ORIGINAL PAPER

Elicitation of situated values: need for tools to help stakeholders and designers to reflect and communicate

Alina Pommeranz · Christian Detweiler ·
Pascal Wiggers · Catholijn Jonker

Abstract Explicitly considering human values in the design process of socio-technical systems has become a responsibility of designers. It is, however, challenging to design for values because (1) relevant values must be identified and communicated between all stakeholders and designers and (2) stakeholders' values differ and trade-offs must be made. We focus on the first aspect, which requires elicitation of stakeholders' *situated values*, i.e. values relevant to a specific real life context. Available techniques to elicit knowledge and requirements from stakeholders lack in providing the context and means for reflection needed to elicit situated values as well as an explicit concept of value. In this paper we present our design of a tool to support active elicitation of stakeholders' values and communication between stakeholders and designers. We conducted an exploratory user study in which we compared the suitability of methods used in social sciences for (1) eliciting situated values, (2) supporting people's expressions of values and (3) being implemented in value elicitation tool. Based on the outcomes we propose a design for a value elicitation tool that consists of a mobile application used by stakeholders for data collection and in-situ self-reflection, and a website used collaboratively by designers and stakeholders to analyse and communicate values. Discussion focuses on contributions to value sensitive design.

A. Pommeranz (✉) · C. Detweiler · P. Wiggers · C. Jonker
Mekelweg 4, 2628 CD Delft, The Netherlands
e-mail: a.pommeranz@tudelft.nl

C. Detweiler
e-mail: c.a.detweiler@tudelft.nl

P. Wiggers
e-mail: p.wiggers@tudelft.nl

C. Jonker
e-mail: c.m.jonker@tudelft.nl

Introduction

Designing for (moral and non-moral) values has become increasingly important for technology development in recent years. In a multitude of technological systems, e.g. medical applications or social networks, human values (health, privacy, autonomy, trust etc.) play a role and are sometimes violated. Consider e.g. the case of medical image technologies used in diagnosis of diseases as discussed by Kraemer et al. (2011). It is practically impossible to design algorithms that can detect healthy or diseased cells with a 100% success rate. Existing algorithms produce false negatives (failing to identify a diseased cell as such) and false positives (identifying a healthy cell as being diseased). Both cases can have devastating consequences for patients depending on the circumstances (e.g. severity of the disease or the effect of treatment) and may be preferred in one context or the other. It is usually the designer's decision to set a threshold that favours either of the two cases. Designers with different sets of values would make different choices on which threshold to use in their algorithms. However, as argued by Kraemer and colleagues, which case may be favoured is based on values of the users. Ethical considerations are unavoidable when designing such systems and, more importantly, should be done together with the stakeholders of the system.

System designers are partly responsible for creating socio-technical systems accounting for human values. The value sensitive design (VSD) framework proposed by Friedman et al. (2006a) tries to guide designers in this process. The framework proposes conceptual, empirical and technical investigations to guide technology design to

better account for human values. These investigations are not to be seen in isolation but are iterative and integrated with each other. Whereas vsd supports ethical and philosophical deliberations on the design of new technology and offers practical guidance to designers, applying it to real life design cases can still be difficult. Among others, problems may include missing some values of importance, e.g. due to a strong focus on the values identified by the designer in conceptual investigations, lack of stakeholders' ability to envision new technology and its influence on their values, or misinterpretation/-communication between designers and stakeholders. An example from current work carried out by one of the authors of this paper reveals some of the difficulties. It deals with the development of an intelligent system supporting elderly in nursing homes, e.g. to detect whether a person has fallen to the ground and can not get up. As privacy is one of the identified values, a virtual model of the room has been created (instead of using real-time cameras images) that can be observed by caretakers or family members. However, the moral burden that is put on the stakeholders by allowing them to track the elderly person at all times had been neglected in the conceptual investigations. In addition, an elderly person might even be willing to trade in privacy for increased health, safety or life quality. These are value tensions that require active discovery of stakeholders' values, a vision of the consequences for stakeholders and a shared understanding of values between designer and stakeholders in the early design stages. Empirical investigations involving the stakeholders actively can overcome these issues. However, for vsd to be carried out well, there must not only be adequate conceptual investigations but also well developed methods carrying out empirical investigations into people's values. While existing HCI techniques may be applicable (see Friedman et al. 2006a, b, 2008a, b) more robust techniques and specific tools are needed to support the stakeholders in understanding the design context and consequences, expressing their values to the designer and arriving at a shared understanding of abstract values instantiated in the given contexts.

This brings up the question of how to elicit values from stakeholders and how to support the above. As acknowledged by (LeDantec et al. 2009), there is an "inherent difficulty in talking about values". Meanings, nuances and interactions of values are complicated to express in a simple ranking of abstract values or interviews and difficult to analyze by system designers not trained in (or aware of) value elicitation techniques (e.g. laddering, photo elicitation interviews).

We argue that two of the crucial aspects in value elicitation are (1) taking real life contexts into account and (2) supporting communication between stakeholders and designers. Expressing general value rankings is difficult for

stakeholders because even a single value can play different roles based on the situational context. Given the example above an elderly person might generally value privacy, but be willing to give up privacy in certain situations in order to increase her safety or health. Therefore, one cannot claim that privacy is the most important value in general. We will call values relating to a context, *situational values*. Furthermore, values are abstract concepts that can be interpreted differently by stakeholders and designers if the context is left out leading to miscommunications. Without a real life context (i.e. situations in which a value serves as a guiding principle for a decision or in which the violation of a value is apparent) elicited value profiles might be based on spontaneous thoughts of a person and by that be biased.

Our main research questions are "In what way do current requirement elicitation methods support the elicitation of situated values?" and "How can we support stakeholders and designers to arrive at a shared understanding of relevant values for the technology in question?". In the following, we first define what we mean by 'values' and point out deficiencies in current requirements elicitation and design methods that include value considerations. We show that standard knowledge gathering methods (e.g. interviews or surveys) lack in giving the real life context needed for people to understand and express their values. Next, we describe methods used in social sciences that support in-situ enquiry. We present an exploratory user study comparing some of these methods regarding their suitability to elicit situated values, to support people to express their values and the use in a value elicitation tool. Based on the results of the study we argue that in order to support elicitation of situational values the stakeholder has to self-reflect on how values guide her everyday actions and decisions. This will help a stakeholder to express her values to the designer and by that enhance the shared understanding of relevant values. As self-reflection on values is best done in-situ we provide ideas for an implementation of a value elicitation tool that relates to personal informatics systems and the experience sampling method used in human-computer interaction. The design of the tool consists of two parts (1) a mobile application supporting stakeholders' in-situ self-reflection and (2) a website providing functionality for a deeper reflection on underlying values as well as communication of the situational values between stakeholders and designers.

What are values?

Values form a central concept of our work. According to Pauls, before the term value came to common use in philosophy in the nineteenth century, "value phenomena were

discussed in terms of the good, the right, beauty, virtue, truth, obligation, moral judgement, aesthetic judgement, etc." (Pauls 1990, p. 7). Such concepts are about *what ought to be* as opposed to *what is* (cf. Hodges and Baron 1992). Values have been described as abstract (see Bardi and Schwartz 2003; Maio 2010) motivational constructs that apply across contexts and time (Bardi and Schwartz 2003). They convey what is good (see Miceli and Castelfranchi 1989; Schroeder 2008) and important to us (see Bardi and Schwartz 2003; Friedman et al. 2006a). By conveying what is good, values can also be said to lead to behavior that supports them, since, as Miceli and Castelfranchi state, "what is good should be pursued" (Miceli and Castelfranchi 1989, p. 189). Values serve as guiding principles in peoples' lives (Schwartz and Bilsky 1990).

We assert that values can have a universal nature—a sense that what is good for or what ought to be pursued by everyone. In this sense, values generate expectations regarding the behavior of others. If one values honesty, one behaves in an honest manner, and expects others to do so as well. However, even if one has a conception of something that ought to be desired by everyone, one still uses this conception to guide one's own actions in the specific situation or context one is in. So, values are made concrete in specific situations (Maio 2010).

Based on these characteristics, in this paper we take values to be conceptions of what is good and (hence) ought to be, which function as guiding principles in specific situations in people's lives. We refer to this concept of values-in-context as *situated values*. To illustrate this situated nature of values, consider an example from the study discussed later in this work. In the study, people were asked what they find important in their work. One might answer that it is important to be allowed to work at home. When asked why this is important, she might answer that this is important because it gives her the flexibiliy to decide for herself whether she works at home or travels to the office. Further questioning might reveal that it is important to her to be able to govern herself. This, in turn, is linked to the value of autonomy. While this person's desire to be able to work from home is in line with her value of autonomy, she does not directly discuss this desire in terms of autonomy. The abstract value of autonomy is made concrete in this specific work situation. Whereas this example shows how a single value is linked to a specific situation, values rarely exist in isolation. Instead, multiple values are at play simultaneously in the same situation, which can lead to value tensions. For example, choosing to use a search engine that personalizes search results, can lead to a tension between privacy (as information about oneself will be used by the personlization algorithm), easy access of knowledge (as the search results will be of one's interest and suited to one's knowledge) and freedom from bias (as the information will be filtered by another party). This complicated web of values poses additional difficulties to the design of technology accounting for several values and may require design trade-offs.

Examples of values frequently implicated in design include human welfare, privacy, trust, autonomy, accountability, identity, environmental sustainability (Friedman et al. 2006a). Further examples of values are achievement, benevolence, conformity, hedonism, power, security, self-direction, stimulation, tradition, and universalism (Schwartz and Bilsky 1990). None of these lists is (or can be) definite, as values are what a person considers important in life and that can be many things.

