [22, p. 65]. This suggests, however, that paternalism can only refer to any form of coercion for actions or behavior change that favors *the person* that gets patronized. As far as the well-being of the individual is concerned, it is very difficult for us to decide what *actually* helps the well-being of a person. This further imposes a challenge on defining the persuasion goal of such a PSS. But what if we extend the term as defined by Dworkin [22], by not only considering the individual's welfare, but also social goals such as sustainability? This was already discussed in Section 4.2 when talking about the conflict between the short-term interests of the users and the mostly long-term collective interests.

It is rather obvious that paternalism presents itself as a strong contrast to most people's understanding of freedom and thus, would require an extensive value shift for a large group of individuals. Even if a paternalistic approach would be somewhat feasible, it is still rather debatable whether we should interfere with an individual's autonomy. Even though certain paternalistic approaches are still relevant today – such as mandatory seat belt wearing – it is also being restricted more and more in many areas. An example of this is the development

in medicine, where the asymmetric relationship between the physician and the patient has moved to one that respects the patient as an autonomous individual who expresses his own will regarding health-related decisions [81]. PSS is no exception to this and should refrain from using heavily paternalistic approaches for persuasion.

4.3.2   Manipulation

Another vague term related to paternalism is "manipulation". Manipulation is often referred to as the influence taken on others in a deceitful way [74]. This means, that the user is mostly unaware of the influence exerted upon him, whereas persuasion requires some kind of interaction between the persuader and the persuadee [34]. At the same time, it is not completely clear where the boundary between persuasion and manipulation lies [82].

There are several arguments for and against manipulation, but the disadvantages outweigh the advantages most of the time. It can be argued that a more manipulative approach might be suitable in certain scenarios. One such example is PT that aims to help people out of their addiction to life-threatening drugs, where a manipulative approach – making use of psychological responses conducive to behavior change – might be more suitable and still considered ethical if the user's informed consent is given [54]. However, history has thought us that the manipulation of behavior holds great potential for abuse. Even though environmental problems such as climate change can be seen as very serious and urgent problems of our time, the deliberate deception of PSS users should, if at all, only be considered as a last resort.

Besides the mostly undisputed ethical downsides, there are – as argued for throughout this work – several more ethically acceptable approaches to address these problems. A lot of these opportunities seem even more promising due to recent technological advancements such as ubiquitous computing [47]. There is no obvious scenario in which such a manipulative approach would seem fitting for PSS. In general, we should use as little manipulation as possible to achieve our persuasion goal. It is also up for discussion whether a system using manipulation as means to persuade should even be labelled "persuasive technology" [83]. Of course, any form of manipulation used by PSS should happen within the consent of the user and the behavioral goal should be in the sense of the user.

4.3.3   Libertarian Paternalism

Although also heavily debated, "libertarian paternalism" acts as a milder approach compared to an entirely paternalistic or manipulative one. Libertarian paternalism supports the idea that it is legitimate to affect behavior while respecting freedom of choice [55]. Because the presentation of information can never be neutral, some designers choose to present information such that it leads the user to act in a certain way, which is called nudging [95]. Nudging follows the idea of libertarian paternalism and is a rigorously debated topic amongst ethicists [7]. Nudges aim to achieve a certain behavior by altering the change environment with the help of certain strategies such as incentives or defaults [95]. The effectiveness of using default nudging was shown by setting the default option for organ donation from opt-in to opt-out, which nearly doubled the number of people consenting to be organ donors [43]. In the context of this work, nudging can be positioned somewhere between rational persuasion and manipulation, because nudging has some ethically debatable traits.

If the designers of PSS decide to encourage the users towards acting in a certain way through nudging, they are also – to some extent – responsible for the outcome (see Section 4.1.2) and the users might want to transfer the responsibility to the designers in case of a bad outcome. An example of this could be a car that automatically uses an eco-friendly driving mode (eco mode) or a computer that starts up in energy-saving mode by default. If the users of such a car expect it to behave as if it were in its *normal* operating behavior, while it is actually in eco mode, they might want to transfer the responsibility to the designers in case of an accident.

When trying to make use of nudging as a means for persuasion responsibly, transparency (see Section 4.6) plays a key role, which means that the user should not only be informed *that* he is being nudged but also *how*. One commonly used argument that supports nudging towards a more sustainable behavior is the inability of technology to be neutral [53, 68, 85]. Because of that, instead of nudging the user towards an arbitrary, unforeseen behavior, we should instead use experts (e.g., social psychologists) to determine an appropriate PSS design to account for the best possible outcome by acknowledging the responsibility that comes with designing PSS and PT in general. Designers should also find ways to disclose the usage of such nudging to the user in an unobtrusive way and make the user aware of the advantages and disadvantages of the behavioral goal towards which the system is nudging the user. That way, the user is not only free to opt out of the nudging, but also informed about what is happening to him.

> **Generic Requirement 3: Nudging**
>
> Nudging serves as an effective, slightly manipulative approach that requires a conscious application. It is justified by the fact that technology can not be neutral.
>
> A lot of responsibility comes with the design of nudges, which has to be accounted for. In contrast to coercion, nudging retains the user's freedom of choice. Nudging has to occur within the informed knowledge and consent of the user, which is supported through transparency-related measures.

4.4 Privacy

As with any system that collects data in one way or another, privacy also poses an area of concern for PSS, which should not be neglected. Just like human persuaders exploit information about people they persuade [8], it is also the case for systems that collect data, such as social networks [96] and PT [8]. That designers of PT should respect users' privacy at least as much as they respect their own privacy has been proposed as a general principle by Berdichevsky and Neuenschwander [8] to address the ethical dilemma of how much privacy invasion should serve as means for PT. This, however, relies on the regard of the designers for their own privacy, which might not be congruent with the views on privacy that the user has [8].

Instead of relying on the designers to decide how much privacy should be given to the user, the designers should select a user-centered approach, in which the user has the last word about how his data is used. In general, users have strong preferences regarding the usage of their data [48]. Instead of opting for a default immutable privacy setting, the system should be designed so that the user has the necessary tools to decide upon the appropriateness of the collection of his data and adjust it accordingly. To achieve this, Hoepman [36] has developed a set of strategies for privacy design based on previous findings regarding ways to help IT architects in designing their software to support privacy.

Table 2 shows six strategies that are especially relevant for PSS design. In contrast to systems

Design Strategy	Description
Minimize	Restrict the amount of personal data processed to the minimum.
Hide	Personal data should be hidden from plain view.
Separate	The processing and storage of personal data should occur in a distributed fashion whenever possible.
Aggregate	Aggregation should be used by processing personal data with the least possible detail while still being helpful.
Inform	Inform users about the processing of personal data.
Control	Give users the ability to control the processing of their personal data.

Table 2 – Excerpt of the Eight Design Strategies proposed by Hoepman [36].

used for the collection and sale of data, minimizing the amount of data collected is especially important for the design of user-centered PSS, because designers should not be inclined to collect more data of the user than necessary to achieve the desired persuasion goals. Hiding data is an important design strategy to prevent the abuse of personal data. These first two design strategies are essential for the means-end oriented collection of personal data for user-centered PSS. If personal data is stored in a distributed fashion (i.e., storing data from separate sources in separate databases) or locally, it hinders the system from creating complete profiles [36]. Designers can reduce the system's ability to attribute information to a single user by aggregating the data over groups of attributes or individuals [36]. Informing the users about the processing of their personal data comes with several ethical and practical benefits (see Section 4.6). Informing should also come with the user's agency over the processing of their personal data so that they can take conscious, informed action for or against it (control) [36].

> **Generic Requirement 4: Privacy**
>
> ---
>
> The careful handling of user data increases the user's trust towards the
> system. When processing personal data for persuasion, PSS designers
> should pay attention to the following privacy design strategies:
>
> - Minimize
> - Hide
> - Separate
>
> - Aggregate
> - Inform
> - Control

4.5   Personalization

Personalization can be used by PT to adapt the ends or means according to the user profile
[45]. Roughly speaking, the ends define *what* the user will be persuaded towards and the
means define *how* he is getting persuaded [45].[5] Personalization comes with many benefits,
but also certain problems that have to be addressed by the designers.

