Designing Human-Centered Systems for Reflective Decision Making

Designing Human-Centered Systems for Reflective Decision Making

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op 7 september 2012 om 10.00 uur
door

Alina POMMERANZ

Master of Science in Interactive Systems Engineering, Royal Institute of Technology Stockholm

geboren te Bochum, Duitsland.

Dit proefschrift is goedgekeurd door de promotor:

Prof. dr. C.M. Jonker

Copromotor:

Dr. ir. P. Wiggers

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. C.M. Jonker,	Technische Universiteit Delft, promotor
Dr. ir. P. Wiggers,	Technische Universiteit Delft, copromotor
Prof. dr. M.A. Neerincx,	Technische Universiteit Delft
Prof. dr. M.J. van den Hoven,	Technische Universiteit Delft
Prof. dr. B. Friedman,	University of Washington
Prof. dr. K. Höök,	Royal Institute of Technology Stockholm
Prof. dr. A. Sutcliffe,	Manchester University
Prof. dr. I. Heynderickx,	Technische Universiteit Delft, reservelid

SIKS Dissertation Series No. 2012-30.

The research reported in this thesis has been carried out under the auspices of SIKS, the Dutch Research School for Information and Knowledge Systems.

This research is supported by the Dutch Technology Foundation STW, applied science division of NWO and the Technology Program of the Ministry of Economic Affairs. It is part of the Pocket Negotiator project with grant number VICI-project 08075.

ISBN 978-94-61913-65-4

In memory of Käthe Pommeranz.
To my parents, Anne and Jürgen Pommeranz.

"Only a life lived for others is a life worthwhile." Albert Einstein (1879-1955)

Contents

Acknowledgments

Throughout the past four years many people have contributed directly and indirectly to my work and made this thesis possible. I may not be able to mention every one of you in person, but I want you to know that I am grateful for all the support I received.

First of all I would like to thank my daily advisor Pascal Wiggers. Thank you for all the interesting discussions, co-authoring papers, helping me find my research direction, always being available and the overall support. Special thanks also go to my promoter Catholijn Jonker, who convinced me that I can reach for the top and always supported my personal development and career ambitions.

I would also like to thank the Pocket Negotiator colleagues with whom I collaborated closely, such as Joost Broekens, Aisling O'Kane and Ulas Ulgen. Thanks also to the rest of the Pocket Negotiator team, including Dmytro Tykhonov, Tim Baarslag, Iris van de Kieft, Wietske Visser, Willem-Paul Brinkman, Birna van Riemsdijk and Koen Hindriks as well as the user committee of the project, especially Mark van Gurp for his valuable feedback on the negotiation scenarios and providing study participants. Special thanks go to Christian Detweiler. It was inspiring and fun to work with you, including the organization of conference workshops and the fun colloquium.

I also thank the rest of the II team, especially Judith Redi for the fun times we had at work and in our free-time and for being my paranymph. Furthermore, thanks to Hani Alers, Hantao Liu, Zhenke Yang, Nike Gunawan and Maaike Harbers for sharing breaks, sports and other fun activities to cheer up our PhD times. Thanks also to Harold Nefs who helped me with parts of my data analysis. I am also grateful for the technical support I received from Bart Vastenhouw and Ruud de Jong, and the support with paperwork from Anita Hoogmoed, who are always positive and helpful.

Support also came from outside the TU Delft. First of all, I would like to thank Petra

Sundström and Kia Höök. Without you I would have never gotten excited about research and done a PhD. I also thank all other Mobile Life and SICS colleagues, especially Baki Cakiçi, Pedro Sanches, Pedro Ferreira and Eric-Oluf Svee, who I stayed in touch with and received feedback on my work from.

Great thanks also go to all my colleagues at the University of Washington and the Value Sensitive Design lab. I had a great time in Seattle thanks to all of you. Batya Friedman deserves special thanks for her close coaching and support. I also enjoyed collaborating and co-authoring with Dave Hendry, Daisy Yoo and Jill Woelfer. Seattle would never have been so much fun without Abi Evans, Kathie Derthick and Chris Hearney. Thank you all!

I thank Liam Bannon and Alistair Sutcliffe for giving me feedback on my work, Maaike Belien and Wolfgang Engels for proof reading parts of my thesis and Eline Hoorweg for fixing my Dutch summary and all other help she offered throughout the process. Thanks also go to Mario Santos-Ramos for helping me with the thesis cover.

Finally, I would like to thank all my friends and family, who have supported me and kept me sane during the PhD process. Thank you for bearing with me while I got absorbed in work and did not keep in touch as much as I should have. Special thanks go to Kathrin Woitha, my dearest friend for more than 25 years of my life. Sharing all the downsides and challenges of research and cultural challenges in the Netherlands with you made life easier here.

Of course, I would also like to thank my parents, Anne and Jürgen Pommeranz, without whom this thesis would have never been possible. Thank you for raising me to be open-minded, ambitious and social, for allowing me to get a good education, for bearing with me whenever I went my own way and for supporting me with your love all along. Big thanks also to my brother, Robin Pommeranz, and my sister-in-law, Sandra Pommeranz, for being there for me and taking part in my studies.

A very special thanks goes to Fredrik Huldtgren for being the most patient person I know, for listening to my endless monologues about work, for his support, encouragement and positive attitude, for always wanting to make me feel good and, above all, for loving me.

CHAPTER 1

THESIS INTRODUCTION

This chapter introduces the main problem, research questions and an outline of the dissertation.

If we knew what it was we were doing, it
would not be called research, would it?
Albert Einstein (1879-1955)

1.1 Introduction

Taking major life decisions, e.g. where to live or what career to follow, is difficult and sometimes emotional. Besides finding out what exactly one wants for oneself, part of the decision making process consists of considering the long-term consequences of the decisions and being empathetic for loved ones affected by the decisions. Furthermore, such life decisions can involve negotiations with another party, e.g. the seller of one's desired house or one's future boss. Especially in the latter case an important aspect of the decision making process is to manage existing relationships or building up future relationships with the other party. Generally, it is important to find an outcome that is satisfying for everyone involved.

Besides these *soft issues*, decision making deals with establishing and browsing different alternatives as well as weighing options according to one's preferences. Depending on the number of issues to be decided upon in a given domain, the alternatives space can be vast, difficult to oversee and easily lead to an information overload. This difficulty is even greater when people are not familiar with the issues in the given domain. Furthermore, balancing one's preferences to come to a good decision is difficult. People may focus, e.g., too much on certain aspects and fail to see options that are better overall. Because of these challenges combined with the soft issues named above, few people are effective at taking decisions and negotiating (Thomson, 2005).

1.2 Decision Support Systems

Fortunately, there can be a remedy to the problem outlined above. Computers are good at handling vast amounts of data, browsing through the data and calculating options fit to the user's preferences within seconds. With the advance of knowledge-based systems and computational intelligence even complex decision making tasks can be done by computers. This has led to the development of decision support systems (DSS) in different domains, such as management science, e-commerce and artificial intelligence (Kersten, 2007; Rangaswamy and Shell, 1997; Schoop et al, 2001; Vetschera et al, 2006). It has been shown that in constrained settings, i.e. a defined domain and an accurate user preference profile, a DSS can significantly improve the human performance in decision processes and, in the specific case of negotiation with another party, increase the number of win-win outcomes (Kersten and Lo, 2003; Bosse and Jonker, 2005). However, current DSS have several drawbacks.

First, the majority of current DSS are built as analytic tools for highly-complex domains, among others medical decision support (Fieschi et al, 2003) or disaster management (Fiedrich and Burghardtl, 2007). Furthermore, the focus lies on domain experts as users (Bellucini, 2006). Systems to support untrained decision makers mainly exist in the form of recommender systems for everyday choices, such as what to wear (Shen, 2007), what movie to watch (Miller, 2003) or what to buy in an online shop (Stolze, 2003). To our knowledge few systems focus on decision support for common people when decision outcomes and long-term consequences are crucial in a person's life. One exception is Choice Point, a tool that helps people to take financial decisions with regard to long term life changes (Fano and Kurth, 2003).

Designing systems for common people that face difficult decisions and negotiations in their lives could greatly improve people's decision making in life choices and lead to more satisfactory decision making processes and outcomes. This would allow a broad population of non-expert users to make use of intelligent decision technology. In order to design for a broad range of people, a human-centered approach is needed. However, current research focuses on technical solutions and has hardly considered the human and the needed cross-disciplinary perspective to account for human cognitive and emotional processes.

Second, the prevailing use of economic models to represent user preferences in current systems is a problem. Current systems developed based on economic models neglect the fact that human decision making does not always follow a rational path. People often do not know what they actually want and how to get there – as wrongly assumed by such models –, but construct their preferences and adapt them as they go along in the decision making process (Payne, 1999). Preferences over concrete issues in a domain (e.g. the size of a house, the salary for a job) are unstable and can easily change, while underlying values (i.e. what is generally important in life to someone) are more stable. However, the majority of current systems do not allow the user to enter underlying values and they do not support the reflection process needed for people to assess their values and preferences. Following from the inaccurate use of rational models the mismatch of preference representation and understanding between the user and system is another problem. Current systems use long lists of questions and utility-based preference elicitation. No system exists yet that is intuitive to use, and has the flexibility to deal with new domains. This can result in inaccurate preference models inside the system and ultimately incomprehensible negotiation or decision advice given by the system.

Third, in many cases, system input and output cannot be scrutinized by the users.

This does not allow the user to gain understanding of (1) why the system asks for certain input, and (2) the user model created by the system and used for decision advice (output).

Altogether, these problems can lead to reduced trust into the systems' capabilities and, thus, low user acceptance of DSS.

1.3 Human-System Interaction

Given the drawbacks of current DSS outlined above, we believe that research and development of DSS needs to focus on the interaction between users and the system.

Humans and computers have, to some extent, complementary capabilities for decision making. Humans are better equipped to understand real life contexts and emotional fluctuations which may influence the decision process. They are capable of finding new relations between concepts, and they have the necessary background knowledge to interpret the decision domain with respect to their own preferences. On the other hand, people can have problems handling their emotions and the complexity of outcome spaces in decision making. Computers provide computational power, data storage and search techniques to handle those spaces. However, they may have problems with handling the huge amount of background and context knowledge necessary to cope with, and understand arbitrary conversations and problems.

It is assumed that combining the strength and mitigating the weaknesses of computers and humans can lead to better decision outcomes (Hindriks, 2008). One important aspect to reach this goal is to design the interaction between the user and the system in a way that is comprehensible to both and allows for collaboration. A challenge of designing interaction for decision support is that often there is not one single right solution to a given (decision making) problem. Take, e.g., deciding a career path and in specific a job negotiation; even if the user was well aware of his or her preferences and entered them correctly into the system the outcome that could be reached would be based also on the other party's preferences and constraints (e.g. the maximum amount of salary the employer can offer) and might be suboptimal according to the user's actual preferences. Furthermore, emotions may influence the process in a way that cannot be accounted for in the system. These are just two examples of the complexity. Psychological processes may play a role in a decision context that can hardly be understood by a computer system. It is thus important that the user can scrutinize the advice from the system, and that the system can in turn ask for more input that may be missing in order to give accurate advice (i.e. lack of context specific

knowledge, e.g. constraints or emotions).

Current systems employ an interaction style based on a black box principle, according to which the user enters required input in a way that fits the system's internal representations; the system then does its calculations and offers the best output based on the given input. This type of interaction leaves little room for humans to feel in control of the decision making process and comprehend what specific information the system's advice was based on.

Furthermore, there is a lack of knowledge about what type of interaction and support is most successful, especially for untrained decision makers who may be insecure and not familiar with a domain. The user-system interaction will have a great influence on how a person or a group of people perceive the decision support system and on the acceptance of decision support systems – an issue under-explored in existing research.

1.4 Research Objectives

The overall aim of this dissertation is to investigate the crucial factors to design human-centered DSS that can be used by untrained decision makers. By putting an emphasis on human-centeredness we would like to highlight a shift from DSS as knowledge-based systems that take over decision making from the users to systems that empower people to take their own decisions. The dissertation's main research question is formulated as follows.

"How can we design user-system interaction for human-centered decision support?"

When talking about design, two main subquestions are relevant. (1) What are we going to design? (2) How are we going to design it?

The first question – the "what" – is of conceptual nature and aims at defining the design space in detail. The precise context in which a DSS would be used, its functionality, design stance (e.g. a persuasive system or a advisory system) and target users still have to be defined. The project in which this dissertation was carried out, i.e. the Pocket Negotiator Project (Hindriks, 2008), defined the type of system and its target users, namely a negotiation support system (NSS) and untrained negotiators respectively. Another given constraint was that the system would be a personal system for one party only instead of a mediating system. In this sense it can be seen as a DSS specialized to support the decision making of one party in a negotiation with added functionality to add opponent preferences and bidding support. Besides the given

constraints, the following questions remained:

1. What functionality is crucial for a NSS from an expert perspective to overcome typical problems in negotiations?
2. What are the needs of end-users with respect to a NSS?
3. In which social situations would people accept the use of a NSS?

Resulting from an investigation of these three questions we found that crucial to the success of DSS, or in particular NSS, is a good preparation in which a decision maker constructs a profile of her preferences, which are in turn influenced by underlying values. Therefore, we narrowed our research focus in the subsequent studies to answer the following questions with regard to preference elicitation:

4. How do people create preferences?
5. What preference elicitation methods exist?
6. What ways do people prefer to express their preferences in interfaces?
7. How can interfaces be designed to fit the user's expression of preferences?

In addition, we explored the following questions with regard to value elicitation:

8. Why are values difficult to assess?
9. How do experts support people in assessing their values?
10. How can we design tools that help people reflect on their values?

The second question – the "how" – is of methodological nature. With regard to human-centeredness, the main focus lies on designing systems that take human decision making capabilities into account and support these in order to arrive at better decisions. Furthermore, to design appropriate interaction between the system and the user, the design process itself needs to be human-centered, i.e. placing the human in the center of design activities. To be more specific, we investigate the role of different stakeholders at different points in the design process and how we can

enhance engagement in the design process of these stakeholders, who are often not designers.

