

# How to approximate users' values while preserving privacy: experiences with using attitudes towards work tasks as proxies for personal value elicitation

Sven H. Koch · Romyana Proynova ·  
Barbara Paech · Thomas Wetter

Published online: 8 February 2013  
© Springer Science+Business Media Dordrecht 2013

**Abstract** Software users have different sets of personal values, such as *benevolence*, *self-direction*, and *tradition*. Among other factors, these personal values influence users' emotions, preferences, motivations, and ways of performing tasks—and hence, information needs. Studies of user acceptance indicate that personal traits like values and related soft issues are important for the user's approval of software. If a user's dominant personal value were known, software could automatically show an interface variant which offers information and functionality that best matches his or her dominant value. A user's dominant personal value is the one that most strongly influences his or her attitudes and behaviors. However, existing methods for measuring a user's values are work intensive and/or interfere with the user's privacy needs. If interface tailoring for very large groups of users is planned, value approximation has to be achieved on a large scale to assign individualized software to all users of the software. Our work focuses on approximating the dominant values of a user

with less effort and less impact on privacy. Instead of probing for a user's values directly, we explore the potential of approximating these values based on the user's preferences for key tasks. Producing tailored versions of software is a separate topic not in the focus here. In this paper we rather describe a method to identify user values from task preferences and an empirical study of applying parts of this method. We are proposing the method in this paper for the first time except for a preliminary version orally presented at a workshop. The method consists of a research process and an application process. In the research process a researcher has to identify key tasks occurring in a context under investigation which have a relationship to personal values. These key tasks can be used in the application process to approximate the dominant values of new users in a similar context. In this empirical study we show that the research process of our method allows us to determine key tasks which approximate values in the shared context of nursing. The majority of the nurses were found to have one of the three following dominant values: *benevolence*, *self-direction*, or *hedonism*. Data confirmed common expectations: that nurses with the value of *benevolence*, when compared to all other nurses, had a higher preference for tasks which helped people immediately or improved their circumstances of the treatment. In relation to all other nurses, participants with *self-direction* disliked tasks which affected their personal freedom, and users with *hedonism* had a lower preference for tasks which involved physical work and preferred tasks which promised gratification. Our findings advance measurement of personal values in large user groups by asking questions with less privacy concern. However, the method requires substantial efforts during the initial research process to prepare such measurements. Future work includes replicating our method in other contexts and identifying value-dependent

---

S. H. Koch (✉) · T. Wetter  
Institute of Medical Biometry and Informatics, Heidelberg  
University, Heidelberg, Germany  
e-mail: Sven.h.koch@gmail.com

T. Wetter  
e-mail: thomas.wetter@urz.uni-hd.de

R. Proynova · B. Paech  
Institute of Computer Science, Heidelberg University,  
Heidelberg, Germany  
e-mail: Proynova@informatik.uni-heidelberg.de

B. Paech  
e-mail: Paech@informatik.uni-heidelberg.de

T. Wetter  
Department of Medical Education and Biomedical and Health  
Informatics, University of Washington, Seattle, WA, USA

tasks for users with other values than the three values our empirical study mainly focused on.

**Keywords** Personal values · Value elicitation · Value approximation · Privacy · Software tailoring · Individualized interface · Value-sensitive design

## Introduction

Tailoring software to individual user needs has become increasingly important for technology development in recent years. Interfaces which are developed with values in mind are better suited for the user's needs e.g. (Nissenbaum 2005; Flanagan et al. 2008; Pommeranz et al. 2011a, b). With "interface" we refer to a specific set of interactions with software including functions that the user can perform and information he or she is being offered.

Software—when delivered out of the box—is often overloaded with functionality and information, because it needs to satisfy the diverse needs of individual users. Customized interfaces, adapted to the individual user's way of working and thinking, could help the individual to better perform his or her tasks and increase the user's satisfaction and the individual utility of the software.

However, many users utilize software which is not customized to their needs, either because identically customized software versions are installed for groups of users, or because customization cannot be performed easily by the users themselves. Identically customized software is often installed for groups of users—such as for all nurses of the department of surgery in a hospital—and the selection of information and features is most likely based on the job role, but not on the individual user's way of working and thinking. Individual customization—the user customizes software him- or herself—is often not feasible, because it requires additional time and profound expertise, or because users fear unwanted side effects.

Software tailoring (Dorn et al. 2009; Kobayashi and Tsunawaki 2007; Sommerville 1981) suggests that interfaces should be automatically customized to the individual's needs. However, current software tailoring methods focus on groups of users (Dorn et al. 2009), so that the individual's needs might not be met. Other tailoring methods, e.g. (Kobayashi and Tsunawaki 2007) are work-intensive or intrusive.

If tailoring could run automatically drawing on easily assessed individual users' characteristics, some of the above problems could be avoided. One promising concept to tailor interfaces individually, which is less intrusive and requires a low workload per user, is to utilize the user's personal values for the interface customization. Personal values or beliefs are concepts which guide individuals

during their life and stay constant over time (Schwartz et al. 2001). They influence a user's goals, decisions, motivation, and preferences. Among such values are the concepts of *benevolence*, *self-direction*, and *power* (Schwartz 1992). Specific values influence the tasks individual users see as essential to reach goals. Therefore, a user's values might indirectly influence which information should be provided through interfaces and which functions are needed to perform associated task. More details on personal values can be found in the Background section.

Examples of value-dependent interface functionality and information are explained in the following: Imagine two physicians: the first one's predominant personal value is *benevolence* and the second one's personal values rather have *power* at the top. Both physicians would need to perform a similar set of basic tasks which all physicians perform and need a similar set of information and functionality for these tasks. However, for some specific additional tasks the information and functionality could differ according to their dominant personal value *benevolence* and *power*. The dominant value is the most important personal value of the individual. Imagine that both physicians are currently seeing a patient and are selecting a medication regime among multiple treatment options. The *benevolent* physician may want to select the least harmful and easiest to follow option to make sure the patient always feels at ease and will be able to successfully follow the treatment regime on his or her own. Potentially preferred functions for *benevolence* include information about the burden of treatment options on the patient and his/her quality of life and how complicated a medication regime is for the patient. The other, *power*-centered, physician might think that the patient needs to follow any kind of strict medication regime, if determined best by his or her physician. Furthermore, after selecting a medication regime, the physician might not want to deal with small details such as explaining every step to the patient, but might want to delegate such explanations to subordinate persons who should then continue the explanation. Possibly preferred functions for *power* include therefore seeing the effectiveness of different medication regimes and the ability of adding tasks to subordinate and other co-workers' to-do lists and seeing their workloads to make sure ordered tasks are performed in time.

However, measuring personal values with the currently used methods is neither easy nor fast, see Background section for more details. Dependent on the selected method, value measurement involves very personal questions and is time-consuming for the person determining the user's value. Personal questions in value measurement questionnaires (Schwartz 1992; Schwartz et al. 2001) include for example details about the user's personal character, opinions, and beliefs such as his or her honesty,

willingness to obey, and responsibility. Other methods, which involve personal interaction between the user and the value-measuring researcher, such as ethnographic observation, require excessive time commitments if the values of many users have to be determined.

Our work focuses on a method for personal value approximation for situations where values of many users need to be determined and users might have privacy concerns. We propose to approximate values based on the user's preference of work tasks. We assume that prior to approximating the user's values, multiple value dependent interface variants of a software are developed (this is a part of interface development and is not subject of this paper). For each user, when initially using the software, we suggest to determine the most dominant personal value. This dominant value would finally be used to select the appropriate interface variant for this user's dominant value. For all subsequent logins, the user would then see this interface variant's tailored information content and functionality. In this paper, we are proposing the method for the first time except for a preliminary version orally presented at the Workshop on Values in Design (Koch et al. 2011).

In the following Background section, we first describe what we mean by the user's personal values, and review currently used methods to elicit user values. The subsequent section explains our method to approximate a user's dominant personal values. Afterwards, we show with an empirical study in the medical domain that our method actually works to approximate nurses' dominant personal values. In the last section, we discuss our results, possible implications and limitations of the method.