Background on requirements elicitation

Requirements engineering is one of the first steps in the larger process of software development. It is the process of identifying stakeholders and their needs, and documenting these in a form that can be analyzed, communicated, and subsequently implemented (Nuseibeh and Easterbrook 2000). Broadly speaking, there are two types of requirements: functional requirements and non-functional requirements (Chung et al. 1999). The former are requirements that define a function of the system, or something that a system will do. The latter define not what a system will do, but how it will do it. Requirements engineering has attention for "soft issues" such as politics and people's values, although dealing with soft issues is problematic as there is little guidance on how to do so (Thew and Sutcliffe 2008).

We consider traditional approaches to eliciting requirements. Our focus is on eliciting values that are relevant to the domain for which the designers are designing. Therefore, the approaches we consider need to be human-centered, have concepts similar to values, and must have some focus on context.

Knowledge acquisition in automated specification (KAOS)

KAOS is a goal-oriented requirements acquisition method consisting of a conceptual model for acquiring and structuring requirements models; a set of strategies for elaborating these requirements models; and an automated assistant to provide guidance in the acquisition process (Dardenne et al. 1993).

In KAOS, goals are identified in the first place by using interviews and analyzing documentation to find problematic issues with the current system, objectives that are stated for the future system, etc. (van Lamsweerde et al. 1998). Van Lamsweerde also mentions obstacle

analysis (van Lamsweerde and Letier 1998), scenario based elicitation (van Lamsweerde and Willemet 1998), and analogical reuse of goal structures (Massonet and van Lamsweerde 1997) as possible techniques for goal identification. Further goals can be identified by a combination of top-down and bottom-up approaches. Offspring goals or sub goals can be identified by asking HOW questions about previously identified goals (i.e. how are identified goals to be achieved), whereas parent goals can be identified by asking WHY questions about goals and operational requirements previously identified (van Lamsweerde et al. 1998).

The concept of goal as used in the KAOS method is similar to the concept of values we use in the sense that goals "lead to the incorporation of requirements components which should support them." (Dardenne et al. 1993, p. 4). This is similar to the idea that what is good ought to be pursued. However, no explicit claims are made regarding the goodness of goals in KAOS. Furthermore, it is not clear in KAOS specification what (or who) the source of these goals is. Goals just are, and so should be achieved. A problem with this is that, since KAOS specifications as such do not indicate how goals relate to stakeholders in the domain, it is unclear who is affected by certain goals not being achieved and how they are affected.

As for context, as pointed out by Kolos-Mazury and colleagues, KAOS' main concepts allow for adequate modelling of the task context of a system (i.e. what the user is doing), but does not provide mechanisms for other types of context including personal (user state) and social (social aspects of the user context) (Kolos-Mazuryk et al. 2006; Krogstie 2001).

Non-functional requirement (NFR) framework

Cycneiros and do Prado Leite (2004) describe an approach to non-functional requirement (NFR) elicitation using the Language Extended Lexicon LEL . The LEL is used to capture the specific language used in the domain for which a system is being developed, or the vocabulary of the *Universe of Discourse*. A Universe of Discourse is defined as "the general context where the software should be developed and operated. The Universe of Discourse includes all the sources of information and all known people related to the software. These people are also known as the actors in this Universe of Discourse".

The LEL is a controlled vocabulary "based on a code system composed of symbols where each symbol is an entry expressed in terms of notions and behavioural responses" (Cysneiros and do Prado Leite 2004, p. 331). Each meaningful term in the Universe of Discourse is captured in such an LEL symbol and must contain at least one notion and one behavioral response. The authors propose gathering LEL symbols by reading documents used in

the domain. Notions and behavioral responses for these symbols should be elicited from stakeholders and users by interviewing the stakeholders and users or sending them questionnaires. The authors do not explicitly mention how to identify stakeholders, although they do suggest adding NFRs desired by customers to the LEL, suggesting that the primary group of stakeholders they consider is customers (of the software under development). They also consider the developers themselves (see below).

To elicit non-functional requirements, the authors propose starting with an existing LEL, or, if one does not yet exist, building one. Then, the requirements engineer is to add NFRs to the LEL by using an NFR catalogue or knowledge base, and for each symbol in the LEL asking stakeholders (customers and the developers themselves) whether each of the NFRs in the knowledge base applies to it, or which NFRs would have to be achieved to be able to consider the symbol fully represented.

Though the authors only discuss direct stakeholders (customers) as a source of NFRs, there does not seem to be any reason why indirect stakeholders [those parties affected by the system who are not direct users (Friedman and Kahn 2003)] could not also be approached as a source of NFRs. These NFRs could be added to the LEL just as those gathered from direct stakeholders.

This approach comes close to our aims by engaging with stakeholders directly and attempting to capture the vocabulary used in the context of use of the envisioned system. However, elicitation is still limited to analysis of domain documentation and stakeholder interviews or questionnaires. Also, there is no explicit mention on how to identify (direct) stakeholders, nor is there any mention of how to identify and involve indirect stakeholders. Even if we assume that indirect stakeholders have already been identified, the method provides no obvious way to distinguish between NFRs of direct stakeholders and those gathered from indirect stakeholders. Furthermore, the approach makes no explicit mention of values or concepts of what is good or ought to be.

Tropos

The Tropos software development methodology supports the agent-oriented paradigm and the associated concepts of actors, plans and goals throughout the software development process (consisting of early requirements, late requirements, architectural design, detailed design, and implementation) (Bresciani et al. 2004; DeLoach et al. 2009; Giunchiglia et al. 2002).

The early requirements phase in Tropos is most relevant to our aims, as it is in this phase that stakeholders and their goals are identified. Morandini and colleagues suggest beginning this phase by asking questions such as "who are the stakeholders in this domain?" and "what are their goals

and how are they related to each other?" (Morandini et al. 2008). The Tropos methodology does not, to our knowledge, offer explicit steps to answer the question of who the stakeholders in the domain are. They make a distinction between hard and soft-goals. Whereas hard-goals can be said to be satisfied, soft-goals have no clear-cut criteria for satisfaction, and as such are used to capture non-functional requirements (see also the NFR framework above). These goals are later refined into subgoals and eventually delegated to the system-to-be.

In light of our aims, one issue with Tropos is that little guidance is provided on how to elicit goals from stakeholders. Also, by modeling social actors and the goal dependencies between them and subsequently considering how these can be delegated to the system-to-be, Tropos does not explicitly consider indirect stakeholders.

Beside these issues, Tropos' concept of soft-goals seems like a useful candidate for representing values. However, values are not the same thing as goals. Miceli and Castelfranchi provide a useful distinction between these concepts. "Values are not goals, they are assumptions (more precisely, evaluations). A value is a judgment, though very general and vague. It says of something that it is good or bad. A goal is a regulatory state in someone's mind" (Miceli and Castelfranchi 1989, p. 179). They illustrate a further important feature of values in discussing the difference between values and norms: "Values in fact offer grounds for, or give rise to norms. Hence the 'normative' facet of values: If something is good, it should be pursued" (Miceli and Castelfranchi 1989, p. 181). If we represent values as soft-goals, the evaluative aspect ("X is good") and the normative aspect ("X should be pursued") are lost. Represented as a soft-goal, a value becomes something that can be satisfied (i.e. sufficiently satisfied). Not achieving a goal is not morally wrong as such. Violating a value, on the other hand, can be seen as morally wrong. Not taking these aspects into account could lead to problems once the design has been implemented and put into practice.

Finally, the Tropos method goes some way in capturing context, especially task and social types of context. It does not, however, seem to capture personal context.

ScenIC

The ScenIC requirements engineering method is based on the idea that requirements documentation or models form the memory of a software development project, analogous to human memory (Potts 1999). Three types of memory schema are used in ScenIC, semantic memory, episodic memory, and working memory. Semantic memory contains information about the system. This type of memory employs a number of entities that are to an extent human-centered and related to values. These are actors, goals, and obstacles.

In ScenIC, actors are defined as "entities that participate in changes of state" (Potts 1999, p. 3). These can include user roles and organizations (as external actors). Potts offers some guidelines for identifying actors. External actors (our main interest here) are those who perform actions or have responsibilities relevant to the system's purpose or interact directly with the system, or people, teams, organizations, devices and systems, and elements of the environment.

There are two kinds of goals in ScenIC. Objectives are expressed as a "trajectory of improvement" or "the preservation or prevention of states of affairs". Tasks, on the other hand, are "stated in terms of achievement of a state or performance of an action". Goals may be hindered by obstacles (Potts 1999).

"Goals are identified by reflecting on the system's purpose, interviewing stakeholders, or inferring goals from background documentation. Goals may also be identified from long-term memory content. Scenarios, for example, may illuminate previously ignored or underemphasized goals" (Potts 1999, pp. 3–4).

Potts offers some guidelines for identifying goals. To answer the question of what the goals are for the system, he suggests obtaining goals from mission statements, questions to stakeholders and the like. He further provides a lexicon of verbs that is useful for identifying objectives and tasks. For example, an objective of goal achievement might be improving a condition, with which the verbs "improve", "reduce" and "maximize", among others, are associated.

The concept of goals in ScenIC is somewhat similar to the concept of values. The objects of achieving goals are improving, maintaining, avoiding or satisfying a condition; bringing about a state; bringing about knowledge; bringing about commitment; providing knowledge; or soliciting knowledge. However, as we discussed above with regard to Tropos, values are not goals (at least not in the sense used here).

Scenario-based requirements analysis method (SCRAM)

In SCRAM, scenarios are used in combination with prototypes to elicit requirements (Sutcliffe 2003). Scenarios are defined as "facts describing an existing system and its environment including the behavior of agents and sufficient context information to allow discovery and validation of system requirements" (Sutcliffe 1998, p. 49).