To summarize, the research questions with regard to the methodological aspect were:

11. Which design and research methods are useful in the design of human-centered DSS?

12. How can we involve end-users and domain experts actively in the design process?

1.5 Dissertation Outline

An overview of this dissertation is presented in Figure 1.1. In accordance with our human-centered viewpoint the following chapter provides an overview of perspectives of human-system interaction and approaches to interaction design. In particular, we outline our participatory design viewpoint and present design methods we used and adapted from user-centered design (UCD) and participatory design (PD). We employed methods of quantitative and qualitative nature focusing both on expert and user knowledge and needs to get a holistic view of the issues at stake.

In the remainder of the dissertation we present our empirical research and design work as well as a reflection on the methods we employed and developed along the way. Chapter 3 (based on previous publications [4,9,12]; see publication list at the end of this chapter) presents an investigation of the design space and success criteria for DSS, in this particular case for mobile NSS. The study consists of several scenario-based focus groups conducted with experts and target users as well as a large online survey. These investigations led to the insight that the preparation phase in a negotiation, including domain knowledge, value and preference elicitation, is the most crucial phase in the decision making and negotiation process. Moreover, it is the phase where users see a support device most fit and socially acceptable.

The insights gained in Chapter 3 led us to focus our subsequent design work on the most crucial component of preparing a negotiation and for decision making in general, i.e. knowing what one wants. Besides understanding the given decision domain, this entails knowing what one wants, what the consequences of decisions may be and why certain things are important (in the long run). What someone wants is usually captured in preferences that need to be entered in the DSS for it to calculate the utility

of decision alternatives. Chapter 4 (based on previous publications [2,10,14]) presents three studies investigating how to design preferences elicitation interfaces that match human capabilities and ways of constructing preferences. Preferences as well as decision making in general are based on values, mostly in an implicit way. However, thinking explicitly about one's personal values before settling on a set of decision alternatives to scrutinize is advocated by value-focused thinking as introduced by (Keeney, 1996). Therefore, we focused our design work presented in Chapter 5 (based on our previous publications [3, 1]) on supporting people to reflect on and thus become aware of their values.

The subsequent Chapter 6 (based on our previous publications [4, 5, 6]) presents the turn from the practical design investigations to the methodological perspective, i.e. on how to design. In this chapter we review our design work presented in-depth in chapters 4 and 5 and two other design works related to this thesis. We aim at providing insights into how to engage and empower different stakeholders to take part in the design process. In specific, we looked at facilitation of design sessions, group compositions and creativity triggers.

Chapter 7 answers the main research questions posed above and concludes the dissertation with presenting the main contributions, limitations and recommendations for future work. Summarizing the contributions, we present design guidelines for preference elicitation and reflection on values for human-centered DSS as well as more specific guidelines for designing NSS. Furthermore, we provide insights for methodological choices for stakeholder participation in design processes. Besides conceptual and methodological contributions, this dissertation results in practical contributions. Concrete prototypes of several interfaces for the job negotiation domain were developed as part of a bigger project focusing on building a pocket device for negotiation support. These include an interface for user-centered preference elicitation and a website for self-reflection on underlying values and constructing preferences in order to prepare for a job negotiation. Finally, we believe, the work can in part be transferred to other areas of research. We have already successfully used our developed design method and guidelines to build interfaces of a decision support system for the creation of science communication strategies by the Dutch government (Heinrichs, 2011). Furthermore, our tool for reflection on people's values could be used in eliciting stakeholder values, e.g. for value sensitive design (VSD) projects (see our published work on VSD (Pommeranz, Detweiler, Wiggers and Jonker, 2011)).

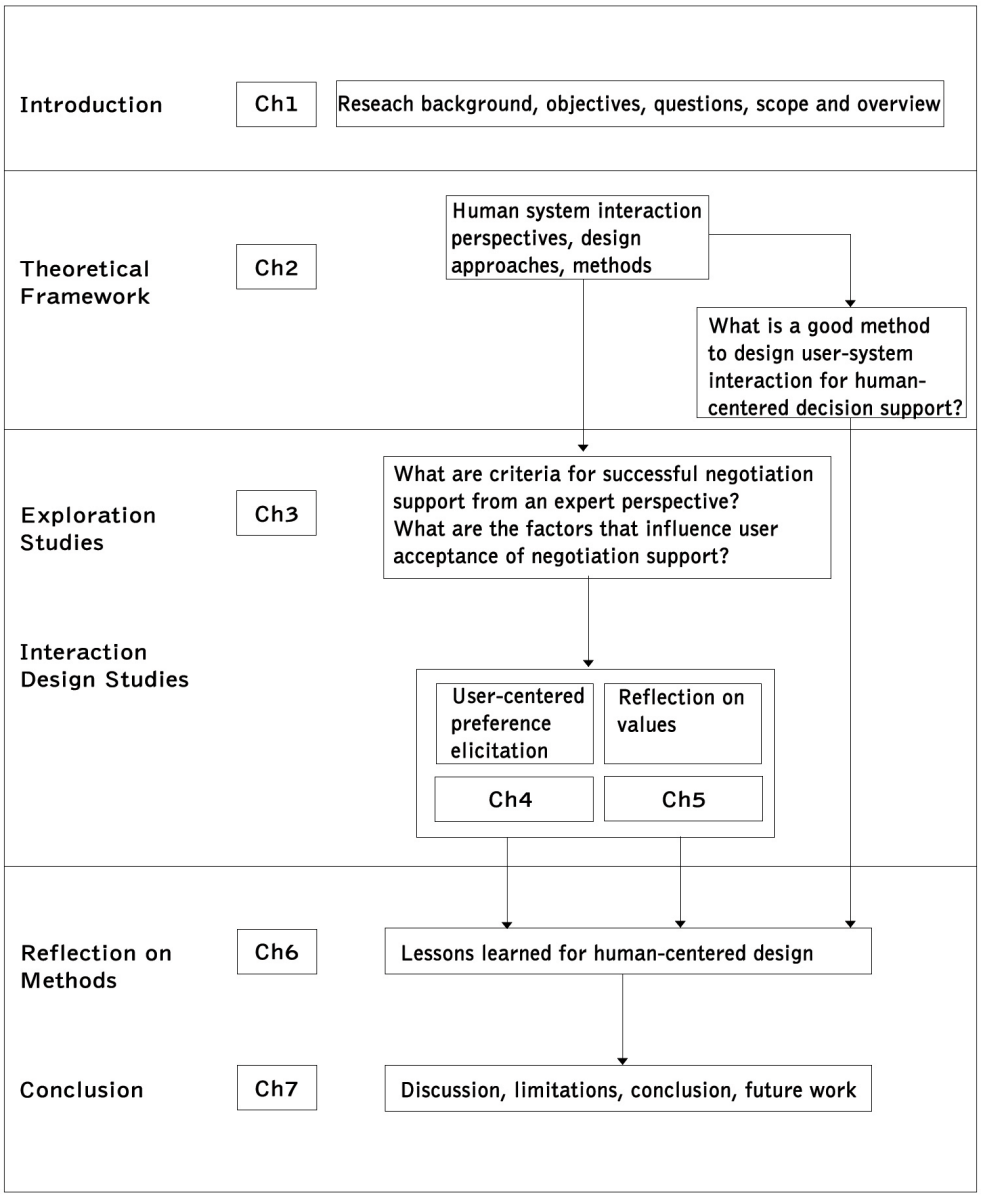


Figure 1.1: Thesis overview

1.6 Bibliography

Bellucini E, Zeleznikow J (2006) Developing negotiation decision support systems that support mediators: a case study of the familywinner system, Journal of Artificial Intelligence and Law 13(2):233–271.

Bosse T, Jonker CM (2005) Human vs. Computer Behaviour in Multi-Issue Negotiation. In Proceedings of the First International Workshop on Rational, Robust, and Secure Negotiations in Multi-Agent Systems, RRS'05, 10–25.

Fano A, Kurth SW (2003) Personal choice point: helping users visualize what it means to buy a BMW. In Proceedings of the 8th international conference on Intelligent user interfaces, IUI '03, ACM, New York, NY, 46–52.

Fiedrich F, Burghardt P (2007) Agent-based systems for disaster management. Commun. ACM 50(3):41–42.

Fieschi M, Dufour JC, Staccini P, Gouvernet J, Bouhaddou O (2003) Medical Decision Support Systems: Old Dilemmas and new Paradigms? Tracks for Successful Integration and Adoption. Methods Inf Med 42: 190–8.

Hindriks KV, Jonker CM (2008) Creating Human-Machine Synergy in Negotiation Support Systems: Towards the Pocket Negotiator. In Proceedings of the First International Working Conference on Human Factors and Computational Models in Negotiation (HuCom'08), Delft, The Netherlands.

Heinrichs R (2011) Designing a designer: on designing decision support systems for designing communication strategies. Masterthesis Delft University of Technology.

Keeney R (1996) Value-focused thinking: Identifying decision opportunities and creating alternatives. European Joumal of Operational Research 92:537–549.

Kersten G, Lo G (2003) Aspire: an integrated negotiation support system and software agents for ebusiness negotiation. International Journal of Internet and Enterprise Management 1(2):293–315.

Kersten H GE and Lai (2007) Negotiation support and e-negotiation systems: An overview. Group Decision and Negotiation 16:553–586.

Miller B, Albert I, Lam S, Konstan J, Riedl J (2003) MovieLens unplugged: experiences with an occasionally connected recommender system. In Proceedings

of Proceedings of the 8th international conference on Intelligent user interfaces, IUI'03, Miami, FL, 263–266.

Payne JW, Bettman JR, Schkade DA (1999) Measuring constructed preferences: Towards a building code. Journal of Risk and Uncertainty 19(1-3):243–270.

Pommeranz A, Detweiler C, Wiggers P, Jonker C (2011) Elicitation of situated values: need for tools to help stakeholders and designers to reflect and communicate, Ethics and Information Technology, 1–19.

Rangaswamy A, Shell G (1997) Using computers to realize joint gains in negotiations : Towards an electronic bargaining table. Management Science 43(8):1147–1163.

Schoop M, Jertila A, List T (2001) Negoisst: a negotiation support system for electronic business-to-business negotiations in e-commerce. Data and Knowledge Engineering 47(3):371–401.

Shen E, Lieberman H, Lam F (2007) What am I gonna wear?: Scenario-oriented recommendation. In Proceedings of the 12th International Conference on Intelligent User Interfaces, IUI'07, ACM, Honululu, Hawaii.

Stolze M, Ströbel M (2003) Dealing with learning in ecommerce product navigation and decision support: the teaching salesman problem. In Proceedings of the Second Interdisciplinary World Congress on Mass Customization and Personalization.

Thomson L (2005) The Heart and Mind of the Negotiator. Pearson Prentice Hall.

Vetschera R, Kersten G, Koeszegi S (2006) User assessment of internet-based negotiation support systems: An exploratory study. Journal of Oganizational Computing and E-Commerce 16(2):123–132

1.7 List of Publications

1.7.1 Journal articles

[1] A. Pommeranz, P. Wiggers and C. M. Jonker. Designing for self-reflection on values for improved life decisions. *Interacting with computers (submitted).*

[2] A. Pommeranz, D. J. Broekens, P. Wiggers, W. P. Brinkman and C. M. Jonker. Designing Interfaces for Explicit Preference Elicitation: a user-centered investigation of preference representation and elicitation. *User Modeling and User-Adapted Interaction, Volume 22, Numbers 4-5 (2012), 357-397. Impact Factor 3.075.*

[3] A. Pommeranz, C. Detweiler, P. Wiggers and C. M. Jonker. Elicitation of situated values: need for tools to help stakeholders and designers to reflect and communicate. *Ethics and Information Technology, Springer, November 2011 (online first). Impact Factor 0.564.*

[4] A. Pommeranz, P. Wiggers, W. P. Brinkman and C. M. Jonker. Social Acceptance of Negotiation Support Systems: Scenario-based Exploration with Focus Groups and Online Survey. *Cognition, Technology & Work, Springer, May 2011.*

1.7.2 Peer-reviewed conference papers

[5] A. Pommeranz, P. Wiggers and C. M. Jonker. Towards value-focused decision support systems. *European Conference on Cognitive Ergonomics (ECCE'12), Edinburgh, August 2012.*

[6] A. Pommeranz, U. Ulgen and C. M. Jonker. Exploration of facilitation, materials and group composition in participatory design sessions. *European Conference on Cognitive Ergonomics (ECCE'12), Edinburgh, UK, August 2012.*

[7] A. Pommeranz, P. Wiggers and C. M. Jonker. Towards Compositional Design and Evaluation of Preference Elicitation Interfaces. *HCI International, Orlando, FL, July 2011.*

[8] A. Pommeranz, C. Detweiler, P. Wiggers and C. M. Jonker. Self-Reflection on Personal Values to support Value-Sensitive Design. *BCS HCI work in progress paper, Newcastle, UK, July 2011.*

[9] A. Pommeranz, P. Wiggers, W. P. Brinkman, and C. M. Jonker. Social Acceptance of Negotiation Support Systems. *USAB2010-HCI in Work & Learning, Life & Leisure, Springer Lecture Notes in Computer Science, November 2010.*

[10] A. Pommeranz, P. Wiggers and C. M. Jonker. User-Centered Design of Preference Elicitation Interfaces for Decision Support. *USAB2010-HCI in Work & Learning, Life & Leisure, Springer Lecture Notes in Computer Science, November 2010.*

[11] A. Pommeranz, W. Visser, J. Broekens, P. Wiggers, K. Hindriks and C. M. Jonker. DUO meta-model for knowledge elicitation and bidding support in NSS. *2nd Working Conference on Human Factors and Computational Models for Negotiation (HuCom'10), Delft, The Netherlands, June 2010.*

[12] A. Pommeranz, W. P. Brinkman, P. Wiggers, J. Broekens, C. M. Jonker. Design Guidelines for Negotiation Support Systems: An expert perspective using scenarios. *European Conference on Cognitive Ergonomics (ECCE'09), Helsinki, Finland, September/October 2009.*

[13] P. Sundström, T. Jaensson, K. Höök and A. Pommeranz. Probing the Potential of Non-Verbal Group Communication. *Group 2009, Sanibel Island, FL, May 2009.*

[14] A. Pommeranz, J. Broekens, W. Visser, W. P. Brinkman, P. Wiggers, and C. M. Jonker. Multi-angle view on preference elicitation for negotiation support systems. *First International Working Conference on Human Factors and Computational Models in Negotiation (HuCom'08), Delft, The Netherlands, December 2008.*

[15] M. Bylund, K. Höök and A. Pommeranz. Pieces of Identity. *NordiCHI, Lund, Sweden, October 2008.*