## Background

First we describe existing competing approaches to explain driving forces of the personality and then focus on the concept of personal values. We summarize previous work which uses values in software development, and finally review methods which are currently used or proposed to elicit user values.

### Driving forces of the personality

"It's a matter of taste" is an often used quote when apparently similar people disagree diametrically in their judgment or liking of the appearance or functionality of an object, procedure or service. Inasmuch as this affects professionalism and work place satisfaction, including satisfaction with software that is used, factors that make people different seem to be a worthwhile research topic. First, it is important to understand and pinpoint the factors. Once determined, the factors can be used to design the workplace

and its processes and tools such that they match the individuals' factors that make up their personality.

Various approaches to explain what drives human behavior have been developed and investigated in the past 30 years. They include the debate between the conscious (Locke and Latham 2002) versus the unconscious (Custers and Aarts 2010) nature of humans making plans and initiating actions, the importance of goals and goal achievement as self-perceived (Bandura 1997) or accommodated to the environment (Oesterreich 1981), and reference systems residing in the individual such as personal values (Schwartz 1992). These will be briefly outlined, trying to argue that personal values suggest themselves as a worthwhile construct to be explored more deeply.

The following analysis will use different dimensions or criteria to assess how versatile a theory of drivers of human behavior is for informing the design of software. We will discuss the *time* frame of scale that is covered by an approach, the question whether a driving factor can be regarded as *innate* in human nature *or culturally trained*, and whether it is a characteristic of *mankind* or comes in *personal* variants. These dimensions are important for our future goal to tailor software: Time plays a role because primarily enduring factors are worth the effort. Results concerning innate factors hold worldwide while culturally trained ones would have to be newly assessed in new cultural environments. Finally, only personally varying factors call for varying software interfaces at all.

According to Custers (Custers and Aarts 2010), unconscious will investigations try to establish primacy of unconsciously processed stimuli over consciously set plans. Presented experimental results include that subjects

- Perform tasks better when unconscious stimuli suggest good performance
- Talk more softly when the setting suggests a library environment
- Do cleaning more thoroughly when there is a scent of cleaning agent in the air.

These and other experimental effects operate instantaneously. If subjects were exposed to other material, other behaviors would be primed. The method of proof is statistical: Individual differences do not matter; it would go unnoticed or at least would not harm the perceived validity of an effect, if some subjects behaved differently. Two stimuli above are *culturally trained*. The cleansing agent setting, though, opens the perspective for *innate* behavior priming, too, because scents directly reach the limbic system as one of the most ancient neuronal units linking perception to behavior. Therefore, although undeniable in their effect upon perception and behavior, unconscious influences don't lend themselves for software variants because they offer no cues to distinguish between users.

Another approach summarizes the attempts to prove the opposite of Custer's conjecture (Locke and Latham 2002): it claims that human behavior can be explained through goals actively and consciously set and rationally pursued.

Goal-setting and achievement of respective tasks operate through *all time ranges*. A *cultural* flavor is definitely present because the goals that the individual judges worth attaining are partially determined through the societal esteem of a goal. Apart from that goals and goal achievement are *individual*: the holder of a goal personally assesses its attainability and attached societal recognition. In this sense the individual is his or her own reference.

The self-efficacy theory, as one flavor of the social-cognitive theory (Bandura 1997), adds a specific form of feedback and self-regulation to theories such as the goal-setting theory. Individuals, who tackle attainable goals and succeed, develop a feeling of achievement or efficacy, and dare to select more ambitious goals. So self-efficacy operates both, *instantaneously* in rejoicing recent achievements or mourning recent failures *and longitudinally* through gradual adaptation of its level to the history of experiences. Perception of success, although felt personally, silently, and without the need to be communicated with others, may well be *culturally trained*: Investigations of reflectivity (Orne 1969) demonstrate that in many settings subjects adjust their personal judgments of what is a success to what their societal peers or seniors would call a success. Self-efficacy as such, however, seems to be a driver throughout *mankind*. To summarize: factors of the goal setting and self-efficacy types are clearly personal but subject to societal reference systems and transient in time.

All these factors are simultaneously present in the individual. Neglecting any of the factors inevitably superimposes noise to observed data. The experimental effort to determine the size of their effects is large, if not prohibitive. We rather search for one factor which is also powerful in predicting behavior and can be assessed with its strengths and with reasonable effort in an individual.

This last factor analyzed here, which also guides the experimental investigation and theory formation presented subsequently, is the portfolio of personal values and their role in driving humans' judgments and behavior.

The word value comes with various meanings. It ranges from face or monetary value of an object through shareholder value of a publicly listed company to personal drivers of behavior and attitudes. This investigation concentrates on the inventory or portfolio of values that individuals have, residing deeply inside themselves (Schwartz et al. 2001). Such values can be regarded as a permanently present reference system. Without necessarily entering consciousness, they set the stage for developing positive or negative feelings about things around the individual.

Values are partially innate and partially driven through culture and society. In investigations among twins that had been brought up separately a 40 % agreement concerning work-related personal values could be shown (Meligno and Ravlin 1998). The values that an individual holds are to a certain extent in agreement with the environment where the person lives and works. For the following analysis it helps to distinguish, what values are and how values "behave" when confronted with properties of the environment; whether values are modified when they permanently collide with environments that are not in accordance or whether individuals avoid or modify such environments. Evidence for individuals' propensity to quit work environments whose organizational culture misfits their personality has been provided by O'Reilly and co-workers (O'Reilly et al. 1991). Without being specific as to how much personality relates to or equals personal values, the authors demonstrate the superiority of an individual's internal states over environmental determinants: at a two-year checkpoint in the presented longitudinal data a majority of employees had left their mismatching organizations.

For the following it helps to distinguish

- Abstract fundamental values (Schwartz et al. 2001)
- Concrete behavioral stereotypes, also called instrumental values (Meligno and Ravlin 1998)
- Job-related role stereotypes, that can help to understand professional work (Meligno and Ravlin 1998)
- Attitudes (Meligno and Ravlin 1998)

Abstract fundamental values are the foundation upon which the others build like instantiations or applications. For example a person who holds *benevolence* as his or her prevalent fundamental value may feel good when playing helping roles in all kinds of settings, i.e. enacts a "helpfulness" action stereotype. For physicians (Meligno and Ravlin 1998) have coined a role stereotype "good Samaritan" and characterized it through excerpts from open ended interviews with anesthetists.

All three—personal values plus the two forms of stereotypes—are regarded as enabling "equipment" that a person has, but which does not transform into observable behavior unless a situation calls for use of the "equipment". In contrast, attitudes are relational; they form as a result of an individual encountering an entity that matches (positively or negatively) the equipment. Typical entities towards which persons have attitudes are work tasks which we will later explore as to their utility to find out underlying personal values. The *benevolent* physician, of whom we assume that he or she is helpful in general and a Good Samaritan in his or her workplace, will develop a positive attitude towards a person or situation where help is needed. He may develop a negative attitude towards managerial situations where resources have to be assigned or withdrawn.

While attitudes come and go as fast as situations come and go, personal values seem to live much longer. This is already obvious from the fact that they are partially innate. However, long term investigations also show that they only start floating when an individual is exposed to an environment hostile to his or her values for a long time (Albarracín et al. 2004). Therefore, if we manage to get hold of a person's values and design software that matches them, we can be confident that the person will be satisfied for a long while.

The following Table 1 gives an idea of how the above constructs relate to each other. It partially draws on investigations with physicians as subjects.

Personal values seem to be suitable characteristics of the individual, and therefore suitable for software tailoring, as they stay constant over time.

#### Definition of personal values and their measurement as established by Schwartz

Personal values describe an individual's basic concepts and beliefs which guide the individual through life. We center our research on the validated personal values theory of Shalom Schwartz (Schwartz 1992). Schwartz' value theory provides us with verified questionnaires for value measurement and specific descriptions of each value concept (Schwartz et al. 2001). We expect that using this theory

will make our research reproducible. The measurement of personal values is described in the section "Methods which are currently used or proposed to elicit user values".