SCRAM consists of four phases. In *initial requirements capture and domain familiarization*, conventional interviewing and fact-finding techniques are used to gather enough information to develop a first concept demonstrator. Subsequently, in *storyboarding and design visioning*, early visions of the intended system are explained to users in storyboard walkthroughs to get feedback on feasibility.

Then, during *requirements exploration*, concept demonstrators and early prototypes are used to present the intended system in a scenario-driven, semi-interactive way, in order to let them critique design and validate requirements. Finally, in *prototyping and requirements validation*, more functional prototypes are developed and requirements are further refined until all users agree a prototype is acceptable (Sutcliffe 2003).

As with the other methods described above, in SCRAM requirements are elicited through conventional interviewing and fact-finding techniques. The added benefit of SCRAM is that initial requirements are used to develop an initial concept demonstrator, which users are exposed to using scenarios. Sutcliffe suggests gathering scenarios of normal system use and looking for commonalities between them to come up with what he calls a "normal use case" (Sutcliffe 2003). Exceptions to this normal use case can then be gathered. The use of scenarios should provide a richer understanding of the context of use, bringing this approach somewhat closer to our aims. However, there is no explicit focus on values.

Value sensitive design (VSD)

The Value Sensitive Design (VSD) approach aims to incorporate knowledge of the ethical impact of a technology into the design process. VSD "is a theoretically grounded approach to the design of technology that accounts for human values in a principled and comprehensive manner throughout the design process" (Friedman and Kahn 2003). To that end, it provides an iterative three-part methodology consisting of conceptual investigations, empirical investigations, and technical investigations.

In VSD, conceptual investigation focuses on "values discovery and informed analyses of these values and potential value tensions" (Miller et al. 2007). Friedman and colleagues suggest beginning with either a value, technology, or context of use, depending on which is most central to the work at hand (Friedman and Kahn 2003). Inspecting the selected item should reveal some value issues surrounding it. The next step is to perform a stakeholder analysis to identify direct and indirect stakeholders, which are the people who interact directly with the technology and those who are impacted by the technology without interacting with it, respectively. For each group of stakeholders, harms and benefits are then identified. The list of harms and benefits can then be mapped onto associated values, especially human values with ethical import. Once these key values have been identified, a conceptual investigation of the values is conducted supported by (philosophical ontological) literature, resulting in working definitions of those values. Potential value tensions, which can constrain the design space, are then examined.

Several of the analyses of the conceptual investigation phase need to be informed by empirical investigations of the technology's context, and to evaluate particular designs. VSD does not prescribe a specific method for this stage, stating that "the entire range of quantitative and qualitative methods used in social science research is potentially applicable" (Friedman et al. 2006a). Asking about values should be done both directly and indirectly, based on the formal criteria specified in the philosophically informed conceptual investigation stage. This conceptualization is needed to define the boundaries of a topic.

Technical investigations then focus on the properties and mechanisms of existing technologies that support or hinder human values. Alternatively, technical investigation can consist of designing a system to support identified human values. Though technical investigations of the first form and empirical investigations seem similar, technical investigations focus on the technology itself, and not on the individuals affected by it, as empirical investigation does. During this stage, it can be helpful to make explicit how design trade-offs map onto value tensions and affect different groups of stakeholders.

While VSD does provide us with the orientation on values we seek, one may raise the concern [similar to issues related to value discovery mentioned by (LeDantec et al. 2009)] that early and rigidly articulated conceptual investigations may predispose researchers from hearing other value considerations arising during empirical investigations. Mitigating this risk could be done by a stronger focus on understanding stakeholders' values in the context in which they are relevant and how they guide judgment in specific contexts. The Value Scenarios approach by Nathan et al. (2007, 2008) provides focus on context, exploring visions of the intended system in various contexts of use and examining the value implications thereof. However, the approach provides little guidance as to how values should be elicited and runs the risk of privileging known values over value discovery. Several studies in the VSD literature have dealt with this issue [for example, (Denning et al. 2010) on usable security for implantable cardiac medical devices; (Miller et al. 2007) on groupware; (Friedman et al. 2008a, b) on longer-term deployment of large displays; (Friedman et al. 2006b, 2008a, b) on privacy in public with webcams in the United Sates and in Sweden, respectively]. However, more work is needed to support stakeholders in reflecting on their own values and creating a tool(kit) that designers can use together with their stakeholders to elicit the relevant values. Especially practitioners would profit from such a tool(kit). We need to provide the designer with robust tools for in-situ value elicitation and creation of shared understanding between designers and stakeholders. Supporting communication between stakeholders and designers is key to successful

technical design and is until now not well addressed within the VSD literature.

Methods used in social sciences and social psychology

As we have seen in the previous section current requirements engineering methods, although some focus on concepts similar to values exist, lack specific methods for elicitation of situated values. Document reviewing, stakeholder interviews and questionnaires are the most used techniques to get information about soft issues and stakeholder goals. These methods provide little context and focus on value discovery. VSD is the only approach that explicitly focuses on values. In the empirical analysis VSD advocates the use of a range of methods used in social sciences. Two questions arise in this context (1) which of these methods are most suited for elicitation of situated values and (2) can we assume that designers or software developers are aware of these methods and able to use them? The latter question is not investigated in this paper by means of an empirical study. However, we assume that the answer is no, which is one of the reasons for developing a value elicitation tool in the first place. In the following we focus on the first question.

We started by investigating the work of social psychologists dealing with understanding values and their interplay with other aspects of human nature and life. Social psychology research resulted in several value inventories and measurement instruments. The Rokeach Value Survey (RVS) (Rokeach 1973) (asking participants to rank 36 values in order of importance), Schwartz Value Survey (SVS) and the Portrait Value Questionnaire (PVQ) (Schwartz and Bilsky 1990) are referred to most often; for others see Cheng and Fleischmann (2010). A majority of inventories and questionnaires ask people to rank or rate values. Lists of values are presented to people out of any real life context. An exception is the PVQ which offers personal descriptions, such as "He likes surprises. It is important to him to have an exciting life", and asks people to rate how similar a described person is to themselves. Whereas this cannot be considered as real life context, it gives details on how values relate to a person's life and personality. In addition, to obtain an importance rating for a value it is not explicitly named but measured with several statements.

From the vast range of social science methods to gather data about human behavior (interviews, surveys, archival research, observations, case studies etc.) several seem to be suited to capture information about people's situated values. In a number of studies on the use of a plasma display window with different foci several in-situ methods were used to get an account of participants' experiences. For instance, interviews in-situ (Friedman et al. 2006b, 2008a, b), surveys in-situ (Friedman et al. 2008a, b), physiological measurements in-situ (Kahn et al. 2008) and diaries (Friedman et al. 2008a, b). While some of these studies (Friedman et al. 2008a, b; Kahn et al. 2008) investigated the overall (verbally expressed and affective) experiences of single users with the plasma display as an office window, including e.g. the value of welfare, others (Friedman et al. 2006b, 2008a, b) were more directly enquiring about personal values such as privacy, welfare and informed consent. As the studies show these in-situ methods clearly provide context and can trigger value deliberations relative to the context of use for an existing technology, which may differ from hypothetical situations (Friedman et al. 2006b). A comparison of current methods using different modalities (pictures, words, physiological reactions etc.) as well as extending these methods by providing more specific support in expressing and sharing value deliberation with designers would add to this body of research.

Ethnography is another method to gain detailed insights in people's life by observation and interviewing informants in a social group. Focus is put on the researcher participating in the group over a longer period of time to gather personal experiences. This is costly and the presence of the researcher/designer can influence the people's behavior. Diary studies are ways to understand people's behavior without being present to observe the person.

Another method used for studying important aspects in people's lives from their point of view is the photo elicitation interview (Harper 2002). In this technique photographs are discussed in an unstructured interview. When comparing photo elicitation to ethnographic field studies, Clark-Ibez (2004) mentioned that photo elicitation can illuminate insights that would stay hidden to the researcher otherwise. Photo elicitation has been used in two ways, discussing pictures that the researcher selected or asking participants to take their own. When participants take their own pictures and discuss them in an interview with the researcher, aspects can arise that the researcher might not have anticipated beforehand.

Photo elicitation seems well suited for elicitation of situated values and has already been employed in human-computer interaction to discover user values in the design phase of new technology (LeDantec et al. 2009). Compared to diary studies, the use of photographs evokes feelings and memories, which can give clues to the importance of certain values situated in real life. In the following we describe how we compared different methods in a first exploratory user study.

Exploratory study of value elicitation

In a first study we explored different ways for people to express personal values. We set out to answer the following

questions; (1) whether the values elicited with different methods are comparable (2) how well different methods can be used for elicitation of situated values, and (3) which methods can be used as a basis for a value elicitation tool. We compared (1) a questionnaire instrument for value elicitation, (2) values elicited by researchers in photo elicitation interviews, and (3) value tags people used for their own pictures. We also recorded affective reactions (skin conductance measures) occuring during the interview to test the utility of such measures in the value elicitation process. However, in the following we focus on the questionnaire results, tags, and interview data. As context is a major aspect in value elicitation, we provided people with a context, i.e. we asked to reflect on values relating to their work environments. This choice was motivated by the focus of the project (decision support for job contract negotiations) in which this work is carried out.