1.7.3 Workshop papers

[16] C. Detweiler, and A Pommeranz. Go slow and beware! A call for reflection on our computational surroundings. *DIS 2012 Workshop on Slow Technology – Critical Reflection and Future Directions, Newcastle, UK, June, 2012.*

[17] A. O'Kane, C. Detweiler, and A Pommeranz. Designing and Evaluating for Trust: A Perspective from the New Practitioners. *Interact 2011 Workshop on Values in Design - Building Bridges between HCI, RE and Ethics, Lisbon, Portugal, September, 2011.*

[18] C. Detweiler, A. Pommeranz, and C. Jonker. Personal Informatics for Reflection on Personal Values *CHI'11 workshop on Personal Informatics & HCI: Design, Theory,& Social Implications, Vancouver, Canada, May 2011.*

[19] A. Pommeranz, P. Wiggers,W. P. Brinkman, and C. M. Jonker. Attitudes towards mobile NSS in different use contexts. *ECCE'10 Workshop: Putting users' first: the importance of human-centred design in the development of mobile applications and services, Delft, the Netherlands, August 2010.*

[20] J. Broekens, A. Pommeranz, P. Wiggers, and C. M. Jonker. Factors Influencing User Motivation for giving Preference Feedback. *Proceedings of the 5th Multidisciplinary Workshop on Advances in Preference Handling (M-Pref) in conjunction with ECAI 2010, Lisbon, Portugal, August 2010.*

1.7.4 Workshop organization

C. Detweiler, A. Pommeranz, and L. Stark. Methods to Account for Values in Human-Centered Computing. *ACM SIGCHI Conference on Human Factors in Computing Systems – CHI'12, Austin, TX, May 2012.*

C. Detweiler, A. Pommeranz, J. van den Hoven, H. Nissenbaum. Values in Design – Building Bridges between RE, HCI and Ethics. *Human-Computer Interaction – INTERACT 2011 13th IFIP TC 13 International Conference, Lisbon, Portugal, 2011.*

CHAPTER 2

BACKGROUND: DESIGNING HUMAN-CENTERED SYSTEMS

*In order to build the theoretical framework needed to answer the main research question of this dissertation "How can we design user-system interaction for human-centered decision support?" this chapter provides background from the literature. Literature that is relevant to specific chapters in the thesis (i.e. Chapter 3-5) will be discussed in-depth in the respective chapters. Therefore, we focus here on giving an overview of Human Computer Interaction (*HCI*) topics relevant to the design of human centered decision support. The human side is often neglected in current support system design, which is rather focused on algorithms and technical solutions than on human characteristics and knowledge. With this thesis we would like to trigger a shift from the technical perspective to a user-centered perspective. In the following we first introduce several perspectives on user-system interaction inspired by different disciplines (section 2.1). Subsequently, we outline our design-oriented standpoint (section 2.2) as opposed to the prevailing engineering standpoint and, last, present our methodological choices (section 2.3).*

> After a certain high level of technical skill is achieved, science and art tend to coalesce in esthetics, plasticity, and form. The greatest scientists are always artists as well.
> *Albert Einstein (1879-1955)*

2.1 Perspectives on Human-System Interaction

HCI is concerned with the study, planning, and design of the interaction between humans (users) and computers. As a research and practice field, HCI emerged in the 1980s, when computers were not anymore merely used by a handful of trained professionals, but became available to a broader public through personal computing. At that time needs for usability and HCI became apparent. Simultaneously, "cognitive science presented people, concepts, skills, and a vision for addressing such needs. HCI was one of the first examples of cognitive engineering." (Carroll, 2009) The term HCI was coined by Card, Moran, and Newell, whose goal was to develop, what they called, a scientific psychology to arrange the interface between humans and computers "easy, efficient, error-free – even enjoyable."(Card et al, 1986)

To investigate how to design human-system interaction for human-centered decision support, a central concept is *interaction*. Generally, interaction can be defined as "mutual or reciprocal action or influence" (Merriam-Webster Dictionary). Through the decades of HCI different perspectives on interaction have emerged which we briefly outline in the following.

2.1.1 From cognitive engineering to alternative views on interaction

Due to the field interdisciplinary roots interaction can be seen from different perspectives. In the following we will touch upon some of these perspectives and models developed in HCI in the past decades and analyze which perspective is most appropriate for the design challenge at hand.

Within HCI several paradigms have emerged inspired by other disciplines to describe interaction. One way is to formalize interactive behavior, e.g. with state diagrams. As stated by Svanaes (2011): "Formal representations of interactive behavior are well suited to describe the technical side of interactivity, but say little of the human side. They are of little value in answering questions like: 'How is the interaction experienced?', 'What does the interaction mean to the user?' To be able to answer such questions about the interactive user experience, we have to leave formal logic and the natural sciences and turn to the humanities and the social sciences."

Since early HCI a prevailing perspective to describe the human side is The Model Human Processor (Card et al, 1986) based on the cognitive science paradigm. In this model the user is seen as an information processor, receiving information or stimuli from the computer, processing it in the "cognitive processor" leading to user actions executed by a "motor processor". A detailed account of user actions, a so-called

execution-evaluation model, is the seven-step action cycle of Norman (1988). The execution side starts with the user having a goal (1), which would lead to an intention to act (2) and subsequently a sequence of actions sent to the motor processor (3) and last, the execution of the actions (4). The results would then be perceived (5), interpreted (6) and evaluated (7) with regard to the initial goal by the user.

"A number of researchers in HCI have argued that the information-processing model reduces the user to a mechanical symbol-processing machine, leaving out important aspects of what defines us as human" (Svanaes, 2011). An alternative to the cognitive science view was introduced by Winograd and Flores (1985) based on Heidegger's phenomenology, in particular his analysis of tools. Instead of taking for granted that human cognition is symbolic and symbols map one-to-one to objects in the world, this view takes into account the human's factual existence in the physical and social environment. Artifacts exist in the world and the specific meaning of the interaction with artifacts depends on the use context and intentions of the user. "Heidegger would also argue that to be able to understand how an interaction is meaningful for a specific user, we would have to understand the lifeworld of that user, i.e. the cultural and personal background that serves as a frame of reference and context for every experience of that person."(Svanaes, 2011)

Based on this view computers can be seen as tools, that reside in the background of the interaction until a breakdown occurs, e.g. software does not work as expected. Similarly, Suchman (1987) viewed interaction as "situated, social and in direct response to the physical and social environment" (Svanaes, 2000). Based on ethnomethodological theories, "meaning is always created in a situation, and how the interpretation of the situation in the next moment constitutes the situation" (ibid).

Another interaction perspective was proposed by Bödker (1989) and is based on Activity Theory. Her perspective distinguishes between human *activities* composed of *actions*, which are in turn composed of *operations*. While actions are consciously undertaken, operations are usually unconscious, until a breakdown occurs. For example, if someone uses a text editor to write a letter (the activity), actions include starting the application, typing in the letter, saving or printing it, while operations are pressing keys on the keyboard or moving the mouse cursor. Actions can be either communicative (directed at a subject) or instrumental (directed at an object). Important here is that communication is not used to denote interaction with objects (computers) and state changes are merely directed at the computer artifact and not to other people's minds (Svanaes, 2000). As Bödker (1989) explained, her interaction perspective implies, that "The use interface cannot be seen independently of other conditions of the use

activity and of the goal or object of the activity. This means that in both analysis and design of user interfaces we must consider the totality of the use situation (i.e. the specific practice of the group of users, the tools and materials that are applied in the process, etc.). Analysis and design cannot be done by outside observers, it has to be done by people who know enough about the professional practice to help interpret the occurring breakdowns."

Resulting from the different views, four perspectives can be distinguished that map out the interaction between humans and digital artifacts. Computers can be seen as tools, media, systems or dialog partners (Bödker, 2000). To summarize, people can use computers as tools (like a hammer), people and computers can be seen as the same components (system perspective), computers can be used as media in human-human communication or, last, the interaction between humans and computers can be seen analogous to a human-human dialog. As defined by Löwgren (2008) tools are "contraptions intended to be used instrumentally, for solving problems and carrying out tasks, and mostly to be used individually." While this view applies to the majority of computer uses by individuals to execute work tasks efficiently, "digital technology in society today is mostly used for communication, i.e., as a medium" (ibid).

While we did not strictly adhere to one of the described perspectives, this thesis is based on a viewpoint – influenced by the above views – that takes into account the human acting in the physical and social environment. While existing DSS seem to focus more on the system-perspective seeing the user as an entity that provides inputs readable by the machine and uses the machine's outputs to execute the following action, we believe that the interaction is influenced by the use context, intentions and characteristics of the user. This has to be considered in the design of DSS that better fit the users. Besides better fit, we are concerned with making DSS tools that give people a chance to enhance their skills, i.e. making them better decision makers, and use them effectively instead of merely relying on the skills of the systems. Throughout our work we considered the overall DSS to be a companion, thus, similar to a dialog partner, offering a number of tools (interfaces) to gain self-knowledge (e.g. about preferences and values).

2.2 Design Approaches to Interactive Systems

Alongside different interaction perspectives, different approaches to design interactive systems exist. In this section we give a brief overview of different approaches. Our goal is to highlight differences between technology-focused approaches, still

prevailing in current DSS design, and user-centered approaches that we suggest for the design of DSS.

The design of interactive computing systems can broadly be approached in two ways: starting from the technology or from the human. "Software Engineering's core concern is software, so people and systems in the socio-technical sense are minor concerns; whereas HCI focuses on people and the user interface as well as on the design of the wider socio-technical system, at the expense of software architecture." (Sutcliffe, 2011) Influenced by cognitive engineering early HCI approaches were rather engineering than design-focused. As Löwgren states, "historically, there has been a tendency in human-computer interaction, usability engineering and human factors to focus on instrumental and technical aspects. Interaction design as a designerly activity would insist that the aesthetical and ethical qualities can never be ignored or factored out. Whether something looks and feels good to use, and whether it makes you comfortable in terms of social accountability and moral standards, has a real impact not only on the overall user experience but also on measurable, instrumental outcomes. For an interaction designer, users are whole people with complex sensibilities and design processes need to be conducted accordingly."(Löwgren, 2008)

2.2.1 Engineering approaches

As said above, engineering approaches focus on technical aspects and are to be situated within the system perspective of interaction outlined in the previous section. As early approaches to technology design were concerned with concrete work tasks, that people carry out to reach a certain goal, analysis, evaluation and design focused on these tasks instead of the user's needs or reasons to carry out the task in the first place. To engineer computer systems, task analyses are carried out in order to model the application domain and specify the system requirements.

"An important characteristic of engineering is that a descriptive analysis for a requirement specification has to be done before creative design begins. In this sense such an approach cannot be considered to be a designerly process such as interaction design" (ibid).

Many existing DSS have been developed in the Artificial Intelligence field as so-called expert systems according to a specific engineering approach, i.e. knowledge engineering (Kendal and Creen, 2007). This is defined as a "discipline that involves integrating knowledge into computer systems in order to solve complex problems normally requiring a high level of human expertise".(Feigenbaum and McCorduck,

1983) In the construction of an expert system two roles are present, the domain expert who provides the knowledge base and the knowledge engineer who transforms it into the system representation and inference rules required to solve the problem at hand. The end-user is usually not involved in the construction phase. We believe that this approach leads to systems that are inflexible as they require specific inputs and do not adapt to the user's characteristics, needs and user context. For that we have to turn to user-centered approaches.

2.2.2 User-Centered design

Opposing system-focused approaches, a well-established approach to designing interactive systems within HCI is user-centered design (UCD) as coined by Norman (1988). In this approach attention is paid to user needs and characteristics at all stages of the design process. UCD approaches (Abras et al, 2004) commonly engage with stakeholders as informants and testers, e.g. to elicit domain knowledge and needs and in usability tests. Common methods to understand the user's needs and work context are observations and interviews, e.g. contextual inquiry (Beyer and Holtzblatt, 1998). Data collected through these methods can then be used to create Personas and Scenarios, as described in the following section. As the main focus is the usability of products, iterative user testing and improvement of prototypes is another core activity in UCD. These activities, however, still leave the main creative design work and decision making to the designer or researcher. Active involvement of end-users and other stakeholders in the creative and decision processes of technology design is the focus of cooperative approaches explained next.

Cooperative Approaches

According to Löwgren, "[f]raming design as exploration also means that it often makes sense to spend time in early phases on divergent work, essentially looking around in a design space of possibilities before committing to a particular direction. Exploring possible futures in interaction design often involves inviting the future users in various forms of participation."(Löwgren, 2008)

Involving participants, in particular future users, in the design of technology has a long-standing tradition in the Scandinavian Participatory Design (PD) (Ehn, 1989). Developed during the labor union movement of the 1970s in Scandinavia, it has a strong focus on empowerment of workers who were confronted with the introduction of new technology in their workplace. Due to its historical scope PD has led to methods that engage in envisioning futures involving changes in the social, technological and

political environment in which they are situated. However, more recently PD has become attractive to the design of single systems as well based on the general belief that "active user involvement in the software development process leads to more useful and usable software products" (O'Neill, 2000).

Co-design (Sanders and Westerlund, 2011), a more recent approach, focuses less on the work domain and more on services and products in general. This creativity-based approach to engaging stakeholders introduces the notion of co-design spaces. This notion is three-fold, referring to the physical design space a team works in, the space constituted by participant activities and the future solutions being developed. Within this approach focus lies on supporting common people's creativity in cooperative design workshops and prototyping sessions, as described below.

Value Sensitive Design

The above mentioned approaches incorporate values such as empowerment, democracy or universal usability. However, many other values (e.g. privacy or liability in DSS) play a role with regard to technology. A design approach that deals with values explicitly is Value Sensitive Design (VSD). "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 integrated and iterative three-part methodology consisting of conceptual, empirical, and technical investigations to incorporate knowledge of the ethical impact of a technology into the design. Furthermore, VSD introduces the notion of direct and indirect stakeholders, which expands user-centeredness to considering all people affected by a technology. For each group of stakeholders harms and benefits are identified, and satisfying value trade-offs are aimed for. By that VSD supports the creation of ethical and human-centered systems.

2.3 Research and Design Methods

We follow a user-centered and in particular participatory approach to our research as briefly outlined above. We deem user involvement necessary to create DSS that correspond to user needs and characteristics and empower them to take their own decisions than blindly follow an expert system's advice.