4.5.1   Benefits

Choosing the right means for persuasion greatly increases the effectiveness of PT compared
to a one-size-fits-all approach [15, 69, 70]. This has also been shown in three case studies by
Kaptein et al. [46]. Adaptive PT can make use of persuasion profiles to increase the persuasion
effectiveness by estimating the effect of certain persuasive strategies on the users through the
collection of user data [45]. Because finding and using an appropriate persuasion strategy for
each user is so rewarding, personalization is often applied to make this possible. As we have
seen in Section 4.1.2, going for a one-size-fits-all approach can also lead to a negative backfire,
often as a result of mistailoring or mistargeting. If a system makes use of user data to create
personalized content, the chance of mistargeting and mistailoring can be reduced. Based on
previous research regarding personalization for PT [46, 69], personalization will most likely be
an enhancement that can be used for most, if not all, design principles covered in this work.
However, this remains an educated guess and is subject to further research.

Personalization can also improve user acceptance, which is shown by referring back to the

---

[5]   In the area of PT, ends are often referred to as "target behavior", and means are also known as "persuasion
strategies".

findings of [92] shown in Section 3.1.2. As we have seen, UTAUT 2.0 includes age, gender and experience as moderators for certain factors that improve user acceptance and use behavior [92]. That means, further evaluating the particular moderating effects of these factors might be a way to increase the potential of personalization for PSS to achieve widespread usage.

Personalization is part of the primary task category [66]. This means, personalization also improves user acceptance by supporting performance expectancy, effort expectancy and facilitating conditions according to UTAUT 2.0 [13].

4.5.2   Problems

The collection and use of data required for personalization do not come without ethical problems. Besides refraining from manipulation and coercion as much as possible, any use of personalization for PSS has to be designed as reliable as possible to avoid unintended, reasonably predictable and unethical outcomes. Several arguments support the view that the use of persuasion profiling may sometimes be the most ethical thing to do [45]. The use of persuasion profiles, however, typically inherits two major ethical concerns: They are end-independent and able to act without making their adaptation known to individuals [45].

End-Independence. Even if the persuasion profiles were obtained and used ethically, there is the potential to use these profiles in unethical contexts if they are shared with third parties [45]. For instance, suppose a public transportation app made by a company wants to convince its users to use public transport more often. It does so by collecting data about their driving behavior to find suitable offers for them. For now, assume that the users know about the collection and use of their personal data. The sharing of this data with third parties can result in actions that were unintended by the users. Suppose the data is shared with a car dealer that uses this profile to persuade its users into buying a car, such as by providing personalized offers that fit their mobility behavior, with the sole intention of increasing the users' maximum willingness to pay. This will result in a backfire that might not be very sustainability-promoting or user-centered by enticing users to spend more instead of promoting sustainable behavior. If the data is shared further, it might be abused for even more unethical practices such as manipulation or deception of the users.

Non-Disclosure. Another problem of means-adaption through personalization occurs when the adaption is not disclosed to the user [45]. Because the behavior change effect might get reduced if the user is informed about the adaption [45], this tends to be considered a legitimate

approach. It can be seen as morally wrong to deceive a user this way, even if the motives are purely user-centered (see Section 4.3). But even if it is seen as legitimate to use undisclosed means-adaption, there is still room for the designers to profit from this non-disclosure for reasons which are not intended by the user [45]. This can occur when the system performs actions that might not be the most effective in terms of behavior change, but instead benefits the designer at the expense of the users, such as a system that focuses on building a persuasion profile that can be sold [45]. Coming back to our example with the public transportation app, we now assume that the means-adaption is not disclosed to the users. The company might be encouraged by the car dealer to collect data about the users' mobility behavior. Instead of providing offers that might be most beneficial both for the users and in terms of sustainability, the app might create offers that aid the system to create a more extensive profile about their mobility behavior.

To reduce the likelihood of misuse through end-independence and non-disclosure, the following privacy design strategies presented in Section 4.4 can be applied:

- Hide: Hiding the data restricts unauthorized access.

- Separate: By separating the data, such as storing certain pieces of information locally, third parties are prevented from creating complete personalization profiles.

- Control: Giving the user control about the usage context and sharing of his personal data prevents misuse.

- Aggregate: Aggregating information reduces the comprehensibility of the data for third parties.

- Inform: Stating the rationale and methods for data collection gives the user insights about potential misuse and promotes trust.

- Minimize: Only collect as much data as necessary to reduce the misuse potential.

> **Generic Requirement 5: Personalization**
>
> Personalization can be used by PSS to enhance the persuasion effect, prevent backfires and support user acceptance. To help against mistargeting and mistailoring, the privacy design strategies should be applied.

4.6   Transparency

A key principle for ethical persuasion is that the persuader's intent should be disclosed to the user [5, 8]. Thinking about the way PSS communicates with the user is not only important for reasons that consider the effectiveness of the persuasion, but also for transparency, which in turn favors ethical PSS design. Transparency is a way to overcome information asymmetry, which is also necessary for highly personalized or AI-based content to help people understand the system's actions.

4.6.1   Trust

Trust is an important concept that comes with several benefits that should be considered when designing ethical PSS. Being transparent about the underlying mechanisms that lead to certain decisions made by the PT has a positive effect on the user's trust towards the PT [77]. Instead of designing a black box that forces the user to blindly trust the system and shifts the responsibility towards the designer, the system should provide information about persuasive strategies, data collection (see Section 4.4) and target behavior, which gives the user back his liberty. Trust can not only be seen as a good thing from an ethical perspective but can also serve as a mechanism to reduce complexity by providing guidance in the case of ambiguity and uncertainty [57]. Further, a lack of trust from the user towards the system might lead the user to refrain from using the system (or any PT) because he feels deceived by the system, whereas trust will probably lead to a stronger usage behavior [83]. As we have seen in Section 3.1, trustworthiness is part of the credibility category, which increases performance expectancy and thus, user acceptance.

4.6.2   Requirements for Transparency

The two requirements "comprehensibility" and "truth" support the transparency of PSS [83]. Even though Spahn [83] does not explicitly impose these conditions for PT, they still relate to this problem.

Comprehensibility. Just like in human communication, the interlocutor has to make his statements comprehensible for the communication to be successful [83]. If the PSS is easy to understand for the user, the designer can ensure effective communication of its persuasion

and transparency-related content. Further, flooding the users with information does not solve this problem. Providing a large amount of unorganized information can lead to information overload, where not all communication inputs can be processed by the user [40]. For instance, there is evidence about negative reading behavior of the Terms of Service due to information overload [65]. This further supports the reduction design strategy proposed in Section 3.1.1. There are effective ways of providing information comprehensibly to the user, such as just-in-time prompts (see Section 5.2).

Truth. For the design of PSS however, this requirement is not as trivial as it may seem. For instance, the persuasion effect might be higher if a system uses false information [83]. We can think about a public transportation app that fakes (or overemphasizes) displayed information of CO_2 savings by using the train to convince the user of using the train more often. Even if only truthful information is shown to the user, hiding other relevant information from the user can also be seen as untruthful. This means that possible disadvantages should not be hidden from the user and brings us back to the statement that PSS should not convince based on hiding facts or displaying untruths, i.e., using deception to convince. Further, there is the problem that the persuasion effect might be improved through the use of exaggeration, but in the long run, if the users learn that the PSS does no show them the truth, it might leave them unwilling to trust the system [83].

> **Generic Requirement 6: Transparency**
>
> Transparency serves as a general requirement for ethical PSS to overcome information asymmetry and promote the trustworthiness of the system. The increase of trustworthiness also favors user acceptance. To achieve transparency, designers have to ensure the availability of comprehensible, truthful information.

4.7   Value Sensitive Design

Because this work considers user-centered and ethical PSS, it seems indispensable to consider the target group's values. In this framework, value refers to what a person considers important in life [31]. Value sensitive design (VSD) is a framework that tries to consider stakeholder's

values through an iterative and integrative tripartite methodology [31]. It does that by acknowledging the fact that technology shapes our values anyway so that it lies in the designer's responsibility to steer the design of the system in a way that supports the values of direct and indirect stakeholders [53], implying that we should be mindful and intentional about the evolution of values through the widespread use of technology [53, 62].