The methodology for this study was greatly guided by the photo elicitation method often employed in social science research. We draw on Harper's work defining photo elicitation as an interview method that not only elicits more (verbal) information than a common interview but also different types of information as visual and verbal processing by the brain takes place simultaneously. "Thus images evoke deeper elements of human consciousness than do words" (Harper 2002). In the light of eliciting values that are (maybe even subconscious) drivers of human behavior this characteristic of photo elicitation seems well suited for value elicitation. However, many photo elicitation interviews use photographs chosen or taken by the researcher in order to enquire certain aspects of the interviewees life. Values and work situations in which they are important are, however, subjective to our participants and can hardly be foreseen by the researcher. Therefore, we adopted the method used in studies by Clark-Ibez (2004) in which the participants were asked to take their own pictures during the study. Given the short time frame we had for conducting the study we could not assure that the participant found enough opportunities to take pictures of every important aspect of their work life. Therefore, we allowed people to add pictures they had taken earlier (which they indeed did). Furthermore, as suggested by Harper (2002), "in addition to deconstructing the portrayed events and scenes," we included the question "[...] what has been left out of this photo sequence?". Using the photo elicitation method allowed us not only to investigate verbal accounts of people's values but also the utility of value tags for pictures. Furthermore, we included a value questionnaire in order to compare the rather time extensive methods of taking pictures, interviewing and tagging to a more efficient one. We chose the Portrait Value Questionnaire (PVQ) as it is a valid, often used method in social psychology and the statements used in the questionnaire provide at least some sort of context (in the form of descriptions of

people) compared to many other questionnaires based solely on ranking or rating abstract values. In the following we describe the study set-up as well as the results from the comparison of the three methods.

Set-up

We asked participants to take at least 10 pictures (work and non-work related) reflecting what they consider important in their everyday lives in the course of 1 week. We did not limit the maximum number of pictures in order not to bias the participants to make a choice of a subset of important values. We determined, however, a minimum number of images in a pilot interview (by comparing the number of pictures to the length of the interview) in order to trigger enough verbal expressions of values. Although we wanted to elicit values related to the work environment, we assumed that some values that may influence the person's work were better elicited in the participant's freetime (such as having freedom of chosing workhours may relate to family values, e.g. spending time with children). Therefore, we asked participants to take pictures at work and in their freetime as formulated in the task description given to the participants:

Task: Take pictures of important aspects in your life. *Please take as many photos as you consider necessary of events, persons or objects in your daily life that you consider important. (This can be anything from a coffee break, a meeting, the coffee mug on your desk to a visit of your grandma, a romantic night out or playing with your kids).We would like you to take a balanced amount of photos at your work and in your free-time. To be able to get the most out of the interview in the end, please take at least 5 photos at work and 5 at home.*

We provided an online tool for people to upload and value-tag their pictures. We did not instruct participants to use specific pre-defined values tags and we did not give a definition of values to avoid influencing participants' interpretion of what values are. After 1 week participants could make a selection of pictures for the upcoming interview. This step was described in the intitial instructions and was chosen to allow participants to take intimate pictures during the study but leave them out of the interview. In this way participants did not have to limit themselves while deciding what moments are important during their daily life.

A day before the interview we asked participants to fill-in the PVQ. We chose this timing carefully in order not to influence the picture taking excersise (by showing the questionnaire first) or the anwers to the questionnaire (by giving it after the interview where people may focus only on situations discussed in the interview).

In the interview each participant was asked to go through the pictures one by one and explain why he or she took a particular picture. The researcher interrupted the participant as little as possible. We filmed the participant for a later analysis.

Participants

Six people (3 female, 3 male) participated in this exploratory study. We ensured that we interviewed people with different work and family backgrounds and in different life phases. Although due to the small sample we did not expect to find patterns of value expressions based on a person's life phase, we still believe that having such diversity would aid the exploration. At least we expected a diversity of values due to different foci in participants' lives related to age and family status. We had two people (aged 29 and 24) in relationships working 40 h per week, one as an IT consultant in a big international company (P1) and one as a communications manager in a research company (P2). One participant (aged 53) worked between 40 and 50 h per week as a senior manager in a major consulting company (P3). This participant was married and had two adult children. Two participants (aged 36 and 37) worked half-time (24 h per week) at a university (P4 and P5). Both were married and had small children. The last participant (aged 44) had recently started his own business on which he spent at least 40 working hours per week (P6). His wife helped with the business. They did not have any children.

Data analysis and results

The study provided data of different types. The PVQ results in ratings of a preset number of values. As we did not ask people to reflect on the same (or any) specific values during the tagging and the interview a comparison of elicited values from the three methods requires a translation step. We had to match the labels people used in their tags and the interviews to the abstract values that the PVQ elicits. As an example, we related labels like fun, enjoyment, happiness to hedonism, and creativity, freedom, work-life balance to self-direction. Next, we compared per participant whether the values elicited with the PVQ were reflected in the tags and the interview. In the following subsections we briefly describe results from each method and then compare the elicited values. Furthermore, we mention the participants' comments on the methods and finally, our conclusions.

Portrait value questionnaire

The PVQ delivered importance ratings per participant for each of the 10 abstract values achievement, benevolence, conformity, hedonism, power, security, self-direction, stimulation, tradition, and universalism. In Table 1 we included the values rated on average with "somewhat like me", "like me" and "very much like me" in order of their importance per participant. Notably when comparing the number of identified values per person is that while some participants, e.g. P4 and P5, rated only a very small number of values (2–3) with at least "somewhat like me", others, in specific P6, rated the majority of values (7) in this way. This could mean that some people could easily identify themselves with the items (descriptions of people) provided by the PVQ while others could only find themselves in very few of the descriptions. In the latter case using the PVQ may not be very successful to elicit all values important to those people. Due to the nature of the questionnaire there is no information in how far the elicited values relate to specific situations in a given context, e.g. in the participant's work life.

Picture tags

Considering the tags people used to describe their pictures, some people were able to use single word tags such as fun, proud, relaxing (see P3) that reflect values, whereas other people (see P6) used full sentences to describe the pictures. In order to compare the values elicited from the different methods we extracted the value-related expressions from the tags (see Table 1) and later mapped them onto the abstract values of the PVQ (see Sect. "Comparison of values elicited with different methods"). In Table 1 we indicate how often each value was used by a participant (numbers in brackets). A higher number of uses could indicate the importance of a value to a participant, but may simply be dependent on how often situations occured during the study in which the value was of importance. At least in case of P3, however, it seems to be the former reason, as he added a picture with the label fun that was not taken during the week of study (although fun is also represented in other pictures). When looking at the type of values elicited in this way we notice that some of them are rather abstract concepts such as accomplishment, fun, health, challenge or self-direction, whereas others are more concrete and specific for the work environment, such as team spirit, quality time with family, work-life balance, work atmosphere etc.

Interview: verbal expressions

During the interview people described different situations that either themselves or the researcher could relate to a higher level value. The interview videos were viewed and statements were extracted that reflected what a person considered important. This included statements explicitly naming a value (e.g. "It is nice to have that freedom", P1) or implicitly referring to a value ("My parents are

Table 1 Comparison of values elicited using different techniques

Participant	Elicited values		
	PVQ results	Interview results (values elicited by researcher)	Tags used by participant
P1	Hedonism	Fun, enjoyment, excitement	Responsibility (2x), fun(2x),accomplishment, frustration about commute (2x), pride, happy, relaxation, calm, productive, love, happiness
	Achievement	Nice colleagues and interesting work (job satisfaction)	
	Universalism	Getting work done, perfection	
	Self-direction	Freedom, flexibility (working from home)	
		Security	
P2	Self-direction	Community/ being social (friends, family, colleagues),	friends(5), parties(2x), healthy, nice working atmosphere, good company (e.g. pets), freedom, self-sufficiency, space, relaxation(2x), self-organization, being creative, learning/study (5x), challenge, meeting people(2x), experience
	Stimulation	Learning/exploring new things (cultures, hobbies, people),	
	Achievement	Independence/freedom (at work and at home), relaxed work atmosphere	
	Hedonism/ universalism/ benevolence		
P3	Hedonism	Fun, enjoyment, team spirit, balance between hard work and fun	fun(3x), results (2x), good discussions, organize, relaxing, proud (2x), recognition, team spirit, quality time, being available, healthy, set goals, rest, moment for yourself
	Self-direction/ benevolence,	Results, recognition, achievement (at work), being proud (work and family),	
	Universalism	Organisation, efficiency,	
		Relaxation, health,	
		Support of others (family and work)	
P4	Self-direction	Family' safety, most important	Safety (2x), love(2x), security, creativity, job satisfaction
	Hedonism/ security	Being social (sporting in big group)	
		Good work atmosphere, social moments, nice colleagues	
		Work-life balance, freedom	
		Proud	
P5	Self-direction	Team work, good atmosphere, fun helping each other	Good teamwork, nice feeling, good work atmosphere, having facilities close by, practical, close relationship to parents, happy, hobby
	Hedonism	Thinking positive (negativity costs energy)	
		Organisation, practical, efficiency	
		Creativity ' perfection	
		Health, challenge (sport)	
P6	Self-direction	Challenge (keep challenging oneself)	Family/relatives, work, hobby, free , challenges
	Stimulation	Self-direction (being able to choose projects that fit own vision)	
	Universalism/ benevolence	Continuous stimulation	
	Power	Achievement (proud about website)	
	Power	Work-life balance: fun, passion, excitement both in private life and work, quality time with family, being social	
	Hedonism		

important to me. I can rely on them." e.g. reflecting trust or safety, P4). The statements were compared to find values that were mentioned most often by the participants. Based on the frequency of naming a value, the context in which it was mentioned (either something positive or negative that is described) and the visible emotional reactions (smiling, frowning) we compiled a list of the most important values (from the view of the researcher).