With this mindset we employed a set of methods and techniques to primarily engage target users in the design process, but also domain experts at stages where we deemed expert domain knowledge necessary to advance the design. To triangulate our data

we mixed quantitative and qualitative methods throughout the thesis work. While the majority of studies were carried out within the lab, we also aimed to get feedback from experts and users in their everyday environment, e.g. in expert interviews at their workplace and diary studies with users (not included in detail in this thesis, for details see (Pommeranz et al, 2010)) or the deployment of the Reflections prototype (chapter 5). HCI methods are vast and we selected carefully which methods would lead us to our goal at each stage of the research. As each chapter describes in detail which methods were used and what procedure was followed we will give brief introductions at this point and our rationale for the choice of each method.

2.3.1 Scenarios

When starting a new design project, in our case the design of a DSS, and in specific NSS, the design space needs to be explored first. As current NSS are mainly used by experts. There are no examples of systems aimed at supporting life-choices and used by untrained negotiators. In addition, applying for jobs or buying a house are not everyday tasks of people. Therefore, it is difficult to interrogate people about their needs with respect to such a system.

As Carroll (2000) stated:"While there is plenty of opportunity to do things that make a difference, it is never unequivocal just what should be done, or even just what the real problems are. The problems can only be definitively analyzed by being solved; the appropriate solution methods must typically be executed in order to be identified; the solutions must be implemented in order to be specified." To explore the design space and narrow down the actual problem we used scenario-based design as it offers a way to envision several use cases and functionalities of an NSS without actually implementing it. The core of this method are descriptions of how people accomplish tasks. As scenarios are stories, they contain elements of stories such as a setting, at least one actor, objectives and a plot (events and actions). This narrative nature of scenarios enables users, experts as well as designers to imagine the use situations and contexts of new or existing technology. Therefore, scenarios are useful in the design process since they capture the consequences and trade-offs of designs (Carroll, 2000). Scenarios can contain different levels of detail with respect to the designed system. Besides the scenarios introduced by Carroll that highlight the goals of the system and users (successful or erroneous) interaction with the system, scenarios can tell the user's story before the introduction of technology (e.g. to enrich so-called Personas (Pruitt and Grudin, 2003) –rich descriptions of fictional target users based on user observations) or they can be very detailed descriptions of the functionality of a

system, thus called use cases. Further adaptations of the method have also focused on tackling the ethical issues that come along with the introduction of new technologies, e.g. value scenarios (Nathan et al, 2007). Scenarios can also come in different forms, e.g. written, drawn or filmed.

In the beginning of our design project we chose Carroll's scenario-based approach as a means to explore the design space with negotiation experts and target users as well as a communication means within the design team. In particular, we were interested in the settings in which such a system would be used and what functionality it should provide. In Chapter 3 we describe how we created scenarios that envisioned the use of a NSS in different social situations, and how we used them in focus groups with experts and users.

2.3.2 Focus groups

Focus groups is a method that consists of group discussions involving several stakeholders. They have been widely used in marketing to exploit the dynamics of group discussions in order to receive attitudes towards ideas or products (Sim, 2001). Bruseberg and McDonagh-Philp (2002) have shown that focus groups are also useful during the design process of new technologies. They help participants to articulate their ideas and provide the researcher with inspiration for the design process. In particular, researchers can present ideas in combination with visual stimuli (e.g. photos, videos) to trigger targeted feedback.

We chose to employ this method with experts and users separately to understand the design problem from different perspectives. Besides the needs of target users from their own and the experts' perspectives we also wanted to explore in which situations it would be socially acceptable to use the system. Our scenarios were therefore, used as triggers in the focus group to help participants envision a future use situation.

2.3.3 Experiments

Experimental research lends itself best to test specific hypotheses. The researcher has great control over the environment and set-up of the study and usually a limited set of variables are manipulated to test several conditions (typically treatment and control conditions). Collected data is usually of quantitative nature and can be analyzed statistically. While this procedure allows for testing of many participants and easy comparisons of data to arrive at general conclusions, the artificial setting often influences the outcomes. It is, therefore, hard to say whether people would act,

e.g. use a system, differently in a real life situation. Despite this problem we chose to use experimental set-ups in some cases to study psychological effects, especially in relation to preference elicitation (Chapter 4), as it allowed us to compare data from participants easily. Besides inviting people to a lab we set-up online experiments to mitigate some of the influences an artificial lab setting has. In the latter cases people were free to choose the time and environment in which to carry out the experiment.

2.3.4 Surveys

A method similar to online experiments, but of a more exploratory nature is the online survey method. With surveys large samples of a population can be asked, e.g. through questionnaires, about preferences, attitudes or opinions. Questions can be of qualitative or quantitative nature. In order to conduct statistical analysis quantitative data is needed and is often obtained in form of Likert ratings of statements constructed by the researchers. Surveys can be used to test several hypotheses or complete models, as was our goal in the research presented in Chapter 3 to predict social acceptance of NSS. We chose online surveys in two cases, as the method allowed us to test a model and several design ideas with a large sample of target users. Thus, we could confirm or reject insights that had been extracted from work with small groups of experts or users.

2.3.5 Prototyping

"Prototyping is a method used by designers to acquire feedback from users about future designs." (Sjoegaard, 2010) Prototypes can vary in form and level of functionality from simple paper mock-ups that visualize the aesthetical appearance and some possible functionality to highly functional digital prototypes that can convey the concrete feel of the interaction with a system. Prototypes can be used for testing specific design ideas with users or as means of communication (so-called boundary objects) between stakeholders. In our research we have used prototypes mainly in two manners, (1) as triggers for critical feedback (e.g. the Reflections prototype in Chapter 5) and creativity (as explained in Chapter 6) and (2) as objectives/outcomes of cooperative design sessions (Chapters 4 and 6). In the first case prototypes were built by the researchers and in the latter case by stakeholders.

2.3.6 Participatory design workshop

We use 'participatory (or cooperative) design workshop' as the umbrella term for the several sessions we conducted with target users and other stakeholders in which

concrete design ideas were developed and usually implemented as paper prototypes. We did not follow the same set-up each time, as part of our research was the investigation of how to optimally set up workshops to engage participants. In most cases participant groups of different sizes worked on a design together and used different materials, e.g. interface elements, in a compositional way to create a paper version of a user interface. While the methods described above (except prototyping) were used to inform design, the workshops were the core activities in our design-based research agenda to arrive at concrete guidelines and interface prototypes. In one case (Chapter 5) we used the Future Workshop technique (Jungk and Müllert, 1987) as an inspiration for the set-up of our workshop into preparation, critique, fantasy and implementation phases.

2.4 Bibliography

Abras C, Maloney-Krichmar D, Preece J (2004). User-Centered Design. In Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications.

Beyer H, Holtzblatt K (1998) Contextual Design: defining customer-centered systems. San Francisco, Morgan Kaufmann Publishers Inc.

Bödker S(1989): A Human Activity Approach to User Interfaces. Human–Computer Interaction 4(3):171–195.

Bödker S, Nielsen C, Petersen MG (2000) Creativity, cooperation and interactive design. In Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques, DIS'00, Brooklyn, NY, 252–261.

Bruseberg A, McDonagh-Philp D (2002) Focus groups to support the industrial/product designer: a review based on current literature and designer's feedback. Applied Ergonomics 33(1):27–381.

Card SK, Moran TP, Newell A (1986) The Psychology of Human-Computer Interaction. CRC Press.

Carroll JM (2000) Making Use: Scenario-based Design of Human-Computer Interactions. MIT Press, Cambridge.

Carroll JM (2000) Five reasons for scenario-based design. Interacting with Computers 13:43–60.

Carroll JM (2009): Human Computer Interaction (HCI). In: Soegaard, Mads and Dam, Rikke Friis (eds.). Encyclopedia of Human-Computer Interaction. Aarhus, Denmark: The Interaction-Design.org Foundation.

Ehn P (1989). Work-oriented Design of computer artifacts. Hillsdale, New Jersey: Lawrence Erlbaum Ass.

Feigenbaum EA, McCorduck P (1983), The fifth generation (1st ed.), Reading, MA: Addison-Wesley.

Friedman B, Kahn PHJ (2003): Human values, ethics, and design. In: Jacko JA, Sears A (eds) The human-computer interaction handbook. Lawrence Erlbaum Associates, 1177 – 1201.

Jungk R, Müllert N (1987): Future workshops: How to Create Desirable Futures. Institute for Social Inventions. London, UK.

Kendal SL, Creen M (2007) An introduction to knowledge engineering, Springer.

Löwgren J (2008) Interaction Design. In: Soegaard, Mads and Dam, Rikke Friis (eds.). Encyclopedia of Human-Computer Interaction. Aarhus, Denmark: The Interaction-Design.org Foundation.

Nathan LP, Klasnja PV, Friedman B (2007) Value scenarios: a technique for envisioning systemic effects of new technologies. CHI '07 extended abstracts on Human factors in computing systems, San Jose, CA, 2585–2590.

Norman DA(1988) The Design of Everyday Things. New York, Doubleday.

O'Neill E (2000) User-developer cooperation in software development: building common ground and usable systems. Springer.

Pommeranz A, Visser W, Broekens JD, Wiggers P, Hindriks KV, Jonker CM (2010) DUO meta-model for knowledge elicitation and bidding support in NSS. In Proceedings of the 11th Group Decision and Negotiation Conference 11:120–123.

Pruitt J, Grudin J(2003) Personas: practice and theory. Proceedings of the 2003 conference on Designing for user experiences, San Francisco, CA, 1–15.

Sanders EB-N, Westerlund B(2011) Experience, exploring and experimenting in and with co-design spaces. In Proceedings of the 4th Nordic Design Research conference, NORDES'11, Helsinki, Finland, 1–5.

Sim J (2001) Collecting and analysing qualitative data: issues raised by the focus group. Journal of Advanced Nursing 28(2):345–352.

Soegaard M (2010). Prototyping. Retrieved 12 April 2012 from http://www.interaction-design.org/encyclopedia/prototyping.html.

Sutcliffe AG (2011) Requirements Engineering. In: Soegaard, Mads and Dam, Rikke Friis (eds.). Encyclopedia of Human-Computer Interaction. Aarhus, Denmark: The Interaction-Design.org Foundation.

Suchman, L. (1987). Plans and situated acttons: The problem of human-machine communication. Cambridge, Cambridge University Press.

Svanaes D(2000) Understanding Interactivity: Steps to a Phenomenology of Human-Computer Interaction. Trondheim, Norway, Norges Teknisk-Naturvitenskapelige Universitet (NTNU).

Svanaes D(2011): Philosophy of Interaction. In: Soegaard, Mads and Dam, Rikke Friis (eds.). Encyclopedia of Human-Computer Interaction. Aarhus, Denmark: The Interaction-Design.org Foundation.

Winograd T, Flores F(1985) Understanding computers and cognition. Norwood, NJ, Ablex Publishing Corp.

CHAPTER 3

SOCIAL ACCEPTANCE OF NEGOTIATION SUPPORT SYSTEMS

*The goal of our first empirical investigations was to understand what is essential for the success and user acceptance of negotiation support systems (*NSS*). In order to understand what kind of support a negotiation support system needs to deliver and in which situations to guide the user through the negotiation process we engaged with experts and users alike. In this chapter we describe how we (1) consulted negotiation experts about common problems, remedies and success factors (section 3.5), and (2) elicited user views on how and when they would like to use a negotiation support system (section 3.6) by using several user-centered methods outlined in the previous chapter, i.e. scenarios, focus groups and surveys. We present 12 concrete guidelines for the design of* NSS *based on the focus group discussions and design implications based on the survey results.*[1]

[1]This chapter is almost equivalent to: Alina Pommeranz, Pascal Wiggers, Willem-Paul Brinkman, Catholijn M Jonker. Social Acceptance of Negotiation Support Systems: Scenario-based Exploration with Focus Groups and Online Survey. Cognition, Technology & Work, Springer, May 2011.

The most precious things in life are not those you get for money.
Albert Einstein (1879-1955)

3.1 Introduction

A skillful negotiator has to carefully balance the issues at stake, have a good understanding of his own and the opponent's needs and since negotiation is a social activity, manage relationships and handle emotions (Thomson, 2005). Often negotiating involves overlooking a vast amount of options, deciding on strategies and evaluating bids with multiple attributes. Computational power can facilitate these processes. Within different research areas, e.g. management science, e-commerce and artificial intelligence (Kersten, 2007; Rangaswamy and Shell, 1997; Schoop et al, 2001; Vetschera et al, 2006), researchers have worked on systems supporting people electronically in negotiations. Existing negotiation support systems (NSS) can significantly improve the human performance in negotiations and increase the number of win-win outcomes if the negotiation space is well-understood (Hindriks and Jonker, 2008; Kersten and Lo, 2003).

Despite these advantages that NSS can offer especially to the unexperienced negotiator, the majority of existing (NSS) are not used in real-life practice, but only for research and training purposes (Kersten, 1999). One reason for this problem may be the technical focus that is prevailing in current NSS development and thereby lacking to address social issues and human factors in the design. We believe that a user-centered design process is the key to understanding such issues and designing solutions that will be accepted by the intended target users.

Another reason may be that current NSS are developed as stand-alone applications (Kersten, 2007) or web-based applications (Kersten and Lo, 2003), and thereby lack in their ability to be applied in real-life negotiation contexts. Negotiation, however is an activity that can take place in almost any setting instead of being tied to, e.g., an office and, therefore, NSS should be designed to support people in these different settings. Imagine a negotiation for buying a new house. Part of this negotiation is e.g. collecting information about different neighborhoods, it involves visiting houses, discussing things with the owners etc. These actions take place in different settings and a NSS should be able to collect the data in these contexts, store them all in a central place and be able to give real-time advice in these settings based on what has been stored earlier.

The advance of mobile technology, especially the recent developments in smartphone technology and usage, opens up a whole new range of possibilities to make this possible. Mobile technology can enable people to have their NSS at hand in any negotiation phase (including e.g. the preparation) independent from place and time.

Devices such as smartphones, mobile phones, PDAs or hand-held computers offer, e.g., opportunities to store and compute large amounts of data, access online sources and show graphical data on color screens. Smartphones are additionally equipped with sensors such as GPS, microphones and cameras that can be employed to capture context and offer intelligent functionality (e.g. sensing the level of aggression during a conversation). The number and diversity of people using portable internet devices is rapidly growing (ITU, 2004), which makes mobile NSS even more feasible and attractive to a wide population of users.