However, interfering with people's values has – as history taught us – a bad reputation attached to it, for which VSD provides designers with a rich set of tools to preserve the users' autonomy and interfere responsibly with the values of the users. VSD can aid the designers in understanding the values of the stakeholders involved and capture the moral reactions to the designed PSS [19]. It also contributes to the design of ethical PSS, because it helps designers exploring reasons a design might be good or bad. This reveals value conflicts and provides methods for exploring the value implications of PSS, thus preventing problems that might otherwise not be seen until the system is deployed [19].
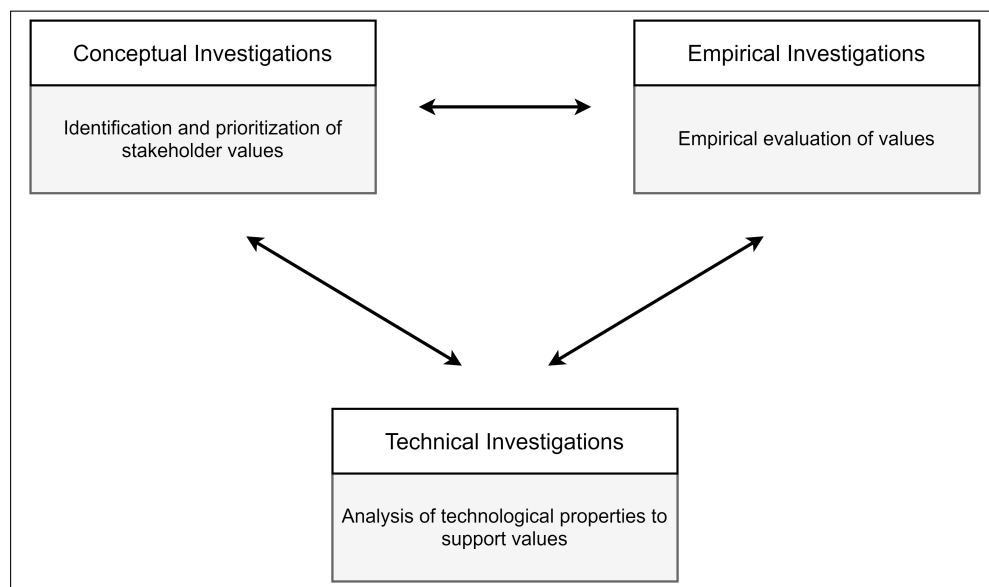
Figure 7 – Simplification of VSD based on the works of Friedman et al. [31].

VSD is similar to stakeholder analysis, but additionally checks how values can be implemented in the design and can find out how values are interpreted by stakeholders [53]. It comprises of three investigations (see Figure 7):

1. **Conceptual Investigation**: Finding out who the direct and indirect stakeholders are, how they are affected by the system, finding trade-offs or even defining the value itself (e.g., what is privacy?) [31].

2. **Empirical Investigation**: Finding empirical information, such as what values are prioritized for the stakeholders [31].

3. **Technical Investigation**: Finding out how suitable a system is for certain values and activities and designing a system that supports the values which were identified during the conceptual investigation [31].

Because the context of PSS can vary, it makes sense to assess the important context-specific values. It makes little sense to define a strict process model which can be universally applied. Because of that, the important values of the stakeholders should be assessed for a specific context. For instance, if a PSS can combine sustainability-oriented benefits with health benefits for the user, it might be appropriate to make use of fear (such as mentioning diseases caused through the ingestion of pesticides) to convince the user to eat more organic food.

The assessment of predominant context-specific values through VSD seems to be a promising approach towards designing ethical PT. One major drawback of VSD, according to Knowles and Davis [53], is that the whole process takes a long time, which is not compatible with the urgent need to solve all-endangering environmental problems. However, the assessment of values is not only important for moral reasons, but can also play a role in the system's effectiveness. As we have seen in the UTAUT 2.0 model, the appropriation of technology is very complex and has to be understood thoroughly to achieve a reliable way of widespread usage. VSD can help us in finding value-bound traits of the PSS that promote user acceptance. While VSD is a mostly theoretical framework, future researchers should recognize VSD as a promising method for ethical PSS and research further potentials of the framework by extending it and evaluating it in more practical contexts.

Generic Requirement 7: Value Sensitive Design

By obtaining a detailed understanding of stakeholders' context-specific values, VSD serves as a way to take responsible actions against arbitrary value influences of PSS design. Consider VSD to improve persuasion effectiveness and promote user acceptance.

4.8   Participatory Design

The goal of participatory design (PD) is to promote a dialogue between the user and the designer [10] by trying to get potential users to participate throughout the design process [6]. However, there is no value in merely asking the users with the sole intention to gain moral acceptance if their contributions are completely ignored during the design of the system [53]. Instead, PD can provide both ethical and practical value for the system. One effective way to make use of PD is to let potential future users brainstorm for ideas without technical constraints to make sure you are not missing out on interesting ideas [60]. In the next step, you can ask the users to think about ideas that could be implemented into a final product and then keep narrowing down the requirements towards a realistic design scenario that considers the available data [60]. An effective method for this could be the creation of mockups [6, 23, 60].

One problem of PSS that PD can help to solve is the question of responsibility, as discussed in Section 4.1.2. Berdichevsky and Neuenschwander [8] state that "[...] the creators of persuasive technology must anticipate and assume responsibility for all 'reasonably predictable' outcomes" [8, p. 2]. PD can not only help in getting an effective system design in terms of usability but can also help in designing an ethical PSS by letting potential users have a say in the form the PT ultimately takes [19]. PD can assist in providing transparency-related benefits as well. It allows for the users to take part in the design process through symmetric relations between designers and users, making it difficult for designers to incorporate falsehoods in their design [20].

By providing potential users with the ability to contribute their ideas during the design process, the designer is partly relieved from the pressure of thinking about every little detail. However, the design of PT is a profession that requires a lot of skill and knowledge. After all, domain experts know about concrete solutions and should make decisions based on the evaluation of the best arguments [53] and not because it seems morally right to implement the ideas of the users. This means that end-users should not take over the role of the designer. Instead, the users should disclose their needs and expectations towards the system, which is then reviewed by the designer.

Van Wynsberghe and Robbins [90] propose that ethicists should also take part in the design [90]. Instead of taking the role of morality police, ethicists can also provide a practical use for the design by discovering the values of the user and turning them into design requirements

[90]. As mentioned before, considering the values during the design process can also open up ways to promote system acceptance. Suppose the designers conclude that target users value personalization and privacy settings when using a certain PSS, and the designers consider this information during the development of the system. Based on that, the PSS appeals to the user and advocates his trust towards the system. Trust can then further increase the persuasion effect [94].

> **Generic Requirement 8: Participatory Design**
>
> Designers should use PD during the design of PSS to gain several ethical and practical benefits. Do not undermine the role of the designer when including users and ethicists in the design of PSS.

5   Design Principles

This section presents a handful of practical design principles, most of which are already widely used for PT or information systems in general. The ethical requirements of Section 4 are applied to the design principles in this section, revealing not only areas that require special ethical attention but also various unused potentials for the design of ethical PSS. The key findings for each section are presented in green boxes, making up a compact set of well-formulated design guidelines.

5.1   Value-Added Design

A significant aspect of creating a superior PSS is to create an impression of added value, resulting in a product that is seen as more attractive to consumers [4]. An example is the colorful illumination of a water sink called "Waterbot" to increase the perceived value of the water [4]. Value-added design can also be applied if the system is designed for long-term usage. In this case, we do not rely on short-term stimulation but offer design that engages the user in using the system regularly, such as giving the user identity and thus, promote the hedonic quality of the system to improve persistence and personal relevance [35].

According to the rational persuasion principle, we still want to provide truthful information and not distract from it through fancy design and illuminations. Further, we still want to emphasize information according to its relevancy and not overemphasize information only because it's what the user needs to hear to reach the persuasion goal. But there is nothing wrong with making the interaction enjoyable and engaging for the user. After all, the persuasive effect will not matter if the system is not being used. Value-added design can also contain feedback information to make the interaction with the system even more engaging for the user by increasing the surface credibility. In the case of the sink, the system used colorful illumination to display the water temperature [4].

> **Design Guideline: Value-Added Design**
>
> ---
>
> **Category**: System Credibility
>
> **Design Principles**: Surface Credibility
>
> To promote the (long-term) usage of the system, the system can make use of Value-added design to make the system more engaging for the user. To generate trust and prevent overemphasizing, the system should not refrain from displaying truthful information according to its relevancy as much as possible.