In the following, we describe concrete examples of how people expressed values. For P1 being able to have flexible work hours and being allowed to work from home meant 'freedom'. Whereas P1 used the word freedom during the interview, P2 mentioned similar aspects of her life that she enjoyed without naming the word freedom, e.g. "Finally getting my own place without any roommates that is really something important." An interesting example of an instantiation of 'relaxation' was given by P3 who took two pictures, one at an airport terminal and one in his car in a traffic jam. Whereas few people would spontatneously connect a traffic jam to relaxation, for P3 being alone in the car and listening to music or watching the surrounding was a welcome relaxation moment compared to his otherwise busy worklife, e.g. "if you are on time it [being in a traffic jam] is relaxed. You have to enjoy that as well." or "I like to see the nice things of the nature. [...] even when I am driving to work, looking to the skies, looking to the area, [...], always trying to enjoy." P6 mentioned that, despite being in need for clients for his new business, he refused clients because what they asked for did not fit well with the philosophy and vision that P6 had for his company. P6 (translated from Dutch) about having his own company: "[I have it] very conciously. I do not go for the money. It has to do with the combination. I have to like the assignment. [...] I refused two contracts because they did not fit with my philosophy." This can be seen as an instantiation of 'self-direction'.

During the interviews all value expressions immediately derive from a concrete situation described by the participants based on a represented photograph. In this way abstract value concepts such as freedom can be described in detail, showing how they play out in concrete situations, e.g. in case of P2 having her own apartment, or being able to have a drink at work.

Comparison of values elicited with different methods

Table 1 shows the values elicited by the PVQ instrument (based on a mean calculation for each value construct ($M \geq 3$, 3 = "somewhat like me"), by the researcher based on the interviews (reflecting the main themes the participants discussed) and people's own tags. In the last column not all tags are shown. Some participants used decriptive texts that did not include a value-related expression. These cases are excluded and we only extracted value-related words to show in the table. Due to this some participants mentioned less than one value tag per image as one would expect from the study set-up.

In the comparison of the elicited values from the three sources we matched the labels people used in their tags and the interviews to the abstract values that the PVQ elicits. These links were mainly based on the definitions of values given in Schwartz' theory underlying the PVQ. As an example, we related labels like fun, enjoyment, happiness to hedonism, and creativity, freedom, work-life balance to self-direction. Next, we compared per participant whether the values elicited with the PVQ were reflected in the tags and the interview. Overall, we found that a majority of values were indeed reflected in all three methods. A notable exception was the value "universalism" which was elicited by the PVQ in four cases but was not reflected in people's accounts of important aspects in their life. Only P3 hinted at universalism, when talking about his tolerance towards his children. Besides the overlap, there are, however, differences in all six cases. Especially in case of P3 and P5 the interview and also the tagging revealed several values that seemed of great importance for the participant but were not highly ranked by the PVQ. P3 mentioned his orientation towards results, recognition, being proud as well as organized and efficient. These aspects clearly relate to achievement, which was not an important value resulting from the PVQ. For P5 several values mentioned in our study, i.e. organization, efficiency, health, challenge, were not revealed by the PVQ. In this case we can tentatively conclude that the PVQ did not work well for the participant, as shown by the small number of values (self-direction, hedonism) resulting in higher than average ratings. A closer look at the internal consistency of the PVQ items for this participant may give deeper insights in this case. On the opposite, in case of P6 the fourth highest ranked value from the PVQ was power which was not at all reflected in the tagging or the interview. Noticable in this case is also that P6 rated 7 out of 10 values at least "somewhat like me", which is a high number compared to the other participants. When tagging his pictures, however, this participant did not use any value related words, but merely descriptions of the picture content. Unless the participant misunderstood the tagging task, this could mean that the PVQ worked well for the participant, but tagging was more difficult for him. The success of the PVQ could also hint at a difference between what he believed to be important (revealed by the PVQ) compared to what seems to be important in his daily life (shown by the pictures).

A general observation we made when analysing the interview data was that it was easy for people to express specific aspects about their work that they found important. Linking such aspects to specific values that they hold, however, was less natural and even impossible for some participants (see paragraph Participants' comments for details). Several participants mentioned social aspects, e.g. team work, nice colleagues, or having fun as one of the important things related to their work without being able to name a specific value. In addition, especially P5 elaborated on the behavior of people around her, which

either impressed or annoyed her. Such narratives relate to values because they show what the person perceives as right or wrong behavior and how people ought to act. Also in this case, the participant did not show any reflection on her own related values. Comparing the tags and the interviews we noticed that in few cases the two methods could be seen as repeated measures. Considering, e.g. the cases where participants were able to find value-tags (mostly single words), P2 and P3 mentioned the same values in the interview, P1 did so for some tags and P4 did not do so at all (her tags were much more to the point than her verbal expressions).

With regard to situatedness, the photo elicitation interview clearly delivers the most detailed account of concrete instantiations of abstract values. The tagging method can also provide a link between a situation and an abstract value, in cases where the participants are able to tag their pictures with a value-related word and the picture clearly reveals the link (tagging a party picture with "fun"). This was not always the case in the study. Some participants did not find value related tags and in some cases the link was not obvious (a laptop on a dinner table tagged with "calm"). The PVQ is generally less suited to elicit situated values, as the link to concrete situations is missing.

Participant's comments Except participant P1 and P6 everyone mentioned that filling in the PVQ was difficult. P2 said that she often felt that the first sentence of the described person fit her well, but when she read on, the person was not similar to her. Generally, people liked taking pictures in their daily lives. P1 thought it was difficult, however, to take pictures of reoccurring every-day situations that are connected to values. Despite this, P1 did not seem to have difficulties talking about the pictures. Although some of his pictures showed rather everyday things, they triggered long elaborations of specific situations important to him. This was the opposite for P2, who took many pictures showing several aspects of her life, but kept the descriptions during the interview rather short. P3 enjoyed taking the pictures and even more talking about important aspects in his life. He mentioned trying to tag the photos with one value-related word was often difficult. The same difficulty was reported by P4. Both succeeded after some time of reflection to put value related tags. In case of P3 the tags reflected well what he focussed on during the interview, whereas in case of P4 some aspects discussed were missing in the tags. P5 and P6 circumvented the tagging difficulty by writing full sentences for each picture. P5 said that she liked that the tasks (taking pictures, answering questionnaire, interview) were easy to follow and, therefore, enjoyed taking part.

Conclusions

The comparison above shows that while the questionnaire gives rankings over abstract values the photo elicitation method provided more descriptive values (being social, work-life balance etc.). Pictures used in the interview provided a good way to link the values to specific situations. In addition, some values were mentioned several times by the same person, but relating to different aspects of life (P3 talking about his own pride and his son being proud). Besides the context-relation of values elicited using photographs we noticed that the tags people used as well as the values the researchers extracted from the interview could be related to the abstract values of the PVQ and matched to a large extent.

Besides this overlap, the comments of the participants showed that there are clearly individual differences between people in how they can express their values and which methods support this process. Based on this we believe that triangulation is an important aspect when designing a tool that helps people to reflect on and express their values. By that we mean that we have to collect different types of data and allow people to analyse the collected data from different perspectives. This could, e.g. mean that a visually oriented person can more easily take pictures related to her values. A person that can express herself better in words could use a storytelling approach.

Overall, we noticed that although participants did not have problems with talking about the pictures and important moments of their lives, few people reflected on how values actually influence their behavior. Furthermore, details about importance of values and possible trade-offs were hard to obtain. Few participants mentioned trade-offs explicitly. P1 mentioned changing to a job that was located less practically but was more fun: "At my previous job I had like 20 min to work [...]. I guess that is a sacrifice you have to make in order to, I mean I wanted a job that I would really enjoy". Such scenarios would give more insights into value rankings and relations, but require deeper reflection. Based on these results we identify self-reflection as a main step in getting to know one's own values which in turn is necessary to be able to communicate them to designers. This self-reflection is easier for some people, but not as natural for others. Therefore, guidance is needed in order to assess meaningful situations in life in terms of values. As described later counselors who give this guidance to their patients/clients may be a good source for knowledge in how to support self-reflection. Summarising the results, we can say that the differences we found between the few participants in this exploration can already help defining what a value elicitation tool(kit) should offer, namely guidance for self-reflection in different ways supporting the users' possibilities to express values.

Towards a mobile tool for in-situ value reflection and elicitation

Motivation

The observations from our exploratory study that people prefer different ways of expressing their values augmented by different methods and reflect to different extents, signals a need for a tool that supports these differences and guides the user in explicit value deliberations. Similar differences of value expressions (e.g. logical, linguistic, interpersonal or bodily expression) were found by Voida and Mynatt (2005) in their study of conveying family values to designers. Based on initial findings that values are situated and given the results of the study in the previous section we developed ideas for a (mobile) tool for support of in-situ self-reflection on and communication of values. As the goal is to advance VSD by adding means to actively elicit values from stakeholders and create a shared understanding of values between designers and stakeholders, we see two audiences of such a tool. First of all, the stakeholders who need support to reflect on their values, understand them and then express them to the designer. At this point, we focus our development on supporting direct stakeholders of the technology to be designed. This does not mean the tool could not generally be used by indirect stakeholders. However, indirect stakeholders may not see the need for spending effort on value reflection if they are not directly involved either in developing or using the new technology. Their involvement may need a different functionality of a tool than proposed in the following. This is outside the scope of this paper. Besides stakeholders the second target group are designers who need to get a grasp of all values at play and make design decisions based on them. Given these two very different target groups the tool needs to operate on two levels, as a personal device used in everyday contexts and as an analysis and communication tool. Therefore, a combination of a mobile phone application used by the stakeholders throughout the day and a website that provides a plattform for the stakeholders and designers to make sense from the collected data would be a good solution. Despite VSD's focus on the design process, the tool we envision at this stage will put more focus on eliciting values and value reflections than integrating these values into the design process. Although the latter will be supported as the designer and stakeholder can communicate and thus share design thoughts, it is not central to the system. In this respect the tool is a general value elicitation tool that could also be used by other user groups, e.g. ethnographers, counselors. It could even be used by people that are generally interested in reflecting on their values (e.g. for decision support). That said, we do not exclude future work on integrating the tool with tools focusing on

the concrete designs of new technology. Furthermore, the openness of a general value elicitation tool also allows for the designer to ask for reflections linked directly to prototypes of the new technology, similar to the idea of Technology Probes (Hutchinson et al. 2003). In the following we will first discuss related work that guides the design of our envisioned tool. Then, we will describe the first steps we have taken in the design of the tool and last, sketch out possibilities of the tool's implementation. At this stage, several research prototypes exist (on paper and digital). However, a final implementation that has been tested with users is still under development.