We would like to take advantage of these trends and develop a new kind of NSS for mobile use, a so-called Pocket Negotiator (PN) as described by Hindriks and Jonker (2008). Our vision is to develop a mobile system that is able to collaborate with unexperienced negotiators in order to reach win-win outcomes in negotiations. The PN will enhance the negotiation skills and performance of the user by increasing the user's capacity for exploration of the negotiation space, i.e. possible bids and deals, reducing cognitive task load and preventing mental errors. The functionality of the device will be focused on handling computational complexity issues and providing bidding- and interaction advice. Our idea is to cover all negotiation phases (preparation, joint exploration, bidding and closure) (Thomson, 2005) with support from the system. Generally, such a system could be used in any negotiation domain. We believe it would be especially useful for negotiations with large possible outcome spaces (that are difficult for people to overlook) and important consequences, e.g. real estate or job contract negotiations.

The mobile nature of the system will allow users to refer to the support not only when they prepare themselves at home, but also when they are on the move or even during the face-to-face situation with the other negotiation partner. This entails several advantages. The users can, e.g., collect relevant information for the negotiation and enter it immediately into the NSS or update information about their preferences in case they change due to new information. They can practice the different negotiation steps and review tips and strategies at any time. In a face-to-face situation it might also be useful to enter information, e.g., revealed by the opponent (i.e. spoken words or information about the opponent's behavior, emotions etc.). Based on this input the NSS will be able to give context-relevant advice or it could just serve as a reminder for information entered by the user during earlier preparation. Also the possibility of connecting to a wireless network enriches the functionality of the NSS, e.g. by providing online market information.

With this new freedom mobile NSS offer, new questions and problems occur. First of

all, the functionality of the system and its interaction with the user needs to be carefully designed to fit the mobile settings. In a face-to-face setting, e.g., the user needs to focus mainly on the interaction with the negotiation party and does not have the cognitive resources available to interact to the same extent with the NSS. Second, the question of social implications arises. When putting NSS into the social setting of a face-to-face negotiation or using it in public spaces, we have to consider appropriateness and acceptance regarding the user, the opponent or bystanders. Entering information or consulting the NSS during a negotiation might interrupt the flow of the communication or bother the opponent for other reasons. Furthermore, the user might be concerned about his or her image when using a mobile NSS in public. These are issues worthwhile investigating.

Currently, the use of NSS is rather focused on preparation than on the actual negotiation, as further explained in the next section. We believe that in order to design NSS that will be successfully used in negotiations we need a human-centered approach investigating the attitudes people have towards NSS, especially given different use contexts. Our main goals are to elicit functional requirements from experts and potential users and to investigate the acceptance of NSS in different social settings.

The paper is structured as follows. Section 2 gives an overview of existing work in the area of NSS and acceptance of mobile devices and services. Section 3 explains our overall approach to eliciting requirements and understanding acceptance of NSS in different contexts. This approach is described in detail in the next sections, including developed scenarios of NSS use (section 4), expert focus groups (section 5), user focus groups (section 6), and a social acceptance survey (section 7). The results from the survey are explained in section 8, followed by possible design implications from the focus groups (design guidelines) and the survey (section 9). Finally, conclusions drawn from our work are presented in section 10.

3.2 Related Work

3.2.1 Existing negotiation support systems

In a recent review, Kersten (2007) give a detailed overview of NSS and E-negotiation systems. Among other things they give a categorization of software systems and a structure of key constructs used in NSS. An NSS developed by Kersten (2004) and used mainly for training and teaching is the Inspire system. The system employs a 3-phase model including pre-negotiation, negotiation and post-settlements. Kersten and Lai conclude that rather few systems were successfully used in real negotiations.

The majority of existing NSS has been used for training and research purposes, but has not been applied to real life negotiations (Kersten, 1999). A recent study on user acceptance of web-based NSS (Vetschera et al, 2006) predicts that 80 percent of the users would use the system to prepare and train for negotiations but only 61 percent would use it in the negotiation. Why is the acceptance for real cases so low?

One possible answer is that NSS development concentrates on technological solutions, while the social problems they intend to solve are secondary or completely neglected (Bui, 1994). Negotiation is inherently a social activity, since it involves communication between at least two parties and is influenced by the social setting in which it takes place. Literature on business science (Havard Business School Essentials, 2003) has, e.g., emphasized the influence of relationships on negotiation processes. Swaab et al (2004) argue for a careful analysis of social and psychological processes in order to design good NSS and claim that the success of an NSS depends on the understanding of the activity that the system will support. They primarily look at two aspects that influence the outcomes of negotiations positively, namely common (cultural) identity and shared cognition. In this sense NSS can help by providing information to the opposing parties to establish a common understanding of the problem and possible solutions. Their studies show that the nature and representation of the information can influence negotiation outcomes.

Another effort to emphasize the importance of social and also emotional issues in negotiation and their consideration for NSS has been made by (Bui, 1994). In his article the author points out problems that evolve from the fact that empirical research focuses only on the rational aspects of negotiation. For instance, the negotiation models that are implemented in NSS assuming strict economic rationalization ignore that people also take decisions based on social acceptability of different means to achieve a deal. Adding reasoning based on ethical and social norms to negotiation models will allow them to better represent the real life negotiation processes. Bui explores socio-emotional aspects such as conflict awareness, thoughts, emotions, intentions, trust and norms and their impact on negotiation. He creates a general list of aspects that NSS should help users with, such as identifying controversy, clarifying issues/criteria, equalizing parties or finding solutions and simulating impacts of potential decision. These can be seen as more generic guidelines for the functionality and design of NSS. These works (Bui, 1994; Swaab et al, 2004) refer to shared NSS used either collaboratively by all parties or as mediators. This is one type of NSS with special requirements. An interesting related research area where social aspects are, however, considered is the design of group decision support systems (Nunamaker et al, 1996). However, also in this research the focus is on collaboration and verbal communication

between the participants rather than other social aspects aspects like context, thoughts, emotions or trust.

3.2.2 Social Impacts of Mobile Technology

Researchers focusing on the adoption of mobile technology in general have recently included social context into their models. Social impacts of mobile technology have been widely studied (Ling, 1997; Love and Perry, 2004; Mallat et al, 2009; Palen et al, 2001), especially the pervasive nature of mobile phones in public places. Most of the literature in this area focuses on the distraction of bystanders by people talking loudly on the phone or by the mix-up of geographic spaces (current physical space the mobile phone user is in and the space created by a phone conversation) (Ling, 1997; Love and Perry, 2004; Srivastava, 2005). In the case of using a mobile NSS, distraction is, of course, especially an issue when the NSS user is in an active, ongoing communication with the other negotiation party (face-to-face or on the phone). The interaction with the device might disrupt this communication and therefore be less socially acceptable. Furthermore, the other party might not accept the interaction with the NSS because it allows the user to have an advantage and other party might feel excluded. In other situations where the NSS is used for preparation, social acceptance might be less of an issue.

3.3 Overall Approach

We aim to build a NSS that supports people that are non-professional negotiators (novices) and may have different levels of negotiation experience. To explore functional requirements and social acceptance in different use contexts we followed a scenario-based approach including three main steps: expert focus groups, user focus groups and an online survey. Although we aim at novices we did expert focus groups because they allowed us to grasp common pitfalls in negotiations that novices may not even be aware of. Since we are in the early stages of designing a PN, we do not have a running prototype at this stage. To be able to communicate our vision of a mobile NSS, that could be used in different contexts, we created a number of scenarios. Each scenario represents a use situation with distinct characteristics (see section 3.4.1). In order to emphasize the different design decisions made while writing the scenarios we did a claims analysis. These claims were used as a basis for short questionnaires used in the focus groups. We created storyboards and short films to visualize the scenarios. These films were used in all three steps of our approach. In the first step we conducted a number of focus groups with negotiation experts. With their expert knowledge we

expected to be able to get insights into common negotiation practices and problems people face, which could be addressed by the functionality of our NSS. Therefore, the focus was on the functional aspects. Second, we conducted focus groups with potential users, i.e. people with various levels of negotiation experience excluding experts. The focus in those discussions was the social acceptance. Focus groups deliver a lot of qualitative data, which is difficult to draw general conclusions from. Therefore, in the third step, we conducted an empirical study of social acceptance. We designed an online survey (a) to find out in which situations people consider a mobile NSS socially acceptable, (b) to find the factors and relationships that influence this acceptance in the different situations and social contexts and (c) to investigate the consequences of people's attitudes towards NSS for their design. In the following we describe the steps in detail.

3.4 Scenarios of Use Contexts

Before designing the concrete functionality of a PN and implementing first prototypes we would like to investigate the attitudes towards mobile NSS in different situations. This will enable us on the one hand to inform the further design process and on the other hand find answers to why current NSS are not used in real negotiations. To be able to give the experts and users an idea of our envisioned system and possible use contexts nevertheless, we used filmed scenarios in the focus groups and the online survey. In the following we will first describe the development of five scenarios representing different use contexts.

3.4.1 Scenarios

Scenarios are useful in the design process since they capture the consequences and trade-offs of designs (Carroll, 2000). The narrative nature of scenarios enables users to imagine the use situations and contexts of new or existing technology. In the project we currently focus on two example domains for NSS use: job contract and real estate negotiations. In order to capture all possible contextual factors in a number of scenarios we identified important dimensions for NSS use in a brainstorming session with the project group. These dimensions include:

1. presence of an opponent, i.e. whether the user is communicating with an opponent while using the NSS. This can be either face-2-face or remote communication (e.g. phone, internet).

2. number of users. Although the PN is meant to support one party in a negotiation, there can be a single user or a number of users (e.g. a couple) forming a party.
3. mobility. The NSS can be used either at home or at work or while being mobile (e.g. on a train).
4. mode of NSS use. The NSS can be used openly, i.e. the opponent knows about it, or in stealth mode, i.e. the opponent is unaware of the NSS use.
5. negotiation phases, i.e. preparation, exploration, bidding and closure. For the scenarios we mainly distinguish between preparation, which is typically done by the user alone and the last three phases which involve interaction with the other party.

Combining all of this dimensions would lead to a high number of use contexts. Therefore, we created meaningful combinations to be able to reveal all aspects and discuss them with our participants. We chose two use contexts illustrating a job negotiation: short preparation being mobile on a train and face-to-face with the boss with concealed use of the NSS. Two scenarios had real estate content: distant negotiation on the phone and collaborative preparation of a couple. The last one illustrated a situation face-to-face with open use at a car dealer.

For each of the five use contexts (Figure 3.1) we wrote a scenario presented in the following in summary. All scenarios were checked by a professional negotiation coach to make sure that they were sufficiently realistic. Each scenario is briefly discussed below. Italic text is taken from the original texts of the scenarios.

Mobile Preparation with Time Constraints (train) Preparation is one of the negotiation phases stressed in the literature, e.g. (Havard Business School Essentials, 2003). In this scenario we describe a preparation situation with special constraints. The job applicant Martin is already on his way to the interview. Therefore, he has limited time to prepare himself. In addition, the mobile setting constitutes another constraint, namely limited resources. Both constraints require special regard when it comes to the functionality of the device. Just before getting on the train Martin has received a mobile NSS from a friend. He uses the device's speed preparation function to prepare himself in the short time he has left. Among other functions the device allows him to receive knowledge about the job negotiation domain.

He wonders how much money he could ask for. He chooses 'expert opinion' on the interface and types in 'salary'. The PN *suggests a website that has a forum where*

Figure 3.1: Scenarios (Screenshots from videos) from left to right, top row: open use at car dealer, collaborative preparation before buying a house, on the phone with real estate agent; bottom row: evaluation talk with boss, preparation for job interview on the train.

you can discuss current average salaries for IT*-consultants with an expert in the field. After reading through the forum Martin has a quite good idea what he can ask for with his kind of educational background and experience. With that knowledge he feels more secure and relieved.*

Later in the scenario Martin makes use of the training module of the NSS which enables him to go through a simulated interview with a virtual agent. He receives on-the-fly advice about his and the opponent's actions. The scenario ends with Martin being more relaxed, knowing what to expect in the upcoming negotiation.

Face-to-Face Negotiation, Secret Use (F-2-F). The situation described in this scenario is a negotiation between an employee, Bianca, and her boss. Bianca is using a mobile NSS. The emphasis in this scenario is the concealed use of the NSS. Bianca is hiding the fact that she has support from an NSS by telling her boss she is using her device only to take notes.

Bianca has been working for a big telecommunication company in The Hague for 2 years now. Today her annual evaluation with her boss is due. Bianca wants to take this meeting as an opportunity to re-negotiate some parts of her contract. Since her

husband got a new job in another city, they decided to move further away. Therefore, she wants to discuss opportunities with her boss to handle the new situation. She knows that she worked hard and well in the last year and should get what she wants, but she does not consider herself a good negotiator. Therefore, she recently got the PN *and prepared herself for this negotiation with the device.*

Throughout the negotiation described in the scenario Bianca receives help from the device. Several functions are described in this scenario including, e.g., the management of emotions, generating new options, and receiving advice from the system. The scenario ends with a deal in which both parties gain something and are satisfied with.

Collaborative Preparation (Coll. Preparation). Negotiation involves a lot of emotions on both sides of the bargaining table, but also within one party, e.g., between two partners buying a house together. In this case the first step is to merge the demands and preferences of both partners before starting a negotiation with the opponent side. Our scenario describes a couple that is planning to buy a house together and uses the NSS during the preparation to sort out their preferences and to download domain knowledge about real estate.

The 'collaborative preparation' module starts up. After a short introduction the PN *asks each of them to put in their preferences for a house separately. Since they also have the* PN *software installed on their laptop they put in their preferences in parallel. From both preference profiles the* PN *creates a matching profile and shows the clashes of their preferences. It advices the couple discussing the clashes and trying to find trade-offs between them that suit both.*

During this process of compromising the couple gets into a quarrel in which both insist on their own wishes without even communicating the underlying reasons in detail. In this case our device takes on a proactive role and interrupts the couple to give advice on how to handle the conflict.

The PN *senses the noise and the angry voices in the room and assumes an argument. The* PN *suggests calming down [. . . and . . .] prompts them to put in an emotional value on a scale from 'I don't care at all' to 'I would die for this' for each variable they have different preferences on.*

After having sorted out all their preferences they start looking for houses. In the last scene of the scenario the couple visits a house and takes advantage of the PN's feature of taking pictures and storing them together with other information about the house in a database.