5.2   Just-in-Time Prompts

As mentioned in the introduction, one of the main potentials of PSS is to provide relevant information at the right time. One effective solution for this is the usage of "just-in-time prompts" [1, 3, 4, 39, 76]. Prompts help the user perform their intended actions by reminding them and providing them with the right information. How close the system places the prompt to the activity and if it contains the relevant information for someone to act appropriately determine the effectiveness of the prompt [58]. But the system requires context data for these prompts to work effectively [39].

Developments in technology such as mobile computing devices (e.g., mobile phones, watches with sensors or wristbands) have further supported the ability to provide the correct information at the right time in an unobtrusive way [39]. Waterbot is an example of effective just-in-time prompts implementation, which uses visual and auditory reminders to help the users to achieve sustainable use of water from the sink [4]. Another example is a system that reduces residential energy consumption by having a context-aware system [80]. This system does not only provide consumption feedback to the user but also points out how to save energy, such as when to use the windows to cool the house instead of the air conditioner [80].

There are, however, some things to consider when designing PSS that make use of these prompts. The first consideration is that the collected data might be sensitive for the user. Hence, the system should protect it from access through third parties [56]. For this, the hide privacy design strategy from Section 4.4 can be applied to address this problem. As long as the system disclosed the collection and usage of context data to the user and the acquired data

remains to be used in an ethical context, this practice seems to be one of the more effective and ethically reasonable methods for PSS.

Another problem concerns the long-term collection of data to generate personalized, context-aware prompts that might raise concerns about privacy [56]. At the same time, the designer wants to fulfill his responsibility of avoiding mistailoring. To solve this dilemma, we can make use of transparency within our design. The first step would be to disclose how user and context data is collected and why it is necessary to achieve the mentioned benefits. Just-in-time prompts can very well use transparency to generate trust for the user. Instead of only providing just-in-time information and behavioral implications, the system can make the user aware of the rationale behind the prompt, which is especially useful for highly personalized content or decisions made by AI-based systems. For a public transportation app, this could be an explanation of a personalized train ticket offer, such as: *"We have provided you with this train ticket offer based on your preference to take the train instead of the bus."*.

Just-in-time prompts can be assigned to the primary task support category, addressing numerous design principles such as reduction. They reduce the information shown to the user by only requiring his attention when it is relevant to reach his target behavior. As we can see from our example with the Waterbot, just-in-time prompts can also help to make the users aware of their water usage through visuals (simulation). They can also make use of personalization to address the relevancy and timing of the prompts. Using context data to provide the user with relevant information at the right time, enhances the ability to provide tailored messages – such as suggesting a bike ride if the weather is suitable – to promote sustainable mobility behavior (tailoring) [2].

> **Design Guideline: Just-in-Time Prompts**
>
> ---
>
> **Category**: Primary Task Support
>
> **Design Principles**: Reduction, Tailoring, Personalization, Simulation
>
> Just-in-time prompts can be used as a way to effectively provide users with information. For this, designers should make sure that the prompts ...
>
> - ... contain relevant information (e.g., through context data).
>
> - ... are located in time and space, such that the user perceives the prompt at the right time to adjust his actions.
>
> To address ethical concerns, designers of PSS should consider the following practices when using just-in-time prompts:
>
> - Make use of personalization to prevent mistargeting.
>
> - Apply the privacy design strategies.
>
> - Disclose the rationale behind the prompts if possible.

5.3   Persuasive Strategies

Besides the more generic, high-level design principles for behavior change, several persuasive strategies have been explored during the research of PT, especially in the field of sustainable urban mobility. Research in that field is of particular importance, because of the massive effects personal transportation has on the environment [32]. These persuasive strategies can often be assigned to one or more categories of conceptual design as discussed in Section 3.1.1. Due to the more practical nature of these persuasive strategies, they can be evaluated in a more realistic setting that provides us with insights into their actual use cases. Three persuasive strategies are especially well suited for the evaluation against the previous findings regarding ethics and usability. The persuasive strategies reward, competition and social comparison are very influential for achieving the targeted behaviors [14, 72]. Table 3 shows an overview of these strategies with corresponding real-world examples[6].

---

[6]   Excerpts of the examples are provided in Appendix 1 (Figure 9 and 10).

Persuasive Strategy	Description	Example
Challenges and Rewards	Providing ways for the user to challenge himself. Rewards can serve as additional incentives for the completion of challenges.	Setting a personal sustainability goal in a retailer's customer loyalty programme.
Social Comparison	Providing social context information as a way to motivate users.	Comparison of own purchasing behavior with the national average.
Self-Monitoring	Keeping track of the user's sustainability contributions.	Showing the development of the user's purchasing behavior.

Table 3 – Three impactful persuasive strategies out of the many found in the literature, such as [14, 72]

Competition and Rewards. One persuasive strategy that is commonly seen in the literature is some form of competition or challenge. Competition is the intrinsic motivation in humans to outperform one another by adopting certain behaviors [68]. PT, however, can not only provide ways to compete against others but also to compete against the system itself [72]. These challenges aim to persuade the user into a target behavior through different kinds of rewards [11]. Users get rewards as a result of accomplishing an individual achievement or task [14]. These rewards are often based on social recognition [9, 37, 97], money [37], or virtual pseudo-rewards, such as points or badges in virtual games [49]. In the area of sustainable urban mobility, most of these challenges aim to persuade users into taking alternative means for transportation other than the car, such as the train or the bike [37]. Especially in the field of urban mobility, additional incentives for the completion of challenges can be created for the user, such as financial incentives or information [37]. An example of such an incentivized challenge could be to offer the user a coupon if he takes the bike for a particular trip instead of the car [37]. Further, the implementation of challenges into PT promotes the long-term usage of PT compared to only providing feedback to the user [78].

Social Comparison. There are several ways to implement challenges in PT, that have shown to be very effective. One approach could be to facilitate behavior change through leaderboards that let users compare their achievements with others [14, 97]. While leaderboards serve as an information source to compete with others, social comparison can also be used

differently. Most notably, social comparison serves as a way for people to evaluate their actions if there's a lack of objective cues [16, 25]. More specifically, people use actions of similar others to assess the correctness of their actions [16]. The actions of others are even more influential if a lot of people do the same thing [16]. For example, think about all the times people started clapping after a concert as soon as someone else did. PT can make use of this by showing the actions of other users [71].

Self-Monitoring. By collecting data of users' mobility behavior, the system can additionally support the ability to track their performance and display the progress they made towards attaining a particular target behavior. Self-monitoring builds on the human needs for awareness [71]. An example for this could be to show graphs of CO_2 emissions caused by the user during different periods when traveling [2]. However, self-monitoring alone is perceived as tedious and should also be combined with rewards to make it more engaging for users [71].
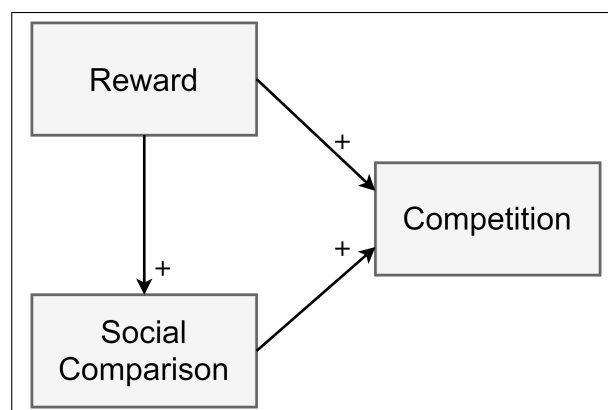

Figure 8 – Interdependency between reward, social comparison and competition based on the findings of [72].

Oyibo and Vassileva [72] conducted a survey to research the correlation between reward, competition, social comparison and social learning. As we can see in Figure 8, reward and social comparison both support competition [72]. Reward also supports social comparison [72]. Based on that, we can draw several implications for PT design. For example, if the target audience of a particular PT is susceptible to reward and social comparison, it makes sense to include some form of competition as well [72].

One effective way to further enhance the usefulness of these strategies is to personalize the challenges based on the combination of personal user data – such as the history and habits of each user – and context information [50]. However, as discussed before, the collection of sensible data comes with certain responsibilities. To address this concern, the system can offer

ways to adjust settings for privacy and data sharing. For example, "tripzoom" is a mobile app that creates individual mobility profiles and is being tested in several European cities [9]. These profiles also contain information about the user's performance regarding CO_2 emissions and they can be shared with the tripzoom-community and other existing social networks [9]. However, the app also provides its users with the ability to adjust settings for privacy and data sharing that allows them to stop using this feature [9].