Related work

Part of the inspiration for creating our mobile tool comes from the young research area of Personal Informatics. Personal informatics systems support people in collecting information about their lives to self-reflect and gain self-knowledge (Li et al. 2010). Li and colleagues identify two core aspects of every personal informatics system, namely collection (collecting information about oneself) and reflection (reflecting on personal information). Although personal informatics systems are usually not used as value elicitation tools, several systems implicitly focus on values, e.g. by trying to improve people's health or happiness (for an overview of tools see http://www.personalinformatics.org/tools). Especially in the area of health applications the use of biosensors has been explored and physiological input is translated to visualizations that support people's understanding of, e.g. their activity or stress levels in the Affective Health project (Sanches et al. 2010; Ståhl et al. 2009). In this case the physiological data is one of the main sources of reflection. Another use of physiological data is as a basis for retrieving meaningful data in the vast amount of data that is gathered in life-logging. Kelly and Jones (2009) found, among others, a correlation between the importance of retrieved events from the SenseCam they used and the user's maximum skin conductance values.

While personal informatics systems offer great opportunities for our work on value elicitation, they focus on self-understanding of people, in our case stakeholders, but do not involve system designers or researchers and therefore do not offer any communication facilities between designers and stakeholders. Methods allowing researchers or designers to study stakeholders' attitudes towards technology by contextual data gathering are Cultural Probes (Gaver et al. 1999) and Technology Probes (Hutchinson et al. 2003) or Experience Sampling (Consolvo and Walker 2003). Cultural and Technology probe studies are longitudinal (at least several days) and require users to fulfill several tasks and report their experiences in a diary. A disadvantage of diary studies is that the stakeholders have

to take initiative to capture, remember and report information. This can lead to problems when the participants forget to report or report on events that are of less interest to the researcher (Khan et al. 2009). The Experience Sampling Method (ESM) solves the latter issue by using sampling periods defined by the researcher. In ESM studies participants receive signals, originally on a pager or nowadays on a smartphone or PDA, and are then asked to fill in short questionnaires administered by the researchers. This procedure occurs repeatedly during the sampling period and by that allows the researcher to collect a high amount of in-situ data. A major downside of ESM is that it requires a lot of time from the participants, interrupts their current tasks and requires them to answer the same questions repeatedly leading to boredom. This can result in a low compliance rate and people ignoring the prompts (Hsieh et al. 2008). It is therefore important to keep the participants motivated. Hsieh et al. (2008), concluded from a field study showing that feedback in the form of visualisations of the gathered data allowed people to learn something about themselves and increased compliance rates by 23%. Another way to keep participants engaged is providing the researchers real time access to the gathered data and by that allowing them to dynamically adapt the questionnaires or contact the participant. Two systems have been developed that support this interaction, MyExperience (Froehlich et al. 2007) running on Windows Mobile and PocketBee (Gerken et al. 2010) for the Android platform. PocketBee additionally allows multimodal input by the participants allowing them to choose the easiest or most appropriate feedback channel for giving feedback. For value elicitation in value sensitive design both self-reflection by the stakeholder and communication of values between designer and stakeholder are needed. Both aspects can be supported by personal informatics and ESM respectively. In the following section we will propose an implementation of a tool combining personal informatics and ESM with a focus on values.

Design process of a value elicitation tool

Based on inspiration from the related work presented above, we developed a general concept for our tool. In the work presented so far we identified two essential aspects for a realistic value elicitation and inclusion in the design context: (1) in-situ data gathering and reflection of the stakeholder about values relevant to a given design context and (2) interpretation and communication of values to reach a shared understanding between stakeholder and designer. Based on these aspects we decided to build a tool consisting of a mobile application and a website as explained in the following. Next, we describe our co-design process involving counselors and the first design

prototypes of the reflection functionality provided by the website.

Mobile application concept

The mobile application (Fig. 1, left) is to be used by the stakeholders in everyday life to gather data connected to their values. This is not as straightforward as data collection in other personal informatics applications, e.g. gathering health information such as calorie intake. Values guide people's actions and choices in everyday life, but it is hardly feasible to ask a user which value was guiding her behaviour each time she takes an action or decision. An important aspect of the application is to identify meaningful moments, which can be user driven or system driven. Ideally the two need to be combined in order not to miss out on important situations that either the system or the user failed to identify. We have implemented a first user-driven version for the Android platform. Using smartphones opens up many possibilities to 'sense' the context and identify the right moments, which we will implement in the next version [e.g. by making use of sensors embedded in the system by using the Sense technology (http://www.sense-os.nl)]. Examples of sources the phone can use to identify moments of importance are location data (GPS), long lasting phone calls or sms conversations (e.g. from a favorite contact). Furthermore, the use of wearable biosensors (e.g. measuring GSR) could give insights into how important (arousing) a situation is for the user, but needs to be handled with care.

Besides identifying meaningful situations, making the user aware of the context is important. Therefore, the application asks the user to take a photograph and tag it with value-related words. To further guide people in explicitly reflecting on situated values we ask stakeholders to give a short assessment of the situation by filling in her emotions and evaluation of the situation in terms of good or bad, right or wrong, beneficial or harmful. The evaluation of the situation will be followed by a request to provide (brief) reasons for the particular evaluation, the underlying assumption being that the evaluation stems from the difference between what is (the situation) and what ought to be (the value). Providing such reasons should be optional, so as not to burden users. Logging quantitative and qualitative (and physiological) context data enables a triangulation of data that can lead to a better self-reflection suited to personal preferences (visual, textual, etc.).

Website: overall concept

The data gathered by the mobile application will be uploaded to a server and represented on a website (Fig. 1, right) that is accessible by the stakeholder and the designer.

Fig. 1 Mock-ups of the concept: mobile application (*left*) and website (*right*)
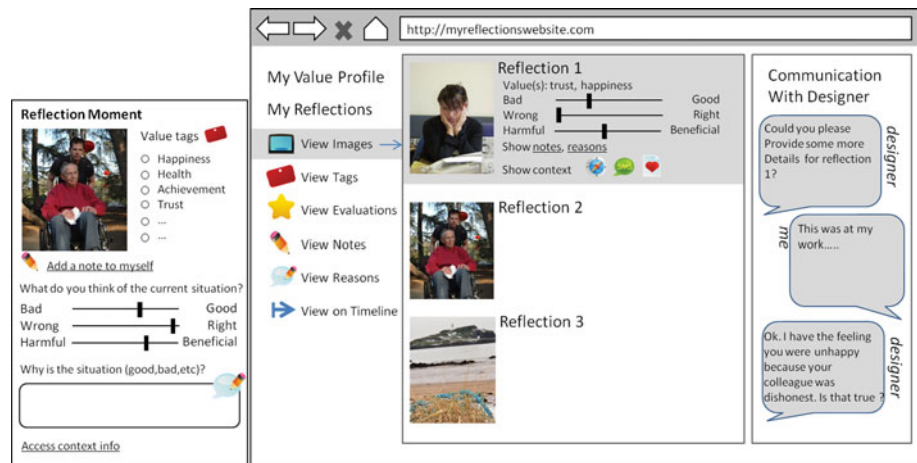
The stakeholder can log onto the website to look at the gathered data, but also enter new reflections. Reflections on the website could be supported by doing small reflection tasks, which will be described in detail below. Based on our study results several modalities should be supported serving a stakeholder's preferences, i.e. providing textual, visual, graphical information. Up to this point the website is similar to current personal informatics systems. It differs, however, in that it supports the communication between the stakeholders (users of the mobile application) and designers (administrators of the mobile application). The designer has access to the stakeholder's data on the website and will play an active role in the meaning making process. It is important to protect the stakeholder's privacy and make sure the decision of which personal reflection data to share with the designer. Some of this data may be sensitive and only be used for personal reflection. Besides being able to show parts of the data to the designer, a message function will be employed that allows private discussions between one stakeholder and the designer. Another option to consider, which is discussed later in this paper, is a community approach in which several stakeholders could share thoughts and enter the same message thread. The communication we generally envision is two-fold, the designer can comment and clarify concrete data released by the stakeholder or start a discussion by asking concrete value related questions and presenting design options. An ultimate goal should be to express the identified values as requirements for the new system's design.

Brainstorm session and guidelines for supporting value reflections

With the general concept in mind we started engaging in a co-design process with experts in order to design the part of the website that supports people's value reflections. Please note, that this does not include the communication with the designer. The implementation details of that part are still being developed and therefore not included in this paper. We chose to interview counselors as they help others articulate their values and how to live their lives (or make at least some decisions) in accordance with them. From our first two interviews and brainstorms we could extract several methods the experts use that focus on the people's own reflection of values. These include, association card exercises, storytelling or symbolic thinking, which similarly engage the person in first describing something concrete (a picture they like, an event or an animal that represents them) and then slowly moving to deliberations about why the person picked a certain thing, what it means to her and what values are reflected in that picture, story or metaphor. Besides these methods, the experts emphasized the uniqueness of every client and every session with a client as well as the difficulty for a client to open up. This results in a need for a trust relationship between counselor and client as well as openness to offer several methods for reflection. From the interview transcripts we could extract the following design guidelines which were sent back to and acknowledged by the counselors:

- The system should offer different ways/tools to reflect as people are unique.
- The system needs to be open in a way that people can choose a reflection method that fits them best and navigate through the system as they want.
- (At the same time) the system should provide guidance so that the user is not lost and does not skip necessary steps in the reflection process.
- The system should avoid using checklists.
- The system should divide the reflection process into small concrete steps, instead of asking the user to reflect on abstract values.
- The system should stimulate the user with questions asking why…? or what…? rather than how…? (Why is x important to you?)