Negotiation on the phone (Phone). A negotiation in which both parties are not situated in a face-to-face setting, but are distant from each other offers different design challenges for a NSS. First of all one party does not see the other party and therefore the use of a NSS can take place without each others' notice. Especially in real estate situations, e.g. when buying a house another aspect to consider is that the negotiation is split into a number of phone calls. This gives the user time in between the calls to use the system in each step of the negotiation. Our scenario describes a couple negotiating for a house. Before the interaction with the opponent they prepare themselves with the help of the NSS.

Furthermore, the PN *has downloaded housing domain knowledge, such as contracts and legal issues and the prices of similar houses in the neighborhood to take into account. Before Mary came to work this morning she had decided with Piet to set a first bid around 450.000 Euro.*

At work Mary calls the agent and starts negotiating. Before and during the phone calls she uses the NSS on her laptop to receive advice about different steps in the negotiation, e.g. the PN advises her to not start the negotiation with offering a price, but instead talk about other issues and options.

The bidding goes on for a while and the PN *shows a visualization of the bids in the outcome space based on the preferences of Piet and Mary and the estimated preferences of the agent. After a while the* PN *detects that the bidding is not reaching a win-win situation.*

After finding new variables to include in the negotiation to reach an agreement that suits both parties they finally close a deal.

Face-to-Face Negotiation, Open Use (Car Dealer). We decided to include another scenario that has a face-to-face setting, but showing an open use of the NSS meaning that the other party is aware of the use. This scenario is about a couple buying a car. Our belief is that the car dealer's setting enables people to use the NSS more openly. When buying a car it is usually not necessary to stick to one specific car dealer. No long-term relationship needs to be considered. Therefore, the couple in the scenario openly states that they will be using the NSS and explain what they can do with it.

The focus of the scenario lies in the advice of time-outs at strategic points during the negotiation. During the process of looking at cars and refining their preferences for the new car, they enter information about the state of the negotiation into the NSS. They receive strategic advice on how to proceed and when to recapitulate.

Figure 3.2: Storyboard for scenario: Mobile Preparation with Time Constraints(train).

He [the car dealer] *shows them a range of more sporty looking family cars and the couple chooses their favorite. They enter that into the* PN. *The* PN *advices them to take a time-out and check whether they have considered all their preferences and whether all the information they need has been disclosed.*

After they have found an interesting car the bidding starts in the car salesman's office. The NSS assists the couple by comparing prices with similar cars online. They disclose to the salesman that the market price is lower than his offer. The salesman drops his price. They negotiate about a few extras and finally leave with a new car and a deal they are satisfied with.

3.4.2 Storyboards and Videos

Due to their illustrative strength scenarios are a good means to communicate design ideas within the project team as well as to users or experts in the field. In order to exploit that strength even more we decided to visualize the scenarios. First we created a storyboard (Figure 3.2) for each of the scenarios. These storyboards then served as a basis for the shooting and editing of short (about two to three minutes) videos (see http://mmi.tudelft.nl/negotiation/index.php/Media for videos (in Dutch) and complete English storyboards per scenario). Using videos we were able to present the use contexts of our NSS very well. Much of the functionality of the NSS was kept open for interpretation to avoid limiting the discussion about the functionality.

3.4.3 Claims Analysis

Due to the scenarios' narrative nature many things are left implicit. Often causal facts and relations underlying the actions described are not revealed. Therefore it is useful to enumerate such causal relations separately. This can be done through claims

analysis (Carroll, 2000). Each claim underlying a certain action or design feature in the scenario is listed together with its tradeoffs. We used the claims as proposed by (Neerincx, 2003), i.e. to test our hypothesis about functionality and use contexts in the focus groups discussions with the experts. We wrote down four to six claims per scenario based on our hypothesis. Due to space limitations we cannot list all the claims here, but only give examples. The first claim was written for the face-to-face scenario with the boss and the second for the negotiation on the phone scenario:

Advice claim the NSS gives generic advice for different negotiation phases in a text-based form (e.g. ask for reason of concern, be sympathetic, and maintain the relationship).

+ Even though the user might know of such things due to a good preparation, the NSS advice serves as a reminder during the negotiation process.

- The user might not be able to put the advice to practice or the way he tries to do so is not effective.

Graphical representation claim the NSS shows the current status of the negotiation graphically including all variables etc.

+ The variables and their influences on the negotiation process are shown, so that the user can understand the process better.

+ The user can recapitulate and learn for future negotiations by looking at the current status and the influences of the variables.

- The number of variables and influences is high and the user finds it hard to learn from the graphical representation.

- The graphical representation is not understood by every type of user.

3.5 Expert Focus Groups

Focus groups (Sim, 2001) have been widely used in marketing to exploit the dynamics of group discussions in order to receive attitudes towards ideas or products. (Bruseberg and McDonagh-Philp, 2002) have shown that focus groups are also useful during the design process of new technologies. They help participants to articulate their ideas and provide the researcher with inspiration for the design process. Lately, HCI researchers have adopted the method and refined the techniques used to stimulate the discussion. As for instance, (Goodman et al, 2004) found out, it is profitable to use

visual help such as pictures and also scenarios in focus groups. Furthermore, tasks can start up a discussion. Based on these findings we used the previously described filmed scenarios in the focus groups.

3.5.1 Set-up and Procedure

In total we had 12 experts divided into three focus groups. We divided the experts into different focus groups according to their expertise. As explained by a number of researchers, e.g. (Sim, 2001), the homogeneity of the group plays an important role. The more similar the group members are the more likely they are to voice their opinions. Therefore, we formed one group with general negotiation experts, such as negotiation trainers, lawyers, a judge etc., and two with job negotiation experts, such as human resource employees and labor union representatives. In the beginning participants were introduced to each other and the project was described. Every participant received a questionnaire that contained two claims from the claims analysis (see section 3.4.3) per video. The claims, however, were reformulated into statements that allowed the experts to specify their level of agreement with. The two claims named in the previous section were presented as the following statements:

- General tips and strategic advice (e.g. try small talk, show sympathy for your opponents concerns) is more useful for the user than specific behavior- and decision-advice.
- The NSS should focus on helping the user to understand the bidding process (e.g. graphical representation of the bidding including history of bidding) rather than proposing the next bid.

After watching each video the participants individually specified their level of agreement with each claim on a 7-point Likert scale, and provided comments. We chose this method to give everyone a chance to think about their own attitudes and opinions in silence. As pointed out by e.g. (Carey, 1995) less confident members may be encouraged to disclose more when having written down their views in advance.

With regard to the organization of the researchers, we had three researchers present in every focus group session. One was appointed to be the moderator and the other two were observing and taking notes to capture what was happening between the members of the group, but they did not interrupt the flow of the discussion between the participants. We chose for this set-up to avoid any influence by the researchers. Once every member finished writing their comments the moderator started a group

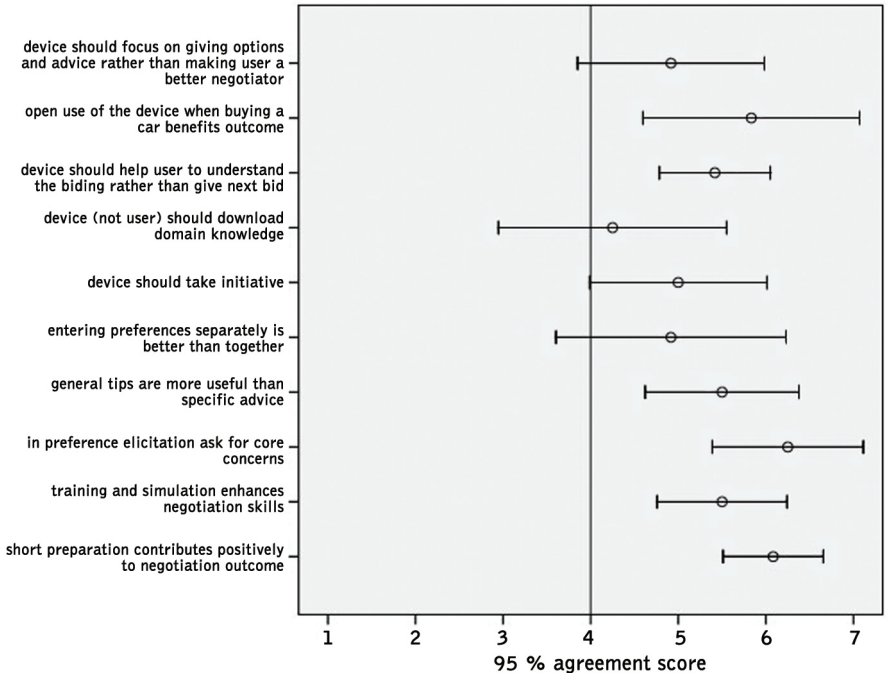


Figure 3.3: Mean values of agreement with claims (1=strongly disagree, 7=strongly agree).

discussion, by asking the participants in turn to react to the claims and discuss their ideas with the others. The moderator stimulated the discussion without enforcing any existing views from the project team. The discussion was audio-recorded for later analysis.

3.5.2 Results

Our approach results in two types of data, i.e., qualitative discussion data in form of written notes and quantitative data from the questionnaires. To analyze the questionnaire data (values on a Likert scale) we used a standard mean value calculation. Figure 3.3 presents the average level of agreement of the experts with the claims that were presented in the questionnaire. Considering the 95 percent confidence interval and the value four as the middle of the scale the results suggest that the majority of the experts leaned towards agreeing with the following claims: 2) open use of the device when buying a car benefits the outcome; 3) the device should help the user

to understand the bidding rather than giving the next bid; 7) general tips are more useful than specific advice; 8) in preference elicitation ask for core concerns (instead of specific values); 9) short training and simulation enhances negotiation skills; and 10) short preparation contributes positively to negotiation outcomes. The qualitative data explains the rationale behind these positions and provides additional ideas.

For the analysis of the qualitative data we used a method similar to interpretative phenomenological analysis (Smith and Osborn, 2003), which is a bottom-up method often used in psychological qualitative research. The idea is to go through the data from one focus group to gather emerging themes from the text. Themes can be recurring ideas, thoughts or feelings from the participants. These themes are then clustered together and superordinate concepts might emerge. This process is repeated for the other focus groups and finally, the superordinate themes are compared and converged to final themes or theories, i.e. in our case transformed into design guidelines.

We analyzed the sessions separately on the basis of the notes by at least two researchers. The recordings from the sessions were only used in case the notes were not clear enough or incomplete. Every idea or attitude was written on a post-it note. Repeated ideas were not written down again, as we were not trying to get empirical generality and furthermore, in groups people tend to agree with or repeat thoughts and ideas.

To define the general themes that can be transformed into design guidelines four researchers independently clustered the post-it notes. We intentionally included one researcher unrelated to the project. Therefore, we could compare unbiased data with the data from the project researchers. Themes thus identified were then compared across all focus groups. Several themes came up that provided first ideas about people's attitudes and requirements towards NSS. In the following we present the main themes (bold) from the discussions in detail.

An NSS device adds higher value in the preparation and training phase than during a negotiation. Training needs to be interactive and the NSS needs to react intelligently. All experts across the groups agreed on the fact that any preparation for a negotiation is useful. However, some experts mentioned that a technical device should add more value to the preparation than just reading a book on negotiation. They emphasized the importance of training and simulation and pointed out that the system needs to be able to respond to the user in an intelligent way. In detail, one idea that was mentioned was that the system needs to make people aware of what they can negotiate about. In addition, the system needs to ask questions to the user similar to

the ones asked in job negotiations. In one group it was mentioned that multiple short sessions of preparation might be better than one long one.

In a face-to-face situation it is hard for the user to focus both on the device and the opponent. Most experts were of the opinion that an NSS should not be used in face-to-face negotiations. Especially the job negotiation experts mentioned that the way the applicant or employee presents him/herself is important as well as focusing on the negotiation partner. While using a device the interaction with the opponent becomes awkward and might be embarrassing. Furthermore, the experts were concerned that understanding and processing the device's information and advice takes too much time and is too much cognitive load for the user in a face-to-face situation.

The context including atmosphere, non-verbal communication and emotions plays a major role for the negotiation process. In two focus groups it was emphasized that especially in job negotiations the non-verbal communication and the atmosphere in the room play an important role. Furthermore, emotions influence the decision-making process and the course of negotiation. This means that the system needs to be able to obtain this context information and take it into account when reasoning about next steps. People are generally better at interpreting emotions, non-verbal communication and atmosphere than computers. One way of enabling the system to understand the context is to build a context model within the system and let the user enter information about the context during the negotiation. To reduce the data that the user needs to feed into the system other techniques like emotion recognition or using (e.g. sound) sensors might be a solution.

The NSS is strong in the rational part of a negotiation, by offering new options and for storing and managing data. It should provide domain knowledge in terms of facts that the user can use to persuade. Most experts agreed that the strength of a device would lie in handling the rational part of a negotiation. It can store and manage vast amounts of data, deal with the computational complexity during the bidding and offer new options to the user. Furthermore, domain knowledge should mainly include facts, such as prices or salaries, which the user can use to persuade his/her opponent.

Both generic and specific advice is useful but needs to be applied carefully. One of our claims was that generic advice is more useful than specific advice. The attitude towards this claim differed between the experts. Many of them saw a danger in specific advice because if the system cannot sense the context specific advice is often inappropriate. Generally both generic and specific advice could be useful but is

dependent on the negotiation phase and the capabilities of system and user.

The NSS needs to adapt to the user's behavior and his knowledge or experience. At several points in the discussion it was mentioned that the system advice or reactions need to be adapted to the experience of the user and his/her behavior. Regarding advice given by the system it was mentioned that novice users who are not good negotiators should get more specific advice whereas more advanced users are able to apply more generic advice. During the bidding the system should adapt its behavior to that of the user and recalculate the next bids in case the user changed his/her strategy.

Interruptions are seen controversial. Time-outs, however, are good. The majority of the experts thought that active interruptions by the system through vibrating and beeping during a tense situation are not useful. The users would either ignore the system or become more upset. However, most experts agreed that time-outs are very useful for reflection of the negotiation process. As the user is not always aware of when to take a time-out the system should suggest it.