Designers of PSS should disclose any collection of personal data to the users since it does not negatively impact the system's effectiveness. If the personal data is not forwarded to third parties without the users' consent and the privacy design strategies are considered appropriately, the Challenges and Rewards design strategy presents itself as a promising opportunity for ethical PSS. I suggest further research on the applicability of this strategy in other areas of sustainable behavior, such as the food industry or energy consumption.

Design Guideline: Persuasive Strategies

Category: Primary Task Support, Dialogue Support, Social Support
Design Principles: Self-Monitoring, Rewards, Social Comparison, Competition

There are several persuasive strategies for PT design. The following persuasive strategies show particularly great potential for PSS:

- Competition and Rewards

- Social Comparison

- Self-Monitoring

---

6 Results

This section briefly lists the results acquired throughout the work, based on findings from the literature as well as conclusions and implications drawn from them. The results will be structured based on the research questions formulated in the Introduction.

Q1: What generic requirements should be considered by designers of ethical PSS?

When designing PSS, designers should consider the eight generic requirements discussed in this work: preventing backfires, rational persuasion, nudging, privacy, personalization, transparency, value sensitive design and participatory design.

Q2: How can these generic requirements be applied to well-known practical design principles and strategies?

The following persuasive design strategies are very promising for the design of ethical PSS: value-added design, just-in-time prompts, competition, reward, social comparison and self-monitoring.

Q3: What opportunities and risks come with the design of ethical PSS?

The design of ethical PSS comes with many risks on one hand, but also with a lot of potentials on the other hand. The relationship between the four categories of design principles by Oinas-Kukkonen and Harjumaa [66] and UTAUT 2.0 (see Section 3.1.3) give designers the ability to work towards an effective and widely accepted PSS. Considering the three factors of behavior (motivation, ability and triggers) for PSS, designers can work even more purposefully towards behavior change. The generic requirements mentioned in this work show that the design of PSS comes with a lot of responsibility and requires caution. However, rational persuasion, personalization, nudging, VSD and PD also offer a large set of effective tools for designers and can contribute towards an ethical PSS design if used correctly.

Q4: How should we value sustainability in contrast to an individual's freedom and welfare?

This research question will be part of the discussion in Section 7.

7 Discussion

This work discussed design principles and strategies that assist PSS design in one way or another. However, some of these practices come with certain ethical issues, which have to be avoided when designing ethical PSS. One of these practices is the design of PSS that leads to harmful outcomes, as discussed in Section 4.1.2. A majority of challenges PSS designers face tend towards finding a golden mean, i.e., balancing costs and benefits of using certain practices. But what exactly are the costs and benefits of ethical PSS? The two main benefits mentioned mostly in this work are sustainability-related improvements based on behavior change and the gains for the users when using PSS, such as health improvements. The costs consist of two main factors: ethical losses and the expenses of the users for using the system.

One of the most significant uncertainties of PSS design lies in the balancing of collective and individual welfare. To answer this, a fundamental question needs to be addressed: How urgent are environmental problems – such as climate change – compared to the users' well-being? Knowles and Davis [53] argue that the current pace of ethical PT is too slow considering the many harmful effects of climate change. This work aims to show that we do not need to put the users' well-being in the background to obtain a more sustainable behavior. Instead, we should apply the mentioned design principles to exploit the full potential of ethical and user-centered PSS design. The growing threats of environmental problems should induce researchers to increase the pace of effective PSS design research and find more potential application areas.

Instead of trying to preserve the users' values, PT should aim to deferentially, but proficiently promote new values and facilitate discourse [53]. Rational persuasion acts as an approach that engages the users into critically questioning their existing beliefs by exposing them to new knowledge and facts. Promoting users' awareness of environmentally problematic behavior will also prevent self-licensing. PSS should not "trick" the users into a specific behavior through manipulation, coercion or the predominant use of nudging. Instead, PSS can make the users aware of solutions regarding environmental concerns. Increasing the users' sense of efficacy could be a way to achieve this.

The best-case scenario for PSS would be to promote the users' motivation to take proactive measures against environmental problems. This can be achieved if PSS causes a fundamental value shift and shows users the relationship between their actions and their effect on the environment. Improving the users' awareness and knowledge of environmental problems and solutions might lead them to a longstanding sustainable behavior, even if they stop using the

system. This exceeds the scope of this work and can certainly be a worthwhile focus of future research.

As mentioned, Knowles and Davis [53] argue for a less "safe" way of persuasion because the current practices are too ineffective considering the imminent danger coming from climate change. While most ethicists consider the use of fear as inappropriate for PT, they see it as an inevitable evil that might not only be the most powerful, but also the most appropriate way towards substantial sustainability-related behavior change [53]. By showing pre-diabetic individuals threatening mental images about the onset of diabetes, fear acts as an effective measure to promote physical activity and a healthier eating behavior [24].

The use of fear for persuasion is not the most appropriate approach for PSS. In Section 4.3, arguments were provided that support the reasoning why manipulation for PSS is generally out of the question. The use of psychological responses for persuasion might be suitable for life-threatening addictions, which are hard to overcome only through rational persuasion. But this is different for PSS. As shown throughout this work, rational persuasion offers a much more promising and less ethically problematic approach for persuasion. But PSS relies on a more or less voluntary usage. This induces PSS designers to develop a system that will be used by a meaningful amount of users and thus, account for non-sustainability-oriented factors (see Section 3.1.2). Because after all, significant change can't be achieved without a substantial amount of users. However, as mentioned before, if we find overlaps between individual goals (such as health) and collective goals (such as sustainability), the usage of fear to achieve a target behavior might be up for discussion again. This is a potential subject for further research regarding this specific application of PSS.

8   Conclusion

This work should serve future designers of ethical PSS as a framework. Through the critical reflection of general and PSS-specific ethical concerns, this work evaluated eight generic requirements for ethical PSS design. The relationship between the design principles proposed by Oinas-Kukkonen and Harjumaa [66] and the UTAUT 2.0 opens the door for designers to indirectly promote user acceptance during the design of PSS.

The main part of this work identified and critically reflected upon several ethical considerations for PSS design by seeking insights from different domains of information system design. Designers of PSS should account for unintended and unethical outcomes caused by their system. Rational persuasion has been identified to be especially promising for PSS because of the potential overlaps of individual and collective goals. Further, the unavoidable influence of technology should be guided in the right direction through conscious nudging, while still maintaining the freedom of the users to decide upon their behavior and actions. Personalization was found to be not only an enhancement for existing practices but can serve as an aid to prevent backfires, such as mistailoring or mistargeting. However, to prevent potential abuse of the acquired persuasion profiles and to protect users' privacy, the privacy design strategies must be applied. Additionally, the users need truthful, comprehensible information to give them the necessary knowledge they need to make qualified decisions about their data usage and traceability of the system.

VSD and PD have been identified to offer great added value for ethical PSS. Including the opinions of stakeholders during the design process supports the designers when facing the challenging task of accounting for different values relevant for the system design. During the design process, ethicists can assist the designer when dealing with value-related design decisions, but should not take over the designer's role.

The acquired knowledge and generic requirements of Section 4 have been applied and evaluated in well-known practices of PT design. Just-in-time prompts have been shown to implement some of the design principles proposed by Oinas-Kukkonen and Harjumaa [66]. By considering a few ethical requirements – such as transparency and privacy – nothing stands in the way of realizing the full potential of this flexible design practice. Personalization is a great enhancement for just-in-time prompts that improves the precision of the prompts and provides measures against mistargeting and mistailoring.

This work pointed out the necessity of urgent measures required to prevent a global disaster through environmental problems, climate change being one the most significant. However, more research is needed to further investigate the effectiveness of PSS design. Also, more work is needed to find a conducive balance of individual values and collective concerns that is up to the difficult task of raising awareness of this imminent calamity. If this awareness is increased, users might be willing to take on more constraints[7], which further increases the potential of ethical PSS.

---

[7] In this context, constraints refer to knowledge or values that lower the amount of possible actions considered ethical by an individual.