Schwartz is one of the leading researchers in psychological analysis of personal values and found that the values of individuals stay constant over time and are present in individuals of different races, nationalities, and social or cultural background. Schwartz' personal values theory was verified through broad empirical research in many countries and individuals with a wide range of different demographics. His value system is commonly used and differentiates between ten personal values.

Table 2 lists the ten personal values which were determined by Schwartz and short descriptions for each. The Schwartz value system is based on two dimensions: (1) focus on the self or not (self-enhancement vs. self-transcendence) and (2) seeking stability or change (openness to change vs. conservation). So if the directions "self" and "change" were binary with sharp boundaries we could draw a two by two matrix where each Schwartz style personal value could be classified as to whether self-enhancement is or is not contained in the value and if openness to change is or is not contained in the value. The category self-enhancement (focus on self) includes the values *achievement*, *power* and *hedonism*, contrasted by the category self-transcendence (not-self) with the values *universalism* and *benevolence*. The category conservation

**Table 1** Comparison of how different constructs relate to each other

Approach	Time	Innate versus cultural	Mankind versus individual
Unconscious will	Instantaneous	Primarily cultural	Mankind
Goal setting	From instance to life span	Cultural	Individual
Self-efficacy	From instance to life span	Cultural	Mankind
Personal values	Life span	Primarily innate	Individual

**Table 2** Personal values determined by Schwartz (Schwartz 1992; Schwartz et al. 2001) and short descriptions for each

Value type	Motivational goals
<i>Achievement</i>	Personal success through demonstrating competence according to social standards
<i>Benevolence</i>	Preservation and enhancement of the welfare of people with whom one is in frequent personal contact
<i>Conformity</i>	Restriction of actions, inclinations and impulses likely to accept or harm others and violate social norms or standards
<i>Hedonism</i>	Pleasure and sensuous gratification to oneself
<i>Power</i>	Social status and prestige, control and dominance over people and resources
<i>Security</i>	Safety, harmony and stability of society, of relationship, and of self
<i>Self-direction</i>	Independent thought and action-choosing, creating, exploring
<i>Stimulation</i>	Excitement, novelty and challenge in life
<i>Tradition</i>	Respect, commitment and acceptance of the customs and ideas that traditional culture or religion provide the self
<i>Universalism</i>	Understanding, appreciation, tolerance and protection for the welfare of all people and for nature



(stability) has the values *security, tradition, and conformity*, contrasted by the category (openness to change) with *stimulation, self-direction, and hedonism* (which belongs to two categories).

Schwartz has created two instruments for determining the personal values of users, the Schwartz Value Scale (SVS) (Schwartz 1992) and the Portrait Value Questionnaire (PVQ) (Schwartz et al. 2001). The SVS was available first. Later, the PVQ was developed as an instrument which is easier to administer, as it requires less abstract thinking, but still delivers results as good as the SVS. We used PVQ for determining the personal values of the participants in our study, as it is the more modern, better developed instrument.

During Schwartz' process of validating the PVQ he found that a large majority of the population has one value that clearly dominates all other values. When measured in an individual, we call this value the person's dominant value.

Personal values are considered to be significant in software development. In the following we describe studies, which use values and related soft issues, and their influence on the user's approval of software.

#### Previous work which uses values in software development

Studies of user acceptance indicate that personal traits like values and related soft issues are important for the user's approval of software, as evidenced for example in (Thong et al. 2004). The idea that human values (personal, ethical and other values) as well as other related constructs (beliefs, attitudes, and emotions) can be employed in software development is not new. Taking these constructs into consideration offers the possibility to better align software with the users' information needs.

Ramos et al. describe a requirements engineering method based on emotions, mentioning values and beliefs as related concepts (Ramos et al. 2005). They provide a valuable framework which shows the importance of knowing the user's personal traits within the requirements engineering process.

Thew et al. focus on the role of "soft issues", for example emotions and motivations in the requirements engineering process (Thew and Sutcliffe 2008, 2011). They describe the impact of such soft issues—as they are exhibited by individual stakeholders—on the elicitation process itself. The emphasis of their research is on the need to shape the communication with stakeholders according to their values, emotions and beliefs.

In the field of Human Computer Interactions there is the concept of Value Sensitive Design (VSD) (Friedman et al. 2006). This concept does not use the social psychology

specific definition of personal values. Instead, it uses a rather society focused approach and centers values around human well-being, human dignity, justice, welfare, and human rights (Friedman and Freier 2005) and accounts for human values in a principled and comprehensive manner throughout the design process (Friedman et al. 2006).

The term *value-based software engineering* has also been used before in a different connotation (Biffel et al. 2006). In this context, value rather focuses on the monetary value of the software product—not on the personal values of users.

In our prior research, we proposed a link between personal values and software requirements based on a literature review (Proynova et al. 2010). We suggested a method to create value-based software requirements (Proynova et al. 2011) and advocated to approximate personal values based on task preference (Koch et al. 2011).

#### Methods which are currently used or proposed to elicit user values

In the following, we describe methods which are currently used or proposed to elicit the personal properties of users to consider them in the context of software design. Table 3 shows an overview of such methods. We included approaches to identify user properties outside the very strict definitions of the personal values as defined by Schwartz, because we are interested in user properties in general as they relate to IT. In the table we estimate for each method the impact of the method on the user's privacy and the time interface developers need for each user. The time interface developers need is relevant if user interfaces need to be adapted for many users.

The privacy impact of the different methods ranges from "low"—when a user would not need to reveal sensitive information—to "high"—when the user is asked to share private information with the researcher or, in case of routine application of a method, with the software engineer. In the following, we describe the different methods in short and explain our reasoning of the privacy impact assessment as given in Table 3. During ethnographic observation, participants are followed by an observer who notes e.g. actions and asks—dependent on the variant of the method—for reasoning and professional as well as personal goals. Continuous observation and personal questions clearly impact the user's privacy—they feel assessed by the observer. They might have the impression that the observer is judging their actions and could share this information with someone else. During user review of scenarios and storyboards, as well as when reviewing prototypes with participants, participants' comments and feedback are used to identify to what extent the system reflects the user's values, motivation, and understanding. If these methods are

**Table 3** Examples of currently used methods to approximate personal user properties and needs, their estimated impact on the users' privacy, and the estimated time commitment for interface developers to determine personal properties for many users

Method	Assumed impact on users' privacy	Estimated time commitment per personal properties elicitation	Proposed or used e.g. by
Ethnographic observation	Medium	High	(Thew and Sutcliffe 2008; Flanagan et al. 2008; Ramos and Berry 2005; Ramos et al. 2005)
User review of scenarios and storyboards	Low	High	(Thew and Sutcliffe 2008)
User evaluation of prototypes	Low	High	(Thew and Sutcliffe 2008)
Discussing users' needs in the design team	Low	High	(Flanagan et al. 2008)
Personal informatics	Medium	High	(Li et al. 2010; Detweiler et al. 2011a)
Interviews	Medium	High	(Friedman 1997; Thew and Sutcliffe 2008, 2011)
Questionnaires	High	Low	(Schwartz et al. 2001; Pommeranz, Detweiler, Wiggers, and Jonker 2011a)

performed correctly, they have a low impact on the participants' privacy, because users only share their opinions about the scenario, storyboard or prototype, but are not asked about personal "secret" details connected to their personality. When discussing user needs in the design team, a shared understanding is built in the team about personal properties of the users' needs. In this method, users are normally not directly involved and above all do not share personal details, so we estimate a low impact on privacy. The method of personal informatics allows participants to collect personally relevant information, for example for the purpose of self-reflection and gaining self-knowledge about their personal values, without directly talking to the developers. Dependent on the concrete questions asked interviews require users to directly reveal private information. Users found it difficult to fill in questionnaires about personal values (Pommeranz, Detweiler, Wiggers, and Jonker 2011a)—but the researchers' workload is low due to automated evaluation.