Preferences of collaborating partner's should be put in separately. Across the focus groups there was a consensus that in the process of generating a preference profile for collaborating partner's, e.g. couples, they should put in the their preferences separately. That avoids that one partner is more dominant than another. In our scenario we proposed that the system then merges the preferences and shows the clashes to the users. The experts did not agree on doing it this way. They pointed out that showing those clashes triggers arguments between the partners instead of a discussion about underlying values. It is more important that the partners talk about such values and come to a conclusion. The system could also directly suggest solutions. It was also proposed that a user indicates the importance of every preference.

Besides these functionality-oriented themes, the discussions showed that the experts' attitudes towards NSS differed widely and that social contexts might play a role when choosing to use a system or not. Social acceptance became a topic in several groups although we did not specifically ask for it in the questionnaire. Especially the question whether it was acceptable in a face-to-face situation was discussed. One hypothesis was that the social acceptance would correlate with the age group of the users. The experts assumed that younger generations due to growing up in a world of mobile technology are more used to people using mobile devices in public and being interrupted by e.g. mobile phones. While being a plausible assumption we conducted focus groups with young people to see if it would be confirmed.

3.6 User Focus Groups

To investigate the attitudes of young people towards mobile NSS, we had focus group discussions with 20 high school students aged 16–18. In these user focus groups we shifted the focus from functionality oriented discussions to the social acceptance of the NSS in the different use contexts.

3.6.1 Set-up and Procedure

The session was split into two parts, i.e. group discussions in smaller groups and a discussion with all students. We first divided the high school students into five groups of four students each. We assigned one researcher of our project to each group to act as a moderator and observer. In order not to bias the participant these researchers were instructed to intervene as little as possible, i.e. only to start the discussion and in cases the discussion stopped. At the same time they were taking notes for later analysis. Each group watched one of the five scenarios. Every participant received a short questionnaire with three statements. Two focused on the social acceptance: (1) I would use the PN in the situation shown in the video, (2) I think that it is socially acceptable to use a PN in this situation. The third statement addressed a functionality aspect of the particular scenario (similar to expert focus groups). All statements were rated by the participants on a 7-point Likert scale (1=totally disagree, 7= totally agree) after watching the video. Before starting the discussions the moderator asked every participant to explain their ratings. Group discussions in the small groups lasted about 15 minutes.

At the beginning of second part of the session we asked one group member from every group to explain the situation shown in the discussed video and the main points of the discussion to the other groups. This was done to make sure every participant knew about all five scenarios and could form an opinion about the social acceptability of each of them. Next, a moderator encouraged a discussion between all 20 students, mainly focused on social acceptance, which took about 30 minutes.

3.6.2 Results

Our initial hypothesis that younger generations think a mobile NSS in public or face-to-face situations is socially acceptable could not be confirmed by the focus groups with high school students. Especially in the job scenario with the boss some students thought a PN would be very strange and unsocial. Others thought that the stealth mode function can be used as a long as the other party does not notice that you have a PN.

In any case, it would stop the communication from its natural flow. This would also be the case on the phone. Nevertheless the students believed it to be more acceptable on the phone, since the other party does not see the NSS. Generally, students tended to see it as more socially acceptable in cases where the other party does not know about the PN. However, if everyone was using a PN the students thought it would be fine to use one. Overall, we could see that the students were very critical towards the PN and its use. Many emphasized that it is important that the user stays independent from the device instead of following its advice blindly. Furthermore, it is of importance that the advice is presented in a way that is comprehensible to the user. However, the students also saw the strength in a PN. They mentioned that it is helpful in the training and to organize things. Some students believed that insecure people would feel more supported and confident with a PN.

In general, focus groups provide large amounts of qualitative data, due to the dynamic nature of the group and the contextual setting. As discussed in detail in (Carey, 1995; Sim, 2001) the data analysis of focus group data is delicate. Researchers have to be aware that focus groups are not meant to find consensus within the group. Therefore, focus groups data is not meant to lead to an empirical generalization but rather give an impression of attitudes of a specific group of people towards a topic or new technology. According to (Sim, 2001), the data from focus group can provide theoretical insights with sufficient level of universality to be projected to comparable contexts. To complement these initial impressions with empirical data and get a deeper insight into what exactly the influential factors to social acceptance are we designed an online survey.

3.7 Social Acceptance Survey

From the focus groups we already got some support for the hypothesis that the use context is influential to the social acceptance. However, other factors were mentioned, such as characteristics of the possible user (age, novice negotiator, etc.), the mode in which the device is used (e.g. stealth mode), or social pressure (“If everyone had a PN it would be okay to use it.”). In the following we present a number of research questions that led our design of a questionnaire to investigate the acceptance of mobile NSS. Next we will describe the underlying model of the questionnaire, the survey and its results (please see (Pommeranz, 2010) for more details).

3.7.1 Research Questions

Overall, the question is: which are the factors that influence the social acceptance of NSS? We looked at several detailed research questions. **RQ 1**: Is there a relationship between the user characteristics and usefulness, attitude towards negotiation, behavioral control and social acceptance? The user characteristics include demographic data and experience in computer usage and with negotiations. We expect that age and possibly gender influence the acceptance of a mobile NSS in different situations. In the focus groups we investigated whether younger people are more open to technology use in public places and social situations than older people because younger generations grow up with technology around them. To get a more definite answer to this question we also included it in this research. This is reflected in **RQ 1a**: Is there a negative impact of the user's age on the acceptance of a NSS in a face-to-face situation?

Based on the results of the focus groups mentioned and groups with 40 middle-aged women, we expect that people with low negotiation skills and a negative attitude towards negotiation are more likely to use an NSS. Due to their own lack of knowledge about negotiations or insecurity they might find an NSS more useful than people, who enjoy negotiating and consider themselves good at it. This leads to the questions: **RQ 2**: Is there a negative relation between a person's attitude towards negotiations and the attitude towards NSS? **RQ 2a**: Is there a relationship between on the one side negotiation skills and experience and on the other side the attitude towards negotiations?

We believe that the acceptance of a NSS in a social context has an impact on the intention to use it. The social acceptance is measured by two variables, one describing how acceptable people find it to use an NSS in a situation (SN1) and the other describing in how far they believe that the opponent would find it acceptable (SN2). Whereas in a face-to-face situation it might play a big role what the opponent thinks, it might become less influential in a phone scenario. Therefore, our last research questions are: **RQ 3**: Is there a relationship between the social acceptance of an NSS and the intention to use it? **RQ 3a**: Does the negotiation situation determine the social acceptance?

3.7.2 The Model

To study social acceptance of mobile NSS empirically, we first developed a model based on existing models and our research questions presented above. This model was the basis for the questionnaire that we used in an online survey.

Since we wanted to predict the intention of people to use a NSS, we could make

use of existing, often used models from social psychology and information systems. The Theory of Planned Behavior (TPB), developed by Ajzen (1991), is a well known model in social psychology to explain the link between attitudes and actual behavior. In this model the behavior is influenced by the intention to perform the behavior. This intention again has three influential factors, namely the attitude towards the behavior, subjective norm and perceived behavioral control. Attitude is defined as positive or negative feelings towards performing the behavior. The subjective norm is an individual's perception of others' beliefs whether he or she should perform the behavior. Perceived behavioral control is an individual's perceived ease or difficulty of performing the particular behavior. The latter also has an influence on the actual behavior.

Whereas the TPB is a general model predicting behavior, the Technology Acceptance Model (TAM) (Davis, 1989) is a more specific model used in information systems research for predicting the acceptance of a technology. The model has been widely used, see e.g. (Wang and Benbasat, 2005; Yu et al, 2003), and extended for specific applications (Shih, 2004; Wixom and Todd, 2005). It identifies perceived usefulness and perceived ease of use as two factors that influence the intention to use a system and its actual use. Both TPB and TAM are extensions or adaptations to the Theory of Reasoned Action introduced by (Ajzen and Fishbein, 1980). Both models predict the actual behavior or use of a system. However, we would like to measure only the intention to use a mobile NSS. In addition, we believe the models need to be extended to fit the more specific negotiation context. Therefore, we used the models as a basis for creating our NSS social acceptance model shown in Figure 3.4. In the next section we will explain how we combined and extended the models in detail.

3.7.3 TPB and TAM extended

Since our study takes place before the implementation of our envisioned NSS and is meant to inform the first designs of it, we are not able to measure the actual use of such a system. Furthermore, other factors that are meant to be perceived by the users, i.e. ease of use, usefulness and behavioral control are not easily measurable either. We decided to leave out the ease of use since this can only be experienced during a real interaction with the system. Usefulness and behavioral control, however, are factors that can be measured by providing the users with detailed visualizations of the system's use. Therefore, we showed videos or storyboards of the five scenarios described above. The remaining factors are, therefore, usefulness, attitude towards NSS, behavioral control, subjective norm and the intention to use the NSS, with their

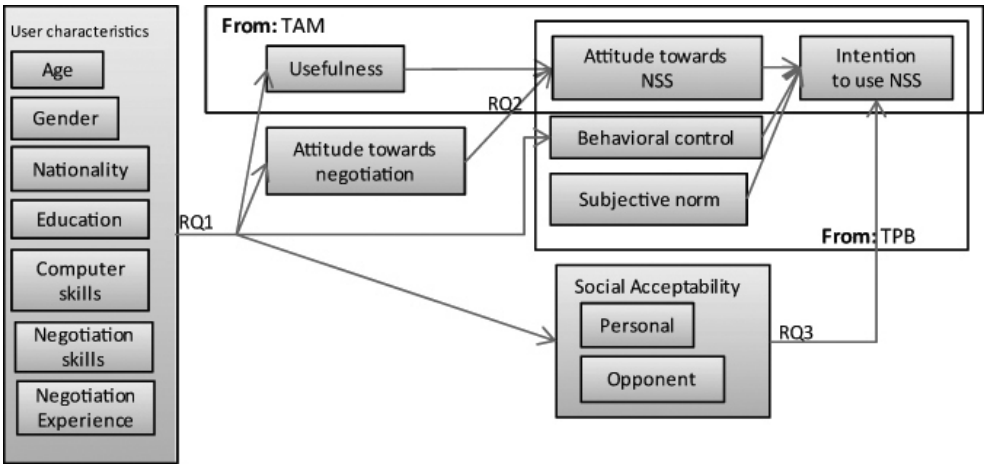


Figure 3.4: NSS Social Acceptance Model

relations taken from the original models as shown in Figure 3.4. Based on our research questions we added a number of factors that might be of influence in the negotiation domain. We added the general attitude towards negotiations as an influential factor of attitude towards NSS. As mentioned earlier the use of such systems might depend on different situations and how socially acceptable it is to use a system in that situation. Therefore, we added social acceptance as an extra factor influencing the intention to use. Last, we added a number of user characteristics including: age, gender, nationality, education, computer and negotiation skills and experience.

3.7.4 The survey

The questionnaire structure

The questionnaire is based on the model shown in Figure 3.4. For details about the constructs and questions, see Appendix A. After a short introduction we collected the user characteristics. The factors intention to use (IU), subjective norm (SN) and social acceptability (SA) were measured after each scenario presented to the respondent. At the end of the survey we collected more general information about the attitude towards NSS (PNA), including behavioral control (BC) and usefulness (USE). For the majority of questions we asked respondents to rate their agreement with a number of statements on a 7-point Likert scale and for an explanation of the ratings after each scenario to explore why people accept the system in one scenario but not in another.

Versions

We setup a Dutch version with short videos (3 min.) and a Dutch and English version each with screenshots from the videos and text explaining the situation. The version with videos took about 45 minutes to fill in and the picture versions 10–15 minutes. To avoid order effects we shuffled the order of scenarios and statements.

Survey Distribution and Response

With NetQuestionnaires (www.netquestionnaires.com) we administered and distributed the survey online. We used an opportunity sample strategy to select participants for the study. We took advantage of personal networks and online forums to invite people to participate. The questionnaire was approached by 365 people. 178 started filling in the questionnaire, 120 (74 male, 46 female) from 18 countries completed it, 72 the English, 31 the Dutch version with videos and 17 with pictures. The most represented countries were the Netherlands (48), Sweden (19), Germany (15) and Greece (10). The age span ranged from 20 to 68 (M = 32.28, SD = 10.36). Participants are mostly familiar with computer usage, with the average number of hours spent at the computer being 44.86 (SD = 20.14) and highly educated (102 with university degrees). The negotiation experience of the sample is rather low. Only about a fourth of the participants are regularly engaged in negotiations in their jobs (31 participants). On average participants have bought 0.65 (SD = 0.97) and sold 0.47 houses (SD = 2.43) and have had less than seven job interviews (M = 6.65, SD = 10.33).

3.7.5 Measurements of Contructs

For an overview of all constructs used in the questionnaire see Appendix A. We used Cronbach's alpha to test the reliability of the constructs usefulness (USE) (.95), and behavioral control (BC) (.72) and calculated aggregated measures for both including all original items. The Cronbach's alpha for attitude towards negotiation (NAT) including all four original items is very low (.04), but increases to .69, if the items NAT 1 and NAT 4 are deleted. Therefore, we decided to keep only the items NAT 2 and NAT 3 and combined them to an aggregated measure. For the construct negotiation skills (NSK) we keep the three items NSK 1, NSK 4, NSK 5 reaching a Cronbach's alpha of .71, while removing NSK 2 and NSK 3. The reliability of social acceptance (SA) was measured per scenario (Cronbach's alpha between .81 and .94). We did not calculate an aggregated measure for the acceptance, but kept them separate in the further analysis.

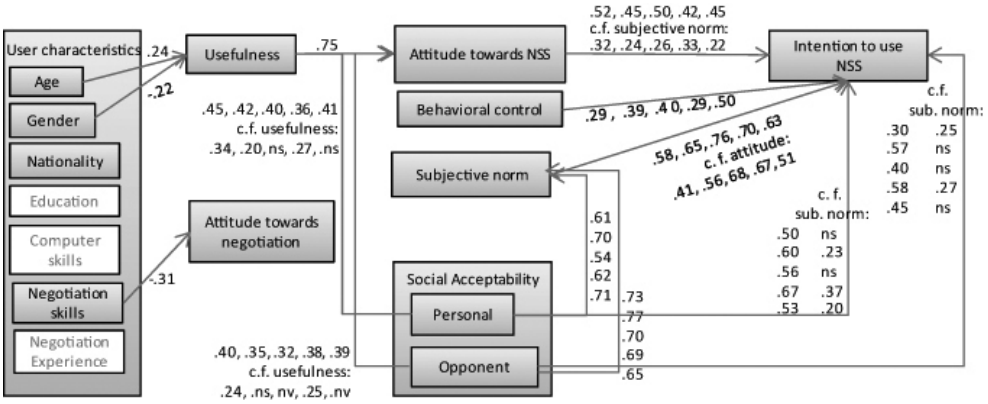


Figure 3.5: Model with (partial) correlations, 5 numbers: per scenario, cf = controlled for, ns = not significant, nv = no value

3.8 Survey Results

3.8.1 Data Analysis

We used correlation analysis to check our hypotheses. Significant correlation coefficients can be found in Figure 3.5.