- The system should support the build up of a user's trust in the system.
- The system should not judge the users input, but promote individual truth. (This means, e.g. that contradictions should not be pointed out as such but the system should ask the user to elaborate in a polite way.)
- The user has to feel comfortable with the system. (can e.g. be achieved by customization, using themes, colors etc.)

Besides these guidelines, another aspect that was mentioned by the counselors was a community-based approach. In their experience group sessions where people can share their thoughts with others and make sense out of different situations and reflections together work better for some people than individual conversations. As a major aspect of our work concerns the sharing of value deliberations with regard to a new technology and finding possible design trade-offs that match most stakeholders' values the community aspects fits nicely. Similar to other networking sites, our website could support a community of users. Sharing would then take place on two levels. First, stakeholders could share more intimate reflections with close ones who they are linked to in order to arrive at a better understanding of their values. Second, stakeholders and designers of the same project could share messages either with the whole group (including all stakeholders) or only with the designer to discuss possible value trade-offs and design solutions. It is important to design this feature of the tool carefully though to avoid invasion of the users' privacy while sharing sensitive information. The user needs to be able to specify in detail who is able to see a message or reflection data. The interface needs to be designed in a transparent way to support a user's understanding of how her data is handled and how to adjust privacy settings in an intuitive way (e.g. clear selection of receipients of a message/other data).

Website prototypes

Given what we have learned from the experts, we (three project members) brainstormed ideas for the part of our website that will support the users in engaging in value reflection excersises. The three researchers independently created a number of paper sketches each (see Fig. 2 for an example). We are also implementing digital versions that are to be shown to the counselors in the follow-up session. As these prototypes are meant for demonstrating the possible design and functionality of the tool the interactivity is limited. We are planning to engage the counselors in designing a new version of the tool drawing from the current ideas and their own. As mentioned before the implementation of this final version and its evaluation is still future work.

Limitations

In our study we did not consider all possible methods of value elicitations. We focused on photo elicitation because it supports people to use context when talking about values but is less obstrusive than e.g. in-situ enquiry. However, it would be interesting to apply different in-situ methods and see in what way people are able to reflect on their values in an everyday life situation. As our study was of exploratory nature we considered only a small number of participants. In order to draw general conclusions about the population's ability to express values, gender specific differences and affective reactions, a larger scale study is necessary. Furthermore, next to skin conductance and heart rate, other affective reactions (facial expression, speech, EEG etc.) can be taken into account in order to identify meaningful moments for self-reflection. Despite the small sample size we still believe that the study shows some interesting findings about the differences in how people are able or prefer to express their values. In this way it influenced the design of our tool to make sure to account for such differences.

Discussion

Conclusions

In this paper we argued that values are situated meaning that they differ in roles and importance to people depending on the situational context in which they guide human behavior or decisions. This entails that general value importance rankings are hard to obtain from people. However, for technology design and especially value sensitive design it is important that a designer is aware of the relevant values that stakeholders hold regarding the technology in question. One of our research questions focused on how current approaches in requirements engineering support the elicitation of these situated values. We analyzed the suitability of traditional requirements elicitation approaches and user centered design methods for value elicitation. We showed that many existing methods are not sufficient to elicit situated values and to provide a shared understanding about the importance of values between designers and stakeholders. Value sensitive design deals with the former aspect (elicitation of situated values) and several studies in this area have shown the suitability of known social science methods used in-situ to elicit values with regard to an existing technology. Our work tried to build on this work by showing how different methods compare in supporting people's value expressions and the elicited values. Furthermore, we found that there is a lack of tools that support self-reflection by stakeholders and a shared understanding of values between stakeholders and designers.
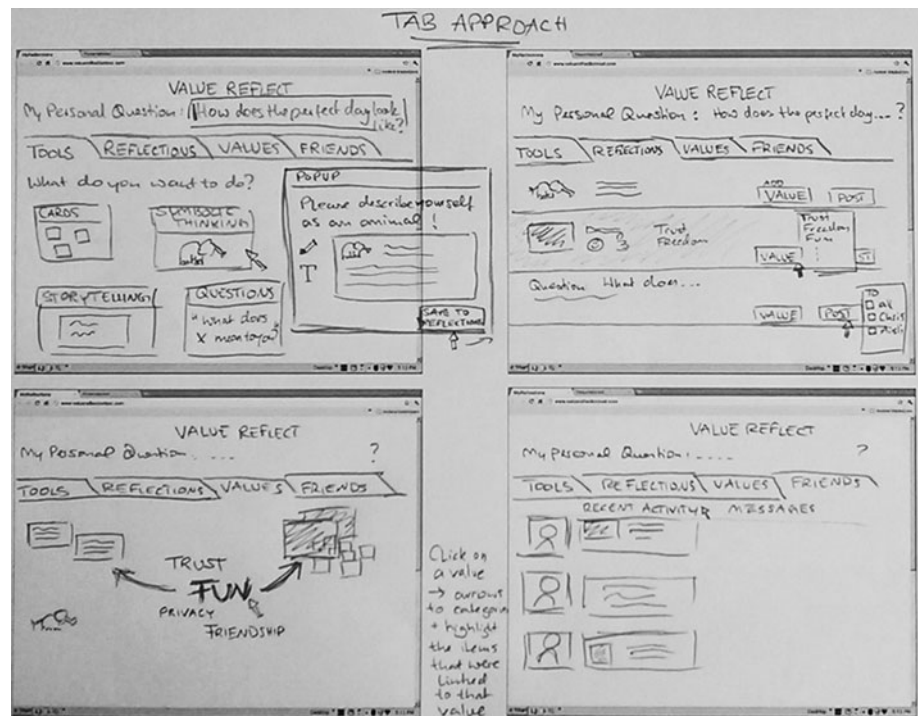
Our goal is to create such a tool and we proposed an initial design. In our initial exploratory user study we focused on understanding how people can express their values and which methods support that expression. We compared a value elicitation questionnaire, verbal expressions of values in a photo elicitation interview and value tagging of pictures made in daily life. The study resulted in the following findings. First of all, we learned that there are great personal differences in which method for expressing values works best. Thus any value elicitation tool needs to allow for triangulation. Second, we learned that key to communicating values to a designer is self-reflection. Talking about important moments in one's life was possible for most participants, however, not everyone could relate such situations and one's own behavior to the underlying values. This requires self-reflection on behavior and decision making over a longer period of time. As this is not natural for all people, triggers to self-reflect and support in analysing the collected data needs to be provided. While the designer can trigger discussions in the presented tool, we also looked into how to guide self-reflection by other means. Interviews of counselors provided interesting insights into triggering self-reflection that we could integrate into several design prototypes of our tool. Inspired by personal informatics and experience sampling the tool consists of a mobile application used by stakeholders for data collection and in-situ self-reflection and a website used collaboratively by designers and stakeholders to further reflect and communicate values. Our study results and the presented design solution are first steps to answer our

second research question: "How can we support stakeholders and designers to arrive at a shared understanding of relevant values for the technology in question?" Future work is, however, needed to evaluate the presented solution.

Future directions

Our work presented in this paper serves as a starting point for elicitation of situated values. We argued that this is needed to advance value sensitive design and generally create a more practical approach to involving ethics in the design of new technologies. By pointing out links between work in different fields, such as requirements engineering, human computer interaction, personal informatics and social sciences we would like to encourage collaborations to further advance value elicitation. We have taken the first step in the design of a value elicitation tool(kit). Possibilities for future research are (1) the investigation of methods such as laddering (used in marketing) or value scenarios, (2) how to support the analysis of the stakeholder's behavior data by means of adequate visualizations or (3) how to further support communication between the designer and stakeholder. In addition, the possibility of using affective reactions in the elicitation process seems promising, but needs further investigation. In a first analysis of our skin conductance data we found that emotional reactions may give an indication of how important certain situations and the affected values are. Therefore, querying participants about their mood or emotion

connected to a situation will help self-reflection. A first exploratory analysis of the data showed that skin conductance measurements can be used to identify (to at least 47.5%) important situations and thus be used to trigger the stakeholder to reflect. As the affect values were obtained in an interview setting in which important events were merely memorized we believe that in real situations affective reactions may be even stronger, but also more noisy. However, we cannot rely on affective reactions alone to trigger self-reflection, as the extent of affective reactions are based on how emotional an individual is.

More specifically, *our* future work will deal with the implementation of the proposed design for a value elicitation tool. Several questions that were left outside the scope of this paper arose, e.g. What additional functionality is needed for indirect stakeholders? How might the tool support stakeholders taking on different roles in relation to a proposed technology (e.g. the same person who in one context is the doctor using a medical record system and in another context the parent of a child in the hospital)? Our final design will be tested and evaluated first in a real life study and later in value sensitive design projects.

References

Bardi, A., & Schwartz, S. H. (2003). Values and behavior: Strength and structure of relations. *Personality and Social Psychology Bulletin, 29*(10), 1207–1220.

Bresciani, P., Perini, A., Giorgini, P., Giunchiglia, F., & Mylopoulos, J. (2004). Tropos: An agent-oriented software development methodology. *Autonomous Agents and Multi-Agent Systems, 8*(3), 203–236.

Cheng, A. S., & Fleischmann, K. R. (2010). Developing a meta-inventory of human values. *Proceedings of the American Society for Information Science and Technology, 47*, 1–10.

Chung, L., Nixon, B. A., Yu, E., Mylopoulos, J. (1999). Nonfunctional requirements in software engineering. In *International series in software engineering* (Vol. 5). Springer.