Our estimated time commitment per personal properties elicitation is based on the time that User Interface (UI) practitioners need per user if values are approximated for a large group of users. Among the methods used, we distinguish between methods which require one-to-one contact between interface developers and user, and methods with no contact required. For direct contact methods, such as participant observation, we assume high time commitments per user and thus per interface developer. For methods without practitioner-user interaction, such as questionnaires, the per-user time commitment is assumed to be low. Questionnaires save time, as they can be administered and evaluated automatically for many users simultaneously and without much effort per individual user.

Although the currently used methods are very valuable for specific applications, none of them seems to be suitable for value elicitation in large and privacy-aware populations and a new approach is needed.

### Towards a method for value approximation

If a user's dominant value were known, a previously developed suitable interface variant could be selected when he or she is starting to use the software. Our method focuses on approximating the dominant value of users. Producing tailored versions of software is a separate process not covered in this paper.

In the following, we first define requirements for value approximation in large groups of users, followed by the utilization of the shared context of having the same profession for value approximation, and finally we describe the steps of our proposed method for value approximation.

#### Requirements of a method for value approximation in large groups of users

A method to approximate values of large groups of users should have a small workload per user and not conflict with the users' privacy requirements. A time efficient method should be suited for fast, self-administered measurements for each user and allow automatic evaluation. Multiple-choice questionnaires satisfy these conditions, if the number of questions is limited to a reasonable amount and questionnaires are administered electronically and evaluated automatically. To minimize the impact on the users' privacy and to limit the risk that users plainly refuse to participate, value measurements should not rely on

questions about very personal (“secret”) information of the users. Measurements should rather rely on information people are willing to talk about such as topics connected to everyday life and work.

#### Utilizing shared contexts to approximate values

We suggest utilizing a context shared by users to allow individual value approximation. Values influence behavior indirectly through attitudes. While individuals are seldom aware of their values, they are aware of their attitudes and use them as rationales for decisions (Tesser and Schwarz 2003; Fishbein 2005). As such, attitudes are one of the value-related concepts which can influence a user’s preferences and expectations about software. An attitude can be expressed as a single statement of the type “I like X” (a positive attitude) or “I don’t like Y” (a negative attitude). Attitudes are formed, among other factors, based on values. For example, if the value *tradition* is very strong in a particular individual, there is a high probability that this individual has a positive attitude towards things considered old-style.

Users are more willing to share their attitude towards work tasks than their personal values. Although the preference for sharing personal information varies from user to user, the willingness or reluctance to reveal personal information depends on the type of information to be shared. During preliminary interviews we found users to be reluctant to reveal personal information such as personal values. However, they were openly talking about what they liked and what they disliked about their work and their attitude towards individual tasks. Approximating personal values through attitudes towards work tasks might be feasible without strong privacy concerns, but not as accurate as directly measuring values.

Groups of users normally share a context and can in this regard be seen as a homogeneous population. A group of users might work in the same work environment and perform similar tasks as part of their job—this is for example the case with the user group of nurses working in a hospital. Another shared context exists within an online community, where the members of the community perform similar actions on the website as part of their visit—imagine, for example users of a discussion forum or an action game. A third shared context might exist in other communities of which humans are part of in their private life—imagine, for example a specific church and the members of the congregation. All these members of a community share a specific context and perform tasks within this context—they have a shared understanding of actions needed to perform each of these tasks. For example,

if nurses administer a medication through a syringe they have a shared understanding of what steps are needed, how difficult it is, and what complications might occur during the task. The shared environment of work probably provides the largest basis of shared tasks compared to online and other private communities as people tend to fulfill more or less complex work roles and spend a considerable amount of time during work.

We suggest that the shared understanding of tasks and a user’s preference for tasks can be used to approximate values within a group of users which share a task context. As described in the Background Section, among others, personal values influence the user’s preferences for performing different tasks. The shared understanding and experience with tasks allows us to use short descriptions of tasks to evoke mental models of tasks in each user—and assure that they are imagining the same kind of task. In this way, mentioning a specific task will make the user recall uniformly which task is meant and make sure that other users—when asked questions about this task—think about the same task. Although attitudes towards work tasks are influenced by other factors, such as the nature of tasks or devices a task is performed with, we pursue the hypothesis that attitudes allow value approximation.

Individual users like or dislike specific tasks more than the average user—this task preference is partially related to the individuals’ dominant personal value, although all other factors outlined above are also present. When asked about their preference of performing a task, some users will like this task better than the average user—and other users will favor this task less than the average user. As such, the user’s preferences or disliking can be used—in comparison with the other users’ preference—to approximate the dominant value of the user. An example of a task which a nurse with the dominant value of *benevolence* might like more than the average could be the task of helping the patient with everyday tasks. In the following, we describe our method to approximate values through task preferences in a shared context.

#### Steps of the proposed method

In order to use preferences for or rejections of tasks to assign a tailored interface to a user, we need to do one major investigation in a research setting (the research process) which delivers a leaner method for the practical context (the application process). We first describe how values and tasks are associated with each other in the effort intensive research process. Subsequently we sketch how the achieved results can be used in a much less effort intensive application process.



*Research process of the method: Identification of value-dependent key tasks in a shared context*

In the first step of the research process, the researcher identifies tasks which suggest themselves as candidates to be value-dependent. Knowledge of Schwartz' value system informs this process. First, a researcher compiles a list of common tasks which are performed by most participants in a shared context. During this step, many tasks are identified, of which only a few will be suitable to approximate values during the later application process of our method. A broad range of methods can be used to generate such task lists. In our case study, we used a combination of literature review, expert consultation, and semi-structured interviews; however, using one of the methods might be sufficient. The literature review provides a list of commonly performed tasks in the domain. Tasks can be validated and enhanced through expert consultation and semi-structured interviews with users and a task list can be derived through these methods, if not available in the literature. Then, only tasks for which the researcher assumes that they relate to one or more of the properties of personal values (e.g., as described by Schwartz, see Table 2) are kept—the other tasks, for which no value dependency is assumed, are removed.

The second step of this process has the goal to determine which of the tasks can actually predict dominant values. These value-predicting tasks are called *key tasks*. To identify these key tasks, a sufficiently large sample of users answers a questionnaire with their preference for each of the previously identified potentially value-dependent tasks on Likert scales where 1 = high preference ... 6 = low preference. Additionally, each user completes a standardized questionnaire to measure personal values, e.g. the PVQ (Schwartz et al. 2001). Then, for each user, the dominant value is determined by the evaluation procedure of the standardized questionnaire, and users are grouped according to their dominant values. For example, value-

related groups of users with *benevolence*, *power*, and *security* could be created. By using a *t* test, the mean preference for each of the potential key tasks is then compared among users of each value-related group and all other users. For example, the mean preference for the task of writing reports is compared among the users with dominantly *Benevolence* versus *all the other users*. This procedure is repeated for all the other tasks and value-related groups. If for a group with a specific value a task is clearly significantly (2-sided *t* test,  $\alpha = 0.05$ ) preferred or disliked compared to the rest of the users, this task is called *key task* and can be used to approximate the value. A table can be used to record such key task-value relationships and the mean rating for the key tasks. Table 6 on page 27 shows key task-value relationships which were determined in our case study.

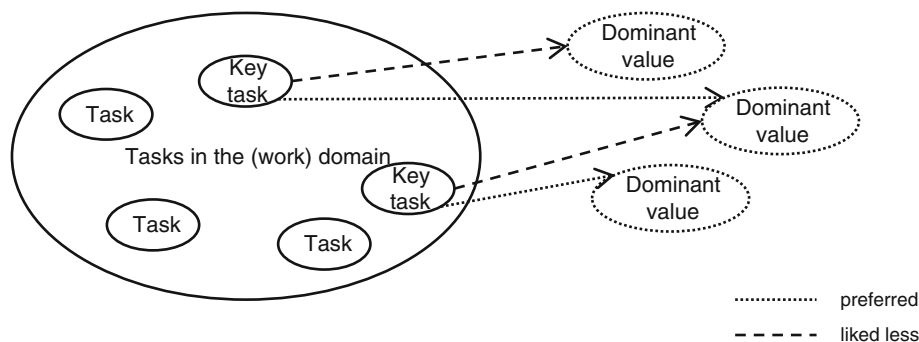
We suggest that key tasks are only determined for value-related groups which are sufficiently large (at least 30 members) to provide meaningful results. Figure 1 visualizes the intended use of key task-value relations to assign dominant values to individuals based on their preferences for key task. It should, however, be noted that the validity of this reverse association has not been shown independently in this investigation.