9 References

[1] C. K. Aitken, T. A. McMahon, A. J. Wearing and B. L. Finlayson, "Residential water use: Predicting and reducing consumption 1", *Journal of Applied Social Psychology*, vol. 24, no. 2, pp. 136–158, 1994.

[2] E. Anagnostopoulou, E. Bothos, B. Magoutas, J. Schrammel and G. Mentzas, "Persuasive interventions for sustainable travel choices leveraging users' personality and mobility type", in *International Conference on Persuasive Technology*, Springer, 2018, pp. 229–241.

[3] E. Aronson and M. O'Leary, "The relative effectiveness of models and prompts on energy conservation: A field experiment in a shower room", *Journal of Environmental Systems*, vol. 12, no. 3, pp. 219–224, 1982.

[4] E. Arroyo, L. Bonanni and T. Selker, "Waterbot: Exploring feedback and persuasive techniques at the sink", in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2005, pp. 631–639.

[5] B. M. Atkinson, "Captology: A critical review", in *International conference on persuasive technology*, Springer, 2006, pp. 171–182.

[6] S. Bφdker, K. Grφnbæk and M. Kyng, "Cooperative design: Techniques and experiences from the scandinavian scene", in *Readings in Human–Computer Interaction*, Elsevier, 1995, pp. 215–224.

[7] A. Barton and T. Grüne-Yanoff, "From libertarian paternalism to nudging—and beyond", *Review of Philosophy and psychology*, vol. 6, no. 3, pp. 341–359, 2015.

[8] D. Berdichevsky and E. Neuenschwander, "Toward an ethics of persuasive technology", *Communications of the ACM*, vol. 42, no. 5, pp. 51–58, 1999.

[9] J. Bie, M. Bijlsma, G. Broll, H. Cao, A. Hjalmarsson, F. Hodgson, P. Holleis, Y. van Houten, K. Jacob, J. Koolwaaij *et al.*, "Move better with tripzoom", 2012.

[10] G. Bjerknes and T. Bratteteig, "User participation and democracy: A discussion of scandinavian research on system development", *Scandinavian Journal of information systems*, vol. 7, no. 1, p. 1, 1995.

[11] A. Bliznyuk, "Green daily guide. easier environmentally friendly transportation with the help of mobile technologies", in *2011 International Conference on Collaboration Technologies and Systems (CTS)*, IEEE, 2011, pp. 612–617.

[12] E. Bothos, G. Mentzas, S. Prost, J. Schrammel and K. Röderer, "Watch your emissions: Persuasive strategies and choice architecture for sustainable decisions in urban mobility.", *PsychNology Journal*, vol. 12, no. 3, 2014.

[13] B. Brauer, C. Ebermann and L. M. Kolbe, "An acceptance model for user-centric persuasive environmental sustainable is", 2016.

[14] M. Busch, J. Schrammel and M. Tscheligi, "Personalized persuasive technology–development and validation of scales for measuring persuadability", in *International Conference on Persuasive Technology*, Springer, 2013, pp. 33–38.

[15] R. B. Cialdini, "The science of persuasion", *Scientific American*, vol. 284, no. 2, pp. 76–81, 2001.

[16] R. B. Cialdini and M. R. Trost, "Social influence: Social norms, conformity and compliance.", 1998.

[17] J. Clapp, "The distancing of waste: Overconsumption in a global economy", *Confronting consumption*, pp. 155–176, 2002.

[18] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology", *MIS quarterly*, pp. 319–340, 1989.

[19] J. Davis, "Design methods for ethical persuasive computing", in *Proceedings of the 4th International Conference on Persuasive Technology*, 2009, pp. 1–8.

[20] J. Davis, "Early experiences with participation in persuasive technology design", in *Proceedings of the 12th Participatory Design Conference: Research Papers-Volume 1*, 2012, pp. 119–128.

[21] S. De Dominicis, P. Schultz and M. Bonaiuto, "Protecting the environment for self-interested reasons: Altruism is not the only pathway to sustainability", *Frontiers in psychology*, vol. 8, p. 1065, 2017.

[22] G. Dworkin, "Paternalism", *the Monist*, pp. 64–84, 1972.

[23] P. Ehn and M. Kyng, "Cardboard computers: Mocking-it-up or hands-on the future", *Design at work: Cooperative design of computer systems*, pp. 169–195, 1991.

[24] H. P. Enwald, T. Kortelainen, J. Leppäluoto, S. Keinänen-Kiukaanniemi, T. Jämsä, H. Oinas-Kukkonen, K.-H. Herzig and M.-L. A. Huotari, "Perceptions of fear appeal and preferences for feedback in tailored health communication: An explorative study among pre-diabetic", 2013.

[25] L. Festinger, "A theory of social comparison processes", *Human relations*, vol. 7, no. 2, pp. 117–140, 1954.

[26] L. Festinger, "Cognitive dissonance", *Scientific American*, vol. 207, no. 4, pp. 93–106, 1962.

[27] K. S. Fielding and B. W. Head, "Determinants of young australians' environmental actions: The role of responsibility attributions, locus of control, knowledge and attitudes", *Environmental Education Research*, vol. 18, no. 2, pp. 171–186, 2012.

[28] B. Fogg, "Chapter 9 - the ethics of persuasive technology", in *Persuasive Technology*, ser. Interactive Technologies, B. Fogg, Ed., San Francisco: Morgan Kaufmann, 2003, pp. 211–239, ISBN: 978-1-55860-643-2. DOI: https://doi.org/10.1016/B978-155860643-2/50011-1. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9781558606432500111.

[29] B. J. Fogg, "Persuasive technology: Using computers to change what we think and do", *Ubiquity*, vol. 2002, no. December, p. 2, 2002.

[30] B. J. Fogg, "A behavior model for persuasive design", in *Proceedings of the 4th international Conference on Persuasive Technology*, 2009, pp. 1–7.

[31] B. Friedman, P. H. Kahn, A. Borning and A. Huldtgren, "Value sensitive design and information systems", in *Early engagement and new technologies: Opening up the laboratory*, Springer, 2013, pp. 55–95.

[32] J. Froehlich, T. Dillahunt, P. Klasnja, J. Mankoff, S. Consolvo, B. Harrison and J. A. Landay, "Ubigreen: Investigating a mobile tool for tracking and supporting green transportation habits", in *Proceedings of the sigchi conference on human factors in computing systems*, 2009, pp. 1043–1052.

[33] R. Gifford and A. Nilsson, "Personal and social factors that influence pro-environmental concern and behaviour: A review", *International journal of psychology*, vol. 49, no. 3, pp. 141–157, 2014.

[34] R. Harré, "Persuasion and manipulation", in *Discourse and communication*, de Gruyter, 2011, pp. 126–142.

[35] M. Hassenzahl, M. Burmester and F. Koller, "Attrakdiff: Ein fragebogen zur messung wahrgenommener hedonischer und pragmatischer qualität", in *Mensch & computer 2003*, Springer, 2003, pp. 187–196.

[36] J.-H. Hoepman, "Privacy design strategies", in *IFIP International Information Security Conference*, Springer, 2014, pp. 446–459.

[37] P. Holleis, M. Luther, G. Broll, H. Cao, J. Koolwaaij, A. Peddemors, P. Ebben, M. Wibbels, K. Jacobs and S. Raaphorst, "Tripzoom: A system to motivate sustainable urban mobility", in *1st Int. Conf. on Smart Systems, Devices and Technologies*, 2012.

[38] W. IJsselsteijn, Y. De Kort, C. Midden, B. Eggen and E. Van Den Hoven, "Persuasive technology for human well-being: Setting the scene", in *International conference on persuasive technology*, Springer, 2006, pp. 1–5.

[39] S. S. Intille, "A new research challenge: Persuasive technology to motivate healthy aging", *IEEE Transactions on information technology in Biomedicine*, vol. 8, no. 3, pp. 235–237, 2004.

[40] T. W. Jackson and P. Farzaneh, "Theory-based model of factors affecting information overload", *International Journal of Information Management*, vol. 32, no. 6, pp. 523–532, 2012.

[41] T. Jackson, "Sustainable consumption", in *Handbook of sustainable development*, Edward Elgar Publishing, 2014.