Clark-Ibez, M. (2004). Framing the social world with photo-elicitation interviews. *American Behavioral Scientist, 47*, 1507–1527

Consolvo, S., & Walker, M. (2003). Using the experience sampling method to evaluate ubicomp applications. *IEEE Pervasive Computing, 2*(2), 24–31

Cysneiros, L., & do Prado Leite, J. (2004). Nonfunctional requirements: From elicitation to conceptual models. *IEEE Transactions on Software Engineering, 30*(5), 328–350.

Dardenne, A., van Lamsweerde, A., & Fickas, S. (1993). Goal-directed requirements acquisition. *Science of Computer Programming, 20*, 3–50.

DeLoach, S., Padgham, L., Perini, A., Susi, A., & Thangarajah, J. (2009). Using three aose toolkits to develop a sample design. *International Journal of Agent-Oriented Software Engineering, 3*(4), 416–476.

Denning, T., Borning, A., Friedman, B., Gill, B., Kohno, T., & Maisel, W. (2010). *Patients, pacemakers, and implantable defibrillators: Human values and security for wireless implantable medical devices* (pp. 917–926). New York: ACM (CHI '10).

Friedman, B., Freier, N. G., Kahn, P., Lin, P., & Sodeman, R. (2008). Office window of the future? Field-based analyses of a new use of a large display. *International Journal of Human-Computer Studies, 66*(6), 452–465.

Friedman, B., Hook, K., Gill, B. T., Eidmar, L., Sallmander Prien, C., & Severson, R. L. (2008). *Personlig integritet: A comparative study of perceptions of privacy in public places in Sweden and the United States* (pp. 142–151). New York: ACM (NordiCHI'08).

Friedman, B., & Kahn, P. H. J. (2003). Human values, ethics, and design. In J. A. Jacko, & A. Sears (Eds.), *The human-computer interaction handbook* (pp. 1177–1201). UK: Lawrence Erlbaum Associates.

Friedman, B., Kahn, P., & Borning, A. (2006). Value Sensitive design and information systems. In *Human-computer interaction and management information systems: Foundations* (pp. 348–372). New York: M.E. Sharpe.

Friedman, B., Kahn, P., Hagman, J., Severson, R. L., & Gill, B. (2006). The watcher and the watched: Social judgments about privacy in a public place. *Human-Computer Interaction, 21*, 235–272.

Froehlich, J., Chen, M. Y., Consolvo, S., Harrison, B., & Landay, J. A. (2007). Myexperience: A system for tracing and capturing of user feedback on mobile phones. In *MobiSys '07*. New York: ACM.

Gaver, B., Dunne, T., Pacenti, E. (1999). Design: Cultural probes. *Interactions, 6*, 21–29.

Gerken, J., Dierdorf, S., Schmid, P., Sautner, A., & Reiterer, H. (2010). Pocket bee: A multi-modal diary for field research. In *Proceedings of the 6th Nordic conference on human-computer interaction: Extending boundaries* (pp. 651–654). New York, NY, USA: ACM (NordiCHI '10).

Giunchiglia, F., Mylopoulos, J., & Perini, A. (2002). The tropos software development methodology: Processes, models and diagrams. In: *AOSE* (Vol. 2585, pp. 162–173). Springer.

Harper, D. (2002). Talking about pictures: A case for photo elicitation. *Visual Studies, 17*, 13–26.

Hodges, B. H., & Baron, R. M. (1992). Values as constraints on affordances—perceiving and acting properly. *Journal for the Theory of Social Behaviour, 22*(3), 263–294.

Hsieh, G., Li, I., Dey, A., Forlizzi, J., & Hudson, S. E. (2008). Using visualizations to increase compliance in experience sampling. In *Proceedings of the 10th international conference on Ubiquitous computing* (UbiComp '08) (pp. 164–167). New York, NY, USA, : ACM.

Hutchinson, H., Mackay, W., Westerlund, B., Bederson, B. B., Druin, A., Plaisant, C., et al. (2003). Technology probes: Inspiring design for and with families. In *Proceedings of the SIGCHI

conference on Human factors in computing systems (CHI '03) (pp. 17–24). New York, NY, USA: ACM.

Kahn, P. H., Friedman, B., Gill, B. T., Hagman, J., Severson, R. L., Freier, N. G., et al. (2008). A plasma display window? The shifting baseline problem in a technologically-mediated natural word. *Journal of Environmental Psychology, 28*(2), 192–199.

Kelly, L., & Jones, G. (2009). Examining the utility of affective response in search of personal lifelogs. In *5th workshop on emotion in HCI*, British HCI conference.

Khan, V., Markopoulos, P., & Eggen, B. (2009). Features for the future experience sampling tool. In *MobileHCI '09*. New York: ACM.

Kolos-Mazuryk, L., Eck, P., & Wieringa, R. (2006). *A survey of requirements engineering methods for pervasive services.* Technical report. University of Twente, The Netherlands.

Kraemer, F., van Overveld, K., & Peterson, M. (2011). Is there an ethics of algorithms? *Ethics and Information Technology, 13*(3), 251–260.

Krogstie, J. (2001). Requirement engineering for mobile information systems. In *Proceedings of the seventh international workshop on requirements engineering: Foundations for software quality (REFSQ'01)*.

LeDantec, C. A., Poole, E. S., & Wyche, S. P. (2009). Values as lived experience: Evolving value sensitive design in support of value discovery. In *Proceedings of the 27th international conference on Human factors in computing systems*. Boston, MA, USA: ACM.

Li, I., Dey, A., & Forlizzi, J. (2010). A stage-based model of personal informatics systems. In *Proceedings of the 28th international conference on Human factors in computing systems*. Atlanta, Georgia, USA: ACM.

Maio, G. R. (2010). Mental representations of social values. *Advances in Experimental Social Psychology, 42*, 1–43.

Massonet, P., & van Lamsweerde, A. (1997). Analogical reuse of frameworks. In *Proceedings of the 3rd IEEE international symposium on requirements engineering* (pp. 26–33).

Miceli, M., & Castelfranchi, C. (1989). A cognitive approach to values. *Journal for the Theory of Social Behaviour, 19*, 169–193.

Miller, J. K., Friedman, B., Jancke, G., & Gill, B. (2007). Value tensions in design: The value sensitive design, development, and appropriation of a corporation's groupware system. In *Proceedings of GROUP 2007* (pp. 281–290). New York: ACM.

Morandini, M., Nguyen, D., Perini, A., Siena, A., & Susi, A. (2008). Tool-supported development with tropos: The conference management system case study. In M. Luck, & L. Padgham (Eds.), *Agent-oriented software engineering VIII, lecture notes in computer science* (Vol. 4951, pp. 182–196). Berlin, Heidelberg: Springer.

Nathan, L., Klasnja, P., & Friedman, B. (2007). Value scenarios: A technique for envisioning systemic effects of new technologies. In *CHI'07 extended abstracts on Human factors in computing systems* (pp. 2585–2590). USA: ACM.

Nathan, L., Friedman, B., Klasnja, P., Kane, S., & Miller, J. (2008). Envisioning systemic effects on persons and society throughout interactive system design. In *Proceedings of the 7th ACM conference on designing interactive systems* (pp. 1–10) .USA: ACM .

Nuseibeh, B., & Easterbrook, S. (2000). Requirements engineering: A roadmap. In *ICSE '00 Proceedings of the conference on the future of software engineering* (pp. 35–46). USA: ACM.

Pauls, R. (1990). *Concepts of value: A multi-disciplinary clarification.* Technical report 20. Centre for Resource Management, Lincoln University, Lincoln.

Potts, C. (1999). Scenic: A strategy for inquiry-driven requirements determination. In *Proceedings of IEEE fourth international symposium on requirements engineering (RE'99)* (pp. 58–65).

Rokeach, M. (1973). *The nature of human values*. New York: Free Press.

Sanches, P., Höök, K., Vaara, E., Weymann, C., Bylund, M., Ferreira, P., Peira, N., & Sjölinder, M. (2010). Mind the body! Designing a mobile stress management application encouraging personal reflection. In *Proceedings of the 8th ACM conference on designing interactive systems* (pp. 47–56). New York, NY, USA: ACM (DIS '10).

Schroeder, M. (2008) Value theory. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy, fall 2008 edn*. Stanford: Stanford University.

Schwartz, S. H., & Bilsky, W. (1990). Toward a theory of the universal content and structure of values: Extensions and cross-cultural replications. *Journal of Personality and Social Psychology, 58*, 878–891.

Ståhl, A., Höök, K., Svensson, M., Taylor, A., & Combetto, M. (2009). Experiencing the affective diary. *Journal of Personal and Ubiquitous Computing, 13*(5), 365.

Sutcliffe, A. (1998). Scenario-based requirements analysis. *Requirements Engineering, 3*(1), 48–65.

Sutcliffe, A. (2003). Scenario-based requirements engineering. In *Requirements engineering conference, 2003. Proceedings. 11th IEEE international* (pp. 320–329). IEEE.

Thew, S., & Sutcliffe, A. (2008). Investigating the role of 'soft issues' in the re process. In *16th IEEE international requirements engineering, 2008 RE'08* (pp. 63–66).

van Lamsweerde, A., & Letier, E. (1998). Integrating obstacles in goal-driven requirements engineering. In *Proceedings of the 20th international conference on software engineering*.

van Lamsweerde, A., & Willemet, L. (1998). Inferring declarative requirements specifications from operational scenarios. *IEEE Transactions on Software Engineering*.

van Lamsweerde, A., Letier, E., & Darimont, R. (1998). Managing conflicts in goal-driven requirements engineering. *IEEE Transactions on Software Engineering*.

Voida, A., & Mynatt, E. D. (2005). Conveying user values between families and designers. In *CHI '05 extended abstracts on Human factors in computing systems*. Portland, OR, USA: ACM.