*Application process of the method: approximation of individual users' dominant personal values*

If a practitioner wants to approximate the personal value of a user (e.g., in order to tailor a user interface to the user's dominant value), the previously determined key task-value relationship is used.

First, the individual user indicates his or her preference for each of the key tasks by using a questionnaire with identical Likert scales as in the research process of the method.

Then the individual values of this user are computed as follows (preferably with a tool): the individual user's



**Fig. 1** From a pool of common tasks in the domain, individual key tasks are determined which have a positive or negative relationship to dominant personal values in the population. The direction of the

arrows denotes how we use answers to task-related questions to approximate users' values. It should be noted that the direction of mental "equipment" and its effects is the other way round.

ratings are centered to deal with an individual's tendency to rate all tasks very high or to rate all tasks very low. To center the individual users' ratings, the mean task rating of all users in the research process of the method is used to adjust the task ratings of the individual user.

Finally, the practitioner determines for each key task, whether the individual's (centered) rating is higher or lower as compared to the tasks mean rating in the key task-value relationship table. If preference or disliking of several key tasks point to the same value, the respective dominant personal value can be approximated. For example, a user dislikes the first key task (centered Likert score = 3.1), and prefers the second (1.7) and the third key task (1.9) where key tasks 1, 2 and 3 all point to the value *power*. The best approximation for his or her dominant value would be *Power* because the majority vote of key tasks point to this value even though the first key task does not. No value approximation can be determined if the user's preferences of key tasks equally point to multiple values, or if no specific difference in task preference can be determined.

We suggest determining dominant values of a user only by a key task-value relationship which was established in a comparable context. For example, a key task-value relationship which was established for nurses could be applied to nurses at another hospital, but not to physicians.

A case study in the following section exemplifies the research process of the method and the identification of value-dependent key tasks.

#### Case study using our method for value approximation

A case study was conducted to show that our proposed method of value approximation can be used to actually identify key tasks which approximate personal values. The study focuses on the research process of the method and shows how key tasks are identified in a shared context. The application process of the method to approximate an individual users' dominant value was not part of the case study.

Our research question was: does our research process allow selecting key tasks suited to approximate user values?

#### Set-up

In the following, we describe the questionnaire development and realization, and give details about its parts. These parts are used to measure personal values, nursing task preferences, and demographics.

*Questionnaire development and realization* Questionnaires were set up which included a total of 179 multiple choice questions of which 95 are relevant to this paper: 40

questions to measure personal values, 45 about common nursing tasks, 10 demographical questions. Due to the high prevalence of female participants, questions were asked in female voice. The questions were developed according to questionnaire guidelines for nurses (Soeken 2004). The questionnaires were tested by 5 medical informatics students and nurses to improve pertinence, relevance and understandability of the questions. The questionnaires were implemented using the Evasys survey system (Electric Paper, Lüneburg, Germany), printed on 10 pages, and later scanned for automatic evaluation. The paper version was selected to potentially increase the return rate by simplifying the questionnaire completion in multiple sessions, so participants could better deal with interruptions. This also ensured that computers at the wards were kept available for patient care, and nurses could easily complete questionnaires at any place of their liking, for example at home. We also wanted to avoid an interaction bias of the kind that the return rate was higher from computer savvy as opposed to computer critical staff.

*Personal values* To determine the personal values of the participants, the portraits values questionnaire PVQ (Schwartz et al. 2001) was used. The questionnaire was developed and validated by Shalom Schwartz, and translated to and validated for German (Hinz et al. 2002). It asks the participants to judge different personal situations and descriptions. It consists of 40 short descriptions of fictional persons and their value related properties in male and female versions. For example it includes for *Power*: "It is important to her to get respect from others. She wants people to do what she says." and for *Benevolence*: "It's very important to her to help the people around her. She wants to care for their well-being." Participants are asked to indicate how similar...not similar to each of the descriptions they think they are on 6-point Likert scales; additionally participants could select that they would not judge specific properties. In a long adaptive process performed by Schwartz and co-workers, questions were developed that equally characterize persons adhering to a certain value across various cultural backgrounds including developed and developing, democratic and totalitarian societies.

*Nursing tasks* We identified 45 common nursing tasks based on literature review and expert consultation, as well as on observation and semi-structured interviews with 8 nurses. This startup part of the questionnaire construction was conducted by one of the authors, who has 4 years of practical experience as a nurse. The questionnaires asked the participants to indicate their liking of each task on 6-point Likert scales ranging from "very much" to "not at all"; additionally the participants could select to not judge specific tasks.

**Table 4** Demographics of the participants

Demographics	Percentage
Gender	77 % female, 23 % male
Function	11 % nurse managers 89 % floor nurses
Workplace	50 % general ward 36 % intensive care unit 14 % functions/clinics
Hospitals	41 % surgery 21 % otorhinolaryngology, neurology, neurosurgery, radiology, and dentistry 28 % internal medicine 3 % woman's health 10 % orthopedics (3 % indicated multiple hospitals)
Years working in nursing	35 % up to 5 years 31 % up to 15 years 34 % more than 15 years
Age	30 % up to 25 years 33 % up to 35 years 16 % up to 45 years 17 % up to 55 years 4 % more than 55 years

*Demographics* 10 questions covered professional experience, computer use and preference, age, gender, and professional position (hospital, role, and kind of ward).

### Participants

265 of the original 1,400 questionnaires were returned, a return rate of 19 %. Of these questionnaires, 212 were usable for the analysis: they were completed by nurses as opposed to other hospital staff, adequately complete, and the PVQ part could be used to determine a dominant personal value of the participant. In our sample of 265, 23 persons differed from that pattern in having two (nearly) equally dominant values and were excluded.

The gender distribution of the participants was representative for the nurses' population at the hospital, though on average more males answered. The nurses' population consisted of 82 % female nurses and 18 % male nurses and 23 % males returned the questionnaires. Table 4 shows more demographics of the participants.

### Procedures

Nurse managers of 63 wards were contacted, and individually and personally visited by the first author. During the visit, they were informed about the aims, procedures, timeframe, potential incentives, and available support

during the study: participants had the option to call in during office hours, a presentation of the study was delivered upon request at regular ward meetings, and the nurse manager could receive a reminder of the study prior to meetings and the submission deadline. Nurse managers could then select the number of questionnaires they needed—dependent on the number of nurses working at their ward. Nurse managers were finally asked to distribute the questionnaires equally among nurses working in each shift and to explain the aims of the study to the participants. At wards where the nurse manager did not respond to attempts to schedule a visit to hand over the questionnaires, the investigators visited the ward without invitation and tried to schedule a meeting. One manager refused to schedule a meeting (in two phone calls, and one personal attempt). This ward was excluded from the study.

Return envelopes were enclosed with each questionnaire. The participants returned questionnaires individually by mail in closed envelopes addressed to the study office. The nurse manager did not know who returned a questionnaire. Each participant could opt for joining a lottery drawing, and wards with high overall return rates additionally received small rewards.

The overall time span from presenting the first questionnaires to closing the data collection was 55 days. Multiple calls were received at the study office: One caller asked a clarifying question about software features, multiple enquired about the deadline, and two had privacy concerns.

### Data analysis

The questionnaires were de-identified upon arrival by administrative assistants who were not part of the research team at the study office. The questionnaires were then scanned and answers were automatically processed and tabulated by the Evasys survey scanning system. A test with a subset of the questionnaires revealed that automatic scanning of the questionnaires resulted in accurate tabulation of the answers. Questionnaires which the system could not automatically process were manually entered by a research assistant.