[42] J. Jariyasunant, M. Abou-Zeid, A. Carrel, V. Ekambaram, D. Gaker, R. Sengupta and J. L. Walker, "Quantified traveler: Travel feedback meets the cloud to change behavior", *Journal of Intelligent Transportation Systems*, vol. 19, no. 2, pp. 109–124, 2015.

[43] E. J. Johnson and D. Goldstein, *Do defaults save lives?*, 2003.

[44] A. Jylhä, P. Nurmi, M. Sirén, S. Hemminki and G. Jacucci, "Matkahupi: A persuasive mobile application for sustainable mobility", in *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, 2013, pp. 227–230.

[45] M. Kaptein and D. Eckles, "Selecting effective means to any end: Futures and ethics of persuasion profiling", in *International conference on persuasive technology*, Springer, 2010, pp. 82–93.

[46] M. Kaptein, P. Markopoulos, B. de Ruyter and E. Aarts, "Personalizing persuasive technologies: Explicit and implicit personalization using persuasion profiles", *International Journal of Human-Computer Studies*, vol. 77, pp. 38–51, 2015, ISSN: 1071-5819. DOI: https://doi.org/10.1016/j.ijhcs.2015.01.004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1071581915000075.

[47] M. Kaptein and A. Van Halteren, "Adaptive persuasive messaging to increase service retention: Using persuasion profiles to increase the effectiveness of email reminders", *Personal and Ubiquitous Computing*, vol. 17, no. 6, pp. 1173–1185, 2013.

[48] M. L. Katz, "Multisided platforms, big data, and a little antitrust policy", *Review of Industrial Organization*, vol. 54, no. 4, pp. 695–716, 2019.

[49] R. Kazhamiakin, A. Marconi, M. Perillo, M. Pistore, G. Valetto, L. Piras, F. Avesani and N. Perri, "Using gamification to incentivize sustainable urban mobility", in *2015 IEEE first international smart cities conference (ISC2)*, IEEE, 2015, pp. 1–6.

[50] R. Khoshkangini, G. Valetto and A. Marconi, "Generating personalized challenges to enhance the persuasive power of gamification", in *Personalization in Persuasive Technology Workshop*, 2017.

[51] R. M. Kidder, *How good people make tough choices*. Morrow New York, 1995.

[52] R. Kight and S. B. Gram-Hansen, "Do ethics matter in persuasive technology?", in *Persuasive Technology: Development of Persuasive and Behavior Change Support Systems*, H. Oinas-Kukkonen, K. T. Win, E. Karapanos, P. Karppinen and E. Kyza, Eds., Cham: Springer International Publishing, 2019, pp. 143–155, ISBN: 978-3-030-17287-9.

[53] B. Knowles and J. Davis, "Is sustainability a special case for persuasion?", *Interacting with Computers*, vol. 29, no. 1, pp. 58–70, 2017.

[54] B. Knowles, P. Coulton, M. Lochrie and B. Wohl, ""Convince us": An argument for the morality of persuasion", in *CSCW'14 Workshop*, 2014, pp. 1–5.

[55] *Libertarian paternalism - wikipedia*, https://en.wikipedia.org/wiki/Libertarian_paternalism, (Accessed on 07/12/2021).

[56] S. W. Loke, J. Singh and H. Le, "Context aware pervasive persuasive systems for managing water and energy usage, and co2 emissions: Multi levelled policies, goals, and an expert systems shell approach", in *Workshop on Security and Privacy Issues in Mobile Phone Use*, 2008, p. 3.

[57] N. Luhmann, *Trust and power*. John Wiley & Sons, 2018.

[58] D. McKenzie-Mohr, *Fostering sustainable behavior: An introduction to community-based social marketing*. New society publishers, 2011.

[59] C. Midden, T. McCalley, J. Ham and R. Zaalberg, "Using persuasive technology to encourage sustainable behavior", *Sustainability WS Pervasive*, vol. 113, pp. 83–86, 2008.

[60] T. M. Miller, P. Rich and J. Davis, "Adapt: Audience design of ambient persuasive technology", in *CHI'09 Extended Abstracts on Human Factors in Computing Systems*, 2009, pp. 4165–4170.

[61] *Nachhaltiges einkaufen — migros*, `https://www.migros.ch/de/einkaufen/nachhaltig.html`, (Accessed on 07/14/2021).

[62] L. P. Nathan, B. Friedman, P. Klasnja, S. K. Kane and J. K. Miller, "Envisioning systemic effects on persons and society throughout interactive system design", in *Proceedings of the 7th ACM conference on Designing interactive systems*, 2008, pp. 1–10.

[63] P. Nicolopoulou-Stamati, S. Maipas, C. Kotampasi, P. Stamatis and L. Hens, "Chemical pesticides and human health: The urgent need for a new concept in agriculture", *Frontiers in public health*, vol. 4, p. 148, 2016.

[64] T. Nyström, "Gamification of persuasive systems for sustainability", in *2017 Sustainable Internet and ICT for Sustainability (SustainIT)*, IEEE, 2017, pp. 1–3.

[65] J. A. Obar and A. Oeldorf-Hirsch, "The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services", *Information, Communication & Society*, vol. 23, no. 1, pp. 128–147, 2020.

[66] H. Oinas-Kukkonen and M. Harjumaa, "A systematic framework for designing and evaluating persuasive systems", in *International conference on persuasive technology*, Springer, 2008, pp. 164–176.

[67] H. Oinas-Kukkonen and M. Harjumaa, "Towards deeper understanding of persuasion in software and information systems", in *First international conference on advances in computer-human interaction*, IEEE, 2008, pp. 200–205.

[68] H. Oinas-Kukkonen and M. Harjumaa, "Persuasive systems design: Key issues, process model, and system features", *Communications of the Association for Information Systems*, vol. 24, no. 1, p. 28, 2009.

[69] F. A. Orji, K. Oyibo, R. Orji, J. Greer and J. Vassileva, "Personalization of persuasive technology in higher education", in *Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization*, 2019, pp. 336–340.

[70] R. Orji, R. L. Mandryk and J. Vassileva, "Gender, age, and responsiveness to cialdini's persuasion strategies", in *International Conference on Persuasive Technology*, Springer, 2015, pp. 147–159.

[71] R. Orji, L. E. Nacke and C. Di Marco, "Towards personality-driven persuasive health games and gamified systems", in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 1015–1027.

[72]   K. Oyibo and J. Vassileva, "Investigation of social predictors of competitive behavior in persuasive technology", in *International Conference on Persuasive Technology*, Springer, 2017, pp. 279–291.

[73]   S. A. Paluska and T. L. Schwenk, "Physical activity and mental health", *Sports medicine*, vol. 29, no. 3, pp. 167–180, 2000.

[74]   *Psychological manipulation - wikipedia*, `https://en.wikipedia.org/wiki/Psychological_manipulation`, (Accessed on 07/08/2021).

[75]   J. Redström, "Persuasive design: Fringes and foundations", in *International Conference on Persuasive Technology*, Springer, 2006, pp. 112–122.

[76]   W. D. Russell, D. A. Dzewaltowski and G. J. Ryan, "The effectiveness of a point-of-decision prompt in deterring sedentary behavior", *American Journal of Health Promotion*, vol. 13, no. 5, pp. 257–259, 1999.

[77]   F. Sattarov and S. Nagel, "Building trust in persuasive gerontechnology: User-centric and institution-centric approaches", *Gerontechnology*, vol. 18, no. 1, 2019.

[78]   J. Schrammel, S. Prost, E. Mattheiss, E. Bothos and M. Tscheligi, "Using individual and collaborative challenges in behavior change support systems: Findings from a two-month field trial of a trip planner application", in *International Conference on Persuasive Technology*, Springer, 2015, pp. 160–171.

[79]   P. W. Schultz, V. V. Gouveia, L. D. Cameron, G. Tankha, P. Schmuck and M. Franěk, "Values and their relationship to environmental concern and conservation behavior", *Journal of cross-cultural psychology*, vol. 36, no. 4, pp. 457–475, 2005.

[80]   C. Seligman, J. M. Darley and L. J. Becker, "Behavioral approaches to residential energy conservation", *Energy and buildings*, vol. 1, no. 3, pp. 325–337, 1978.

[81]   M. Siegler, "The progression of medicine: From physician paternalism to patient autonomy to bureaucratic parsimony", *Archives of internal medicine*, vol. 145, no. 4, pp. 713–715, 1985.