*Personal values* For each remaining participant the dominant value was determined by applying the procedure as described by Schwartz to individual participants. The dominant personal value was determined to be the highest ranked value according to the procedure described by Schwartz. The participants where one value dominated were then assigned to the respective group. Three such groups made up the vast majority of our sample (see Table 5) and were further investigated.

**Tasks** For each key task the answers of the participants in one dominant value group were compared to the answers of all other participants (*t* test for independent samples, using SPSS 19.0.0, IBM, Armonk, New York). A total of three tests were performed for each task, one for each of the three dominant value groups, and participants with missing answers were excluded test by test. A 95 % confidence interval was used. Similar or different variance was determined through a Levene test and the appropriate *t* test statistic was used dependent on the variance.

**Table 5** Value distribution of the dominant values of the nurses

Dominant value	Number with this value	Percentage with this value (%)	Inclusion in our study
<i>Benevolence</i>	77	36	At least 30 members—used in our study
<i>Self-direction</i>	41	19	
<i>Hedonism</i>	34	16	Groups with infrequent values—not used to determine key tasks for value prediction in our study
<i>Universalism</i>	26	12	
<i>Security</i>	19	9	
<i>Stimulation</i>	8	4	
<i>Achievement</i>	3	1	
<i>Conformity</i>	4	2	
<i>Power</i>	0	0	
<i>Tradition</i>	0	0	

We focused on groups with dominant values with at least 30 members to assure sufficient validity

## Results

72 % of the participants had their dominant values in *benevolence*, *hedonism*, or *self-direction*, see Table 5. These were later used for analysis. To increase statistical validity, only participants with dominant values which occurred in at least 30 other participants were considered for the analysis. As a consequence, the data presented below cannot be regarded as representative for the whole nursing population in Germany and not even for the nursing population in the hospital analyzed. However, they are primarily meant to demonstrate the feasibility of the method. For that purpose it is appropriate to concentrate on a subset that delivers the largest amount of data.

14 tasks emerged which could be used to predict the three dominant values which occurred in the majority of the population, see Table 6. *Benevolent* nurses liked specific tasks better than the other nurses—they did not like any task less; *self-direction* and *hedonist* nurses both liked and disliked specific tasks compared to the others. The mean task rating of all participants calculated over all key tasks was 2.4.

*Benevolent* nurses preferred tasks connected to helping (helping the patient and colleagues), and talking to the patient and relatives (giving advice and talking to the patient, searching contact information of the patients' relatives) compared to nurses with other dominant values.

*Self-direction* nurses were found to like asking the physician for information and disliked tasks connected to

**Table 6** Results of the case study: Key tasks to predict the users' dominant personal value (1 = high preference ... 6 = low preference)

Task	<i>Benevolence</i>	<i>Self-direction</i>	<i>Hedonism</i>
Help the patient during everyday tasks (e.g. bathing, combing)	Preferred (2 vs. 2,4)		
Talk to the patient	Preferred (1,6 vs. 1,9)		
Provide advice to the patient	Preferred (1,8 vs. 2,1)		
Help colleagues (e.g. repositioning the patient)	Preferred (1,6 vs. 1,8)		
Search for contact information of patient's relatives	Preferred (2,4 vs. 2,8)		Liked less (3 vs. 2,6)
Query the patient (e.g. for symptoms)	Preferred (2 vs. 2,1)		Liked less (2,4 vs. 2)
Instruct the patient			Liked less (2,3 vs. 1,9)
Reposition patient			Liked less (2,8 vs. 2,4)
Mobilize the patient			Liked less (2,7 vs. 2,1)
Ask the physician for information		Preferred (1,9 vs. 2,3)	Liked less (2,5 vs. 2,1)
Order medication		Liked less (3,3 vs. 2,7)	Preferred (2,2 vs. 2,9)
Check medication stock		Liked less (4 vs. 2,9)	
Prepare medication		Liked less (2,9 vs. 2,4)	
Organize the ward (e.g. responsibilities of the nurses)		Liked less (3 vs. 2,4)	

For each group (e.g. Benevolence), the mean rating of the group is compared with the rating of the other participants (e.g. 1.6 vs. 1.9 for the task "Talk to the patient"). Only tasks with a statistically significant difference are included in the table

When comparing some of the value descriptions (see Table 2) to the tasks which were found to be relevant for these personal values, some relatedness can be seen

medication (ordering-, checking-, and preparing medication) and to organization of the ward.

*Hedonist* nurses preferred ordering medication, and disliked asking the physician for information, tasks with patient contact (querying-, instructing-, repositioning-, and mobilizing the patient) as well as searching for contact information of the patients relatives.

Overall, the case study could show that our proposed method to identify key tasks was applicable in a nursing context. We identified key tasks which relate to the personal values of nurses, and which would be used in the next step of the method during the application process to approximate the dominant value of individual users.

## Discussion

Through our case study we could show that tasks in a shared context can be used to approximate the user's

personal values. A return rate of 19 % is not outstanding and selection biases should be considered. However, with regard to the length of the questionnaire (179 items) the return rate seems to be acceptable and was sufficient to identify key tasks to approximate values. Also, the absolute number of 212 completely filled questionnaires allowed aggregating reasonably sized subgroups for statistical analysis. The questionnaire was presented in paper form. Therefore, biases towards technology savvy subjects are unlikely. Rather could there be biases towards older, more conservative subjects. This, however, does not show in the data. At least, the value "Tradition" is not dominant in any of the 212 participants. Therefore, selection biases, if present, would be on a more subtle, tacit level.

The findings of our case study intuitively make sense when comparing the characteristics of people with specific personal values and the implications of individual tasks, for example how effort-intensive they are, how much fun they provide, and which results performing a specific tasks will

**Table 7** Key tasks are compared to descriptions of values according to Schwartz PVQ value questions (Schwartz et al. 2001)

Value	Value description	Key task to predict the value
<i>Benevolence</i>	It's very important to him to help the people around him. He wants to care for their well-being	Help the patient with everyday tasks (e.g. bathing, combing) (preferred)
	It is important to him to be loyal to his friends. He wants to devote himself to people close to him	Help colleagues (e.g. repositioning the patient) (preferred)
	It is important to him to respond to the needs of others. He tries to support those he knows.	Provide advice to the patient (preferred)
	Forgiving people who have hurt him is important to him. He tries to see what is good in them and not to hold a grudge	No suitable task determined
<i>Self-direction</i>	Thinking up new ideas and being creative is important to him. He likes to do things in his own original way	Check medication stock (liked less) ( <i>Comment: checking medication stock is very repetitive</i> )
	He thinks it is important to do lots of different things in life. He always looks for new things to try	Order medication (liked less) ( <i>Comment: medication ordering might be boring as it needs a very structured approach</i> )
	It is important to him to make his own decisions about what he does. He likes to be free to plan and to choose his activities for himself	Organize the ward (e.g. responsibilities of the nurses) (liked less) ( <i>Comment: this task centers around planning for others</i> )
	He thinks it's important to be interested in things. He likes to be curious and to try to understand all sorts of things	Ask the physician for information (preferred)
<i>Hedonism</i>	It is important to him to be independent. He likes to rely on himself	Prepare medication (liked less) ( <i>Comment: medication preparation is very rule-based</i> )
	He seeks every chance he can to have fun. It is important to him to do things that give him pleasure.	Order medication (preferred) ( <i>Comment: medication ordering could be gratifying as it results in receiving a later delivery—comparable to a present</i> )
	Enjoying life's pleasures is important to him. He likes to 'spoil' himself	Mobilize the patient (liked less) ( <i>Comment: Mobilizing the patient is effortful and repetitive</i> )
	He really wants to enjoy life. Having a good time is very important to him	Ask the physician for information (liked less) ( <i>Comment: asking the physician for information will likely result in questions back and potentially additional tasks to do</i> )

Nurses with benevolence liked key tasks better and value descriptions sound comparable to key tasks. For self-direction and hedonism nurses liked key tasks less than the average and the value descriptions contradict the character of the tasks. Schwartz' original male formulation is used here, while our questionnaires were in female voice throughout



have, see Table 7. As such, tasks which tended to help the users' surroundings were preferred by users with *benevolence*. Tasks, which did not allow much personal freedom, were liked less by users with *self-direction*. Tasks which involved physical work, or might result in additional work, were liked less, and tasks which had some gratification were preferred by users with *hedonism* as compared to the average.

It is interesting to note that users with *self-direction* and *hedonism* liked the majority of key tasks less than the average, whereas users with *benevolence* liked key tasks more. These findings might as well imply that users with *benevolence* fit better in the nurse profession. A potential explanation for these findings could be a difference in attitudes in users with different values. Disliking a task might be triggered if performing the task violates value properties, but users might be indifferent towards the task if such properties are not violated.

The users' comments during the case study confirmed that a method to approximate values instead of asking for them directly would be beneficial, if used in large groups of users. Individual participants expressed privacy concerns related to the direct value measurement section during the research process of the method. If used in another context, the research process could still evoke refusal in individual participants. However, during the application process of the method no direct value measurements are performed and therefore no refusal would occur.

An interesting result of our study was that only 14 of the initially 45 tasks were found to be value-specific for the main three values. Potential explanations could be that tasks for all ten values were selected or that nurses need to perform many tasks that do not correspond to their primary value.

#### Comparison with other research

Our method to approximate personal values based on task preference has not been proposed previously. In our case study we could show that we could identify preferences for types of tasks related to the three most frequent personal values in nurses.

When comparing our personal values measured in nurses with other publications, our findings are primarily congruent to the values identified in nurses in other countries—but show differences compared to the values of the general population in Germany and to the values of German physicians. In nurses from Hungary, Italy, the United Kingdom, and the USA (Glazer and Beehr 2002), the nurses' dominant values were benevolence, self-direction, hedonism, and universalism.

Compared to the average population in Germany, our findings confirm the importance of *benevolence* and

*universalism* (Hinz et al. 2002), but do not replicate their findings of *security*. In our nurses' sample, *universalism* (12 %) was fourth in counting, not far behind *hedonism* (16 %), while *security* was the dominant value of only 9 % of our sample. We see two possible explanations for the difference in prevailing values. The nursing subpopulation of the German population may not be random but rather may that profession primarily draw universalists and hedonists while it rejects security driven individuals. Or we witness a change in generation: Hinz' German sample was taken a decade before our sample and prevailing values may have changed nationwide. On the other hand, within our nursing population the sample was very broad and presumably representative. Therefore, we believe that at least for nurses in academic medical centers the results are valid.

When comparing our approach to approximate values with the methods described in the Background section, we find that our research complements existing work. Additionally to the work of Ramos et al. (Ramos et al. 2005), who suggest a knowledge-based framework for the requirements engineering process without providing detail how such knowledge should be collected, we suggest a way to identify the users' dominant values by drawing on their work experience. Compared to the approach offered by Sarah Thew which relies on soft issues measurements based on interviews (Thew and Sutcliffe 2008, 2011), we suggest an approach which offers better performance in situations where a large number of users and privacy issues needs to be considered. Compared to the very broad definition of "values" used in the field of value-sensitive design e.g., by (Friedman 1997), our understanding of "values" centers on personal values. Compared to our work presented at the Workshop on Values in Design (Koch et al. 2011), this paper describes the approach in more detail and proves that it is feasible by providing a concrete set of questionnaire items suitable to distinguish nurses as to their dominant value *benevolence*, *self-direction* or *hedonism*.

#### Implications of our findings

##### *Interface development for nurses*

Our findings that nurses had different values compared to other occupational groups such as physicians imply that specific interfaces for nurses are needed. Although this demand is not new, the reasoning based on the nurses' differences in personal values is novel. The demand is seconded by other research based on other reasoning, for example depending on the nurses' expertise (Cho et al. 2010).

## Interface tailoring in general

Our findings that users with different values prefer different kinds of tasks, might imply that individualized interfaces might give an advantage to the specific user. Such interfaces should, however, allow the possibility for each user to opt out of the tailoring to meet potential privacy demands, stigmatizing due to the exposed value, and fears of potential missing information and functionality based on the individualization as discussed above.

## Value approximation

Our findings that users with different dominant values showed a value dependent preference for different tasks could be used to approximate values based on such dependencies. To increase the chances of success of the method, it is essential to select an appropriate set of key task candidates as discussed above. However, the research process requires a considerable effort. Therefore, such tailoring is only feasible, if interface tailoring for very large groups of users is planned.

One ethical dilemma needs to be mentioned. The strategic idea behind the presented approach is

- That knowing users' personal values is a worthwhile long term investment to design software to their satisfaction,
- That users feel concerned to answer Schwartz' (Schwartz et al. 2001) personal questions to help reveal their values,
- That instead we ask proxy questions about work task which users are not concerned about,
- That we are still probing for values, but covertly.

If users are willing to reveal their personal values but feel concerned about Schwartz' questions the dilemma does not show. If, however, users feel concerned about their personal values becoming known, by whatsoever method, the dilemma is there: We are cheating the persons we ask unless we inform them that, in order to find out their values, we ask them about attitudes towards tasks. If, however, we inform them that we will use their task preferences to find out their personal values, this may affect the results, because some subjects are now equally concerned about work task-related questions. To study such confounding behavior a human subjects Internal Review Board approved investigation should be made, where subjects are randomized to one of two experimental settings: informed versus not informed about the fact that work task related questions are used to find out personal values. This is one of the few unique settings where seeking informed consent must be waived, because

informed consent would centrally interfere with the research question.

## Limitations

The generalizability of our study is limited. Our study focused on a single user group with a shared context: nurses in an individual academic hospital. Therefore, our findings of the nurses' task preferences cannot be generalized to the nursing population in general or to other professions. Threats to validity include that nurses' values were self-reported. This may have induced a selection bias if, for example users with security-related values might not have responded to our survey out of privacy concerns and especially nurses with the value *hedonism*, who could have been keen on the incentive, might have responded. But as we know through the seminal work of Rosenthal (1969), newly published in 2009 volunteering and volunteer subjects are full of surprises. Together with appealing results about the effect of incentives on volunteering we also find just the opposite effect in other studies. So, in our investigation it may as well be the case that the hedonistic individuals found the moderate incentive an insufficient stimulus and would rather not respond and the benevolent individuals might prefer being benevolent towards their patients over being benevolent towards the researchers, i.e. both form an even bigger part of the nursing workforce than shows in our data.

Furthermore, users might have not told the truth when answering value measurement questions to keep their personal characteristics to themselves. Therefore, our results might not be valid for the nursing population in general. Task specific judgments may be affected through the way the tasks are implemented in different clinical context, e.g. in internal medicine versus surgery. However, as our sample included nurses from a broad range of clinical context we do not think that these judgments would significantly influence our results concerning personal values measured over the whole population.

## Future directions

Future research could replicate our method with users from another shared context to verify, whether it can identify key tasks which can approximate personal values in that environment as well. Such studies could focus on users with a different profession, nurses at a different hospital, or users with another shared context such as the same social group or online forum. Furthermore, studies could focus on users with other values than the three values our empirical study focused on. Additionally, research is needed to identify value specific interface variants based on their diverse needs for information and functionality.

## Conclusions

We showed that in a well-defined setting of the shared work context of nurses in an academic hospital dominant personal values can be approximated by having subjects mark their preferences or disliking for a few selected work task items. To apply the approach with other target groups our process will have to be replicated by generating target group pertinent items and selecting those key items that relate to dominant values as of the approved PVQ value classification. Wherever such context-dependent sets of key items have emerged from a high effort high privacy exposure investigation, they can be used for low effort low privacy exposure user classification and subsequent generation of user value-tailored software requirements.