[82]   J. Smids, "The voluntariness of persuasive technology", in *International Conference on Persuasive Technology*, Springer, 2012, pp. 123–132.

[83]   A. Spahn, "And lead us (not) into persuasion...? persuasive technology and the ethics of communication", *Science and engineering ethics*, vol. 18, no. 4, pp. 633–650, 2012.

[84]   A. Stibe and B. Cugelman, "Persuasive backfiring: When behavior change interventions trigger unintended negative outcomes", in *International conference on persuasive technology*, Springer, 2016, pp. 65–77.

[85] C. R. Sunstein, "Nudging and choice architecture: Ethical considerations", *Yale Journal on Regulation, Forthcoming*, 2015.

[86] K. Torning and H. Oinas-Kukkonen, "Persuasive system design: State of the art and future directions", in *Proceedings of the 4th international conference on persuasive technology*, 2009, pp. 1–8.

[87] N. Tromp, P. Hekkert and P.-P. Verbeek, "Design for socially responsible behavior: A classification of influence based on intended user experience", *Design issues*, vol. 27, no. 3, pp. 3–19, 2011.

[88] G. Tsai, "Rational persuasion as paternalism", *Philosophy & Public Affairs*, vol. 42, no. 1, pp. 78–112, 2014.

[89] P. A. Van Lange and J. A. Joireman, "How we can promote behavior that serves all of us in the future", *Social Issues and Policy Review*, vol. 2, no. 1, pp. 127–157, 2008.

[90] A. Van Wynsberghe and S. Robbins, "Ethicist as designer: A pragmatic approach to ethics in the lab", *Science and engineering ethics*, vol. 20, no. 4, pp. 947–961, 2014.

[91] V. Venkatesh, M. G. Morris, G. B. Davis and F. D. Davis, "User acceptance of information technology: Toward a unified view", *MIS quarterly*, pp. 425–478, 2003.

[92] V. Venkatesh, J. Y. Thong and X. Xu, "Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology", *MIS quarterly*, pp. 157–178, 2012.

[93] P.-P. Verbeek, "Materializing morality: Design ethics and technological mediation", *Science, Technology, & Human Values*, vol. 31, no. 3, pp. 361–380, 2006.

[94] W. N. Wan Ahmad and N. M. Ali, "A study on persuasive technologies: The relationship between user emotions, trust and persuasion.", *International Journal of Interactive Multimedia & Artificial Intelligence*, vol. 5, no. 1, 2018.

[95] M. Weinmann, C. Schneider and J. Vom Brocke, "Digital nudging", *Business & Information Systems Engineering*, vol. 58, no. 6, pp. 433–436, 2016.

[96] K. Wong, A. Wong, A. Yeung, W. Fan and S.-K. Tang, "Trust and privacy exploitation in online social networks", *IT professional*, vol. 16, no. 5, pp. 28–33, 2014.

[97] M. Wunsch, A. Millonig, S. Seer, K. Schechtner, A. Stibe and R. C. Chin, "Challenged to bike: Assessing the potential impact of gamified cycling initiatives", *evaluation*, vol. 23, p. 24, 2016.

[98]   F. Yetim, "A set of critical heuristics for value sensitive designers and users of persuasive systems", 2011.
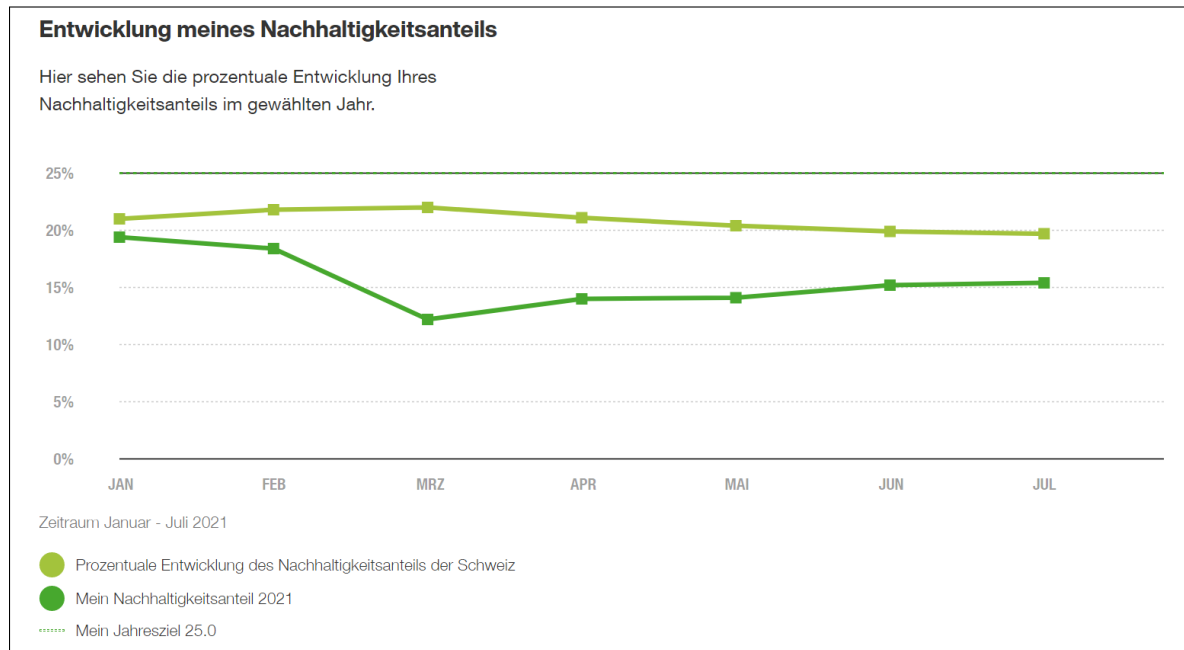
1 Appendix

Figure 9 – Development of sustainable purchasing behavior (taken from the Migros website [61]). The green lines show the development of the share of sustainable Migros purchases of the user compared to Switzerland's average.
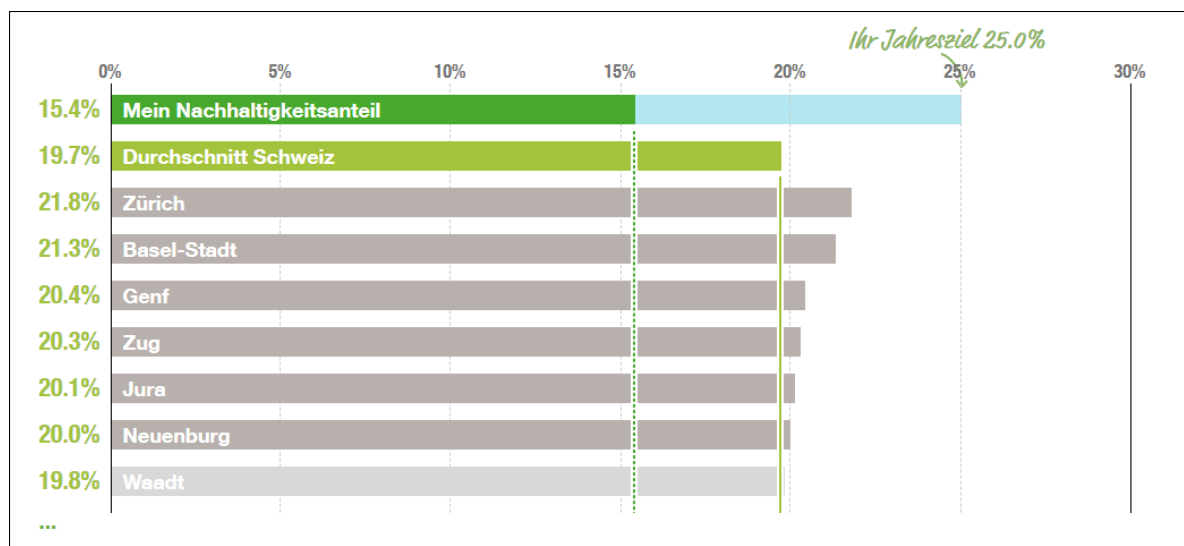

Figure 10 – Comparison of the share of purchasing sustainability (taken from the Migros website [61]). The green bars show the current share of sustainable Migros purchases of the user and Switzerland's average. The blue bar shows the target goal set by the user. The grey bars show the average of each canton.