**Acknowledgments** We would like to acknowledge the funding of the Deutsche Forschungsgemeinschaft (DFG), grants PA 1387/3-1 and WE 2467/7-1, and the help of the participating nurses, nurse managers, the Employee Committee, the Director of Nursing at the Heidelberg University, Doris Keidel-Müller and Steffi Schiemang.

## References

- Albarracín, D., Wallace, H., & Glassman, L. (2004). Survival and change in judgments: a model of activation and comparison. *Advances in Experimental Social Psychology* 36.
- Bandura, A. (1997). *The exercise of control*. New York: Freeman.
- Biffi, S., Aurum, A., Boehm, B., Erdogmus, H., & Grünbacher, P. (Eds.). (2006). *Value-based software engineering*. New York: Springer.
- Cho, I., Stagers, N., & Park, I. (2010). Nurses' responses to differing amounts and information content in a diagnostic computer-based decision support application. *Computers, informatics, nursing: CIN*, 28(2), 95–102. doi:10.1097/NCN.0b013e3181cd8240.
- Custers, R., & Aarts, H. (2010). The unconscious will: How the pursuit of goals operates outside of conscious awareness. *Science*, 329(5987), 47–50.
- Detweiler, C., Pommeranz, A., & Jonker, C. (2011a). *Personal informatics for reflection on personal values*. Paper presented at the CHI'11 workshop on "Personal Informatics & HCI: Design, Theory, & Social Implications, Toronto, Canada.
- Dorn, K., Hewett, A. J., & Ukis, V. (2009). Method of providing tailor-made software for hospital departments. In U. S. P. Application (Ed.), (Vol. 20100179826). USA.
- Fishbein, M. (2005). The influence of attitudes on behavior. *The Handbook of Attitudes*, 173.
- Flanagan, M., Howe, D., & Nissenbaum, H. (2008). Embodying values in technology: Theory and practice. In J. v. d Hoven (Ed.), *Information technology and moral philosophy* (pp. 322–353). Cambridge: Cambridge University Press.
- Friedman, B. (1997). Social judgments and technological innovation: Adolescents' understanding of property, privacy, and electronic information. *Computers in Human Behavior*, 13(3), 327–351.
- Friedman, B., & Freier, N. (2005). Value sensitive design. In K. Fisher, & S. Erdelez (Eds.), *Theories of information behavior: A researcher's guide*. (pp. 368–372), Medford.
- Friedman, B., Kahn, P. H, Jr, & Borning, A. (2006). Value sensitive design and information systems. In P. Zhang & D. Galletta (Eds.), *Human-computer interaction in management information systems: Foundations* (pp. 348–372). New York: Armonk.
- Glazer, S., & Beehr, T. A. (2002). Similarities and differences in human values between nurses in four countries. *International Journal of Cross Cultural Management*, 2(2), 185–202. doi:10.1177/1470595802002002867.
- Hinz, A., Albani, C., Gießler, A., & Brähler, E. (2002). Welche Werte sind den Deutschen etwas wert? Ergebnisse einer repräsentativen Umfrage. *Psychosozial*, 25(2), 21–30.
- Kobayashi, I., & Tsunawaki, T. (2007). Intelligent help system that synchronizes application software operation: Toward intelligent tailor-made information provision. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 11(10), 1216–1223.
- Koch, S. H., Proynova, R., Paech, B., & Wetter, T. (2011). Approximating user values to preserve privacy - a proposal. Paper presented at the 1st International Workshop on Values in Design - Building Bridges between RE, HCI and Ethics, Lisbon (p. 32).
- Li, I., Dey, A. and Forlizzi, J. (2010) A stage-based model of personal informatics systems. In *Proc. CHI'10*, ACM Press. (pp. 557–566).
- Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation—a 35-year Odyssey. *American Psychologist*, 57(9), 705–717. doi:10.1037//0003-066X.57.9.705.
- Meligno, B. M., & Ravlin, E. C. (1998). Individual values in organizations: Concepts, controversies, and research. *Journal of Management Information Systems*, 24(3), 351–389.
- Nissenbaum, H. (2005). Values in technical design. In C. Mitcham (Ed.), *Encyclopedia of science, technology and society*. New York: MacMillan.
- Oesterreich, R. (1981). *Handlungsregulation und Kontrolle*. München, Wien, Baltimore: Urban und Schwarzenberg.
- O'Reilly, C. A., Chatman, J., & Caldwell, D. F. (1991). People and organizational culture: A profile comparison approach to assessing person-organization fit. *Academy of Management Journal*, 34(3), 487–516.
- Orne, M. T. (1969). Demand characteristics and the concept of quasi-controls. In R. Rosenthal & R. L. Rosnow (Eds.), *Artifact in behavioral research* (pp. 143–179). New York: Academic Press.
- Pommeranz, A., Detweiler, C., Wiggers, P., & Jonker, C. M. (2011a). *Self-reflection on personal values to support value-sensitive design*. Newcastle: In Forthcoming in BCS HCI work in progress track.
- Pommeranz, A., Detweiler, C., Wiggers, P., & Jonker, C. M. (2011c). Elicitation of situated values: need for tools to help stakeholders and designers to reflect and communicate. In *Ethics and information technology, special issue on requirements engineering* (pp. 1–19). Berlin: Springer. doi:10.1007/s10676-011-9282-6.
- Proynova, R., Wicht, A., Paech, B., & Wetter, T. (2010). Use of personal values in requirements engineering - a research preview. Paper presented at the 16th International Working Conference on Requirements Engineering: Foundation for Software Quality, Essen, Germany.
- Proynova, R., Paech, B., Koch, S. H., Wicht, A., & Wetter, T. (2011). Investigating the influence of personal values on requirements for health care information systems. Paper presented at the SEHC '11 Proceeding of the 3rd workshop on Software engineering in health care. ACM New York (pp. 48–55). doi:10.1145/1987993.1988004.
- Ramos, I., & Berry, D. M. (2005). Is emotion relevant to requirements engineering? *Requirements Engineering*, 10(3), 238–242. doi:10.1007/s00766-005-0014-5.
- Ramos, I., Berry, D. M., & Carvalho, J. (2005). Requirements engineering for organizational transformation. *Information and Software Technology*, 47(7), 479–495.

- Rosenthal, R., & R. L. Rosnow (1969). In R. Rosenthal, R. L. Rosnow (Eds.), *The volunteer subject. Artifact in behavioral research* (pp. 59–118). Academic Press.
- Schwartz, S. H. (Ed.). (1992). *Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries (advances in experimental social psychology)*. San Diego: Academic Press.
- Schwartz, S. H., Melech, G., Lehmann, A., Burgess, S., Harris, M., & Owens, V. (2001). Extending the cross-cultural validity of the theory of basic human values with a different method of measurement. *Journal of Cross-Cultural Psychology*, 32(5), 519–542.
- Soeken, K. L. (2004). Guidelines for writing multiple-choice items. In C. F. Waltz, O. L. Strickland, & E. R. Lenz (Eds.), *Measurement in nursing and health research* (3rd ed.). New York, NY: Springer.
- Sommerville, I. (1981). *Providing the user with a tailor-made interface*. Paper presented at the Sixth ACM Europ. Regional Conf, London; Engl 30 Mar.-1 Apr. 1981.
- Tesser, A., & Schwarz, N. (Eds.). (2003). *Intraindividual process (1 ed., Vol. 1, Blackwell handbook of social psychology, Vol. 1)*. Malden: Blackwell.
- Thew, S., & Sutcliffe, A. (2008). Value-based requirements engineering. *REFSQ 2008*. Barcelona.
- Thew, S., & Sutcliffe, A. (2011). *Elicitation of values, motivations and emotions: The VBRE Method*. Paper presented at the 1st International workshop on values in design—building bridges between RE, HCI and Ethics, Lisbon.
- Thong, J. Y. L., Hong, W., & Tam, K. Y. (2004). What leads to user acceptance of digital libraries? *Communications of the ACM*, 47(11), 78–83.