User's background

Our first research question was "Is there a relationship between the user characteristics and usefulness, attitude towards negotiation, behavioral control and social acceptance?" With regard to user characteristics (left column in Figure 3.5) we only found a significant positive correlation (.24) between age and usefulness and a negative one (-.22) between gender and usefulness. Computer skills and negotiation experience were not correlated with usefulness, attitude towards negotiation or behavioral control. We removed the item education from the model, since our data was not heterogeneous enough to draw any conclusions on the effects of education level. We also removed nationality because the data was not equally distributed. Furthermore, the second set of research questions was "Is there a negative relation between a person's attitude towards negotiations and the attitude towards NSS?" and "Is there a relationship between on the one side negotiation skills and experience and on the other side the attitude towards negotiations?" We did not find a significant correlation between a person's attitude towards negotiations and the attitude towards NSS. With regard

to the second question we found that negotiation skills are negatively correlated (-.31) with the attitude towards negotiation opposing our initial hypothesis. However, negotiation skills were rated subjectively by the respondents themselves, which might not correspond to their actual negotiation skills. This issue needs further research.

Usefulness, Subjective Norm and Social Acceptance

We found a positive correlation (.75) between usefulness and the attitude towards NSS, which confirms the relationship predicted by TAM. Considering our third research question "Is there a relationship between the social acceptance of an NSS and the intention to use it?" we can say the following. We found that social acceptance, (personal (SA 1) and opponent (SA 2) view), is correlated with the attitude towards NSS and the intention to use for all scenarios. However, when controlled for usefulness in the first case and subjective norm in the second, the correlations are either weaker or not significant. This suggests that the attitude towards an NSS is mainly influenced by how useful people consider it. The intention to use the system depends mainly on the subjective norm, i.e. whether others relevant to the respondent believe he or she should use it.

The dominance of subjective norm was further supported by a regression analysis for each individual scenario. We used a stepwise method with the dependent variable intention to use NSS in a particular scenario and the following independent variables: attitude towards negotiation (NAT), behavioral control (BC), subjective norm (SN) and social acceptance (SA) (see right part of Figure 3.5). Table 3.1 and table 3.2 given an overview of the regression models. Table 3.1 shows the strength (R) of the relationship between the intention to use the NSS and the independent variables included in the model, which can be seen in table 3.2. R^2 represents the extent to which the included independent variables can predict the intention to use. Besides the included variables table 3.2 also shows the coefficients. We can see that subjective norm has the major influence in predicting intention to use in all scenarios. In the car dealer scenario it is even the only variable included in the model (β= .58, t(118) = 7.67, p $<$ .001). In the collaborative preparation and the phone scenarios behavioral control were also included in the model. In the face-to-face and the train scenario behavioral control as well as social acceptance was included in the model. Whereas face-to-face the social acceptance is the second strongest indicator before behavioral control, in the train scenario it is the other way around. This is not surprising since in the situation with the boss social rules are much more important and can have stronger consequences than when sitting on a train.

Table 3.1: Results of regression analyses *per scenario*, R = strength of the relationship between the intention to use NSS and the independent variable(s) (see table 3.2). R^2 = extent to which the independent variables can predict intention to use.

Scenario	R	R^2	Adj. R^2	SE (Std error)	df_{reg}	df_{res}	F	p
Train	.684	.467	.453	1.341	3	116	33.90	<.001
F-2-F	.762	.580	.569	1.143	3	116	53.39	<.001
Coll. Preparation	.674	.455	.439	1.434	2	69	28.78	<.001
Phone	.764	.584	.577	1.151	2	117	82.15	<.001
Car Dealer	.577	.333	.327	1.521	1	118	58.83	<.001

Table 3.2: Estimated coefficients of regression models for each scenario, B and β are the regression coefficients, unstandardized and standardized (same units) respectively. VIF stands for Variance Inflation Factor and measures the impact of collinearity among the variables

Scenario	B	SE	β	t	p	VIF
Train						
Constant	-.77	.644		-1.19	.24	
SN	.46	.111	.394	4.10	<.001	2.01
BC	.38	.126	.237	3.04	.003	1.33
SA	.22	.105	.188	2.11	.04	1.73
F-2-F						
Constant	-1.25	.523		-2.39	.02	
SN	.52	.097	.441	5.33	<.001	1.89
SA	.36	.088	.339	4.08	<.001	1.91
BC	.24	.095	.157	2.56	.01	1.04
Coll. Preparation						
Constant	-.41	.794		-.51	.61	
SN	.67	.102	.595	6.53	<.001	1.05
BC	.34	.144	.215	2.36	.02	1.05
Phone						
Constant	-.37			-.71	.48	
SN	.82	.076	.704	10.88	<.001	1.18
BC	.21	.102	.131	2.03	.05	1.18
Car Dealer						
Constant	1.01	.383		2.63	.01	
SN	.70	.092	.577	7.67	<.001	1.00

People using mobile devices on a train are a common sight and therefore social acceptance has less influence. More interesting is that in the other three scenarios social acceptance is not included in the model. In the phone and collaborative preparation scenario this might be due to the lack of a public setting.

In the survey participants were asked after each scenario whether they wanted to explain their ratings. Per scenario between 55 and 61 respondents entered comments. At the end of the questionnaire another entry field allowed to give overall feedback (41 respondents entered comments here). Looking at the comments respondents gave voluntarily, we get deeper insight into how people see social acceptance considering the opponent's view in the different scenarios. People tend not to care whether the opponent accepts the NSS if they are not in eye contact ("This [on the phone] seems like the best application of the NSS, because it is invisible to the 'opponent'.") In the face-to-face scenarios people value the opponent's opinion highly. In the car dealer scenario some respondents doubt the acceptance of the NSS by the opponent. However, usefulness, the competitive situation ("I think the opponent will accept it because otherwise people would go to the competitor.") or the ability to put pressure on the opponent ("I like the secret weapon!") cause people to care less about the opponent. In the job scenario between an employee and her boss, most respondents are worried about the opponent's opinion on the use of an NSS. The comments show different views considering not being honest ("I think it is not acceptable because she lies about using an NSS."), impolite ("It's very impolite to use an electronic device during a face-to-face negotiation."), embarrassed ("I would be embarrassed to use an NSS in this situation."), nervous ("Stealth mode would make me extremely nervous.") or appearing weak ("In a face-to- face negotiation this would make you look like you cannot think for yourself."). A dominant opinion was that the interaction with the device will interrupt the communication flow ("The boss could get angry for not paying attention, the communication would be disturbed").

With regard to our last research question "Does the negotiation situation determine the social acceptance?", we found that the social acceptance indeed depends on the situation in which the NSS is used as shown in Figure 3.6. Whereas most scenarios have an average rating above the scale's mean (4), the face-to-face situation with the boss got a low rating (3.06) lying significantly below the average ($t(119) = -6.25$, $p < .001$). This means, in the latter scenario people do not accept the use of an NSS. The situations which are most favorable for NSS use are negotiations on the phone and preparation on the train. At the car dealer or during the collaborative preparation NSS are accepted, but the average rating is closer to the neutral value.

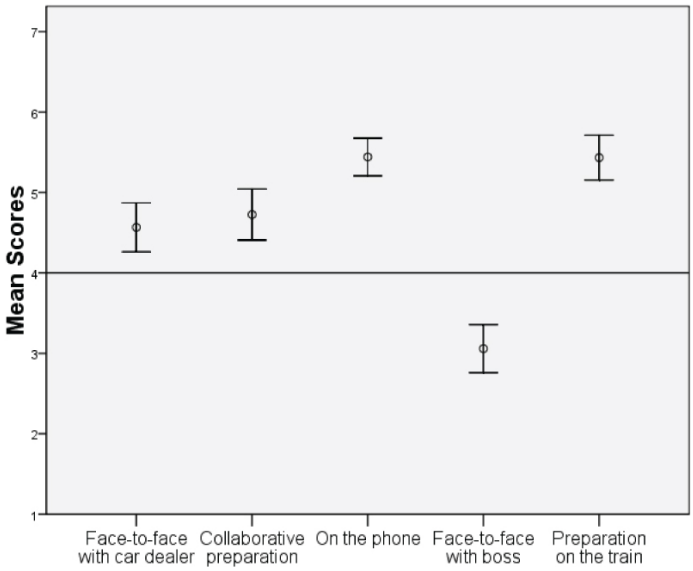


Figure 3.6: Mean social Acceptance ratings (1=low to 7=high)

3.8.2 Limitations

The online survey presented has a few limitations. First of all, the participants were not offered the chance to interact with an implemented system. We used the TAM model because it is well-known and a valid model to predict acceptance of new technology. We have to emphasize, however, that this model is based on constructs which can be perceived by the user when interacting with a real system. We are at the beginning of the development of a novel NSS. Therefore, no implementation was available. Furthermore, this study intended to inform the design process of a new NSS, instead of evaluating an existing design. To avoid misinterpretations we excluded variables from the model that could not be perceived by only watching videos or seeing pictures, e.g. perceived ease of use. However, we would like to emphasize that we have to bear in mind that generally the added value a system can bring to the user's activity may strongly influence its acceptance. As ease of use was not measured and usefulness was not perceived directly by using the system, we cannot make general claims about this aspect. In our study we focused rather on the use situations than the functionality the NSS could offer. We believe that by showing scenarios of use contexts in the questionnaire we found a good way to give participants a vision of what the system could be able to do, but on such a level that it does not distract from the focus on

the situation. We believe that people could get a feeling for the usefulness of the system and judge whether they would be able and willing to use it. The results of the survey pointed to social acceptance and subjective norm as major factors influencing the intention to use the NSS. There were only little indications (positive correlation between usefulness and attitude towards NSS) that people believed in the added value of the system. The fact, that our hypothesis that people with less negotiation skills and negative attitudes towards negotiations would have a positive attitude towards the NSS could not be confirmed, may signal that these people did not find the NSS particularly useful. In the user focus groups the students were critical towards the presented NSS functionality in the different scenarios. While they had the opinion that it was only useful if people were not dependent on the system and if the advice was intelligent and comprehensible, they were positive towards using it as a trainer and to organize data. We believe that the issue of usefulness needs to be investigated in more detail with a follow-up study using first prototypes of the system.

Further limitations concern the number of participants in the survey and the opportunistic sample. Unfortunately, these aspects did not allow us to make any general claims about the acceptance of NSS with regard to cultural or educational backgrounds or differences depending on age groups. Despite this, we believe that we offer interesting results that put NSS into a different light. The fact that both subjective norm and situation dependency were major influential factors needs to be taken into consideration when designing new NSS, especially for mobile use.

3.9 Design Implications

In the following we will point to several design implications resulting from the focus group discussions and the survey results.

3.9.1 Implications from Focus Groups

From the focus groups with negotiation experts we could extract several themes, mostly focused on the functional requirements for a mobile NSS. In summary: the preparation phase of a negotiation and the actual negotiation with an opponent require different interaction styles. In the preparation phase NSS should provide a negotiation training that is rich, content-full and contextual. Preferably it should make use of an adaptive scenario including socially intelligent opponents to provide a real setting. During the negotiation with an opponent, on the contrary, the system should provide concrete, personalized advice regarding offers and generic advice regarding

the negotiation process with easy interpretable hints. The interaction style in this case should be as little interrupting as possible. The major implication of these guidelines is that NSS need to have intelligence and reasoning capabilities in order to process the information entered by the users and give personalized output. Furthermore, the system needs to possess an accurate user model that is updated during the interaction to be able to adapt to the user. Furthermore, the interaction styles need to be carefully selected for each phase of the negotiation. Based on these themes we constructed the following 12 design guidelines for NSS development (Pommeranz, 2009):

1. An NSS should support interactive preparation sessions of different lengths.
2. The preparation module should have a simulation mode in which the user interacts with an intelligent negotiation agent.
3. The cognitive load of the information representation provided by the NSS during a face-to-face negotiation should be minimized.
4. In the training module the user should be trained on being aware of the context.
5. Advice from an NSS should consider information about the context of the negotiation.
6. An NSS should support the user by calculating bids and offering new options to negotiate on.
7. It should have a data storing and managing function that gives the user easy access to the information needed at a certain point in time.
8. An NSS should generally provide the user with more generic advice that the user can apply to the situation he/she is in.
9. An NSS should be able to adapt to the user's skill level and experience and more in specific to the user's bidding behavior.
10. System advice should be based on the capabilities of the user to apply them in practice.
11. An NSS should suggest time-outs at appropriate stages in the negotiation process.
12. Partners should put in their preferences separately and assign an (emotional) value to each preference.

3.9.2 Implications from Social Acceptance Survey

From the social acceptance survey we learned that not only functionality and usefulness play a role, but also social aspects like the subjective norm and social acceptance. An NSS is not only a tool people use to fulfill a certain task but it is a social device depending on the use context. Therefore, the designer has to determine in which context the device should be used and fit the design to the context and its social norms. Furthermore, our survey has shown that the respondents value the opinions of close friends or family highly, both for deciding whether to use an NSS and when taking decisions during the negotiation. Some respondents mention explicitly that they consult others before an important negotiation. ("I would take others' opinions into consideration as well, [...]", "In buying something like a car [...] I get advice for prices online, from friends.") This behavior made us contemplate about the idea to create NSS that are connected to social networks. Friends using the same type of NSS could be connected to each other, and whenever one needs to take a decision they could provide help or generally comment on each others' actions.

Another idea is storing negotiations within this network in a database that every NSS can access. This will enable users to see what strategies friends used in similar negotiations. These ideas fit social computing trends (Parameswaran and Whinston, 2007) by bringing mobile information spaces to the user and using social networks to enhance the system's functionality. Also, if people like to ask friends for advice when negotiating, a good NSS should be designed to behave in a similar manner. Surely, there are more ways designers can think of to make NSS more social devices.

3.10 Conclusion

We presented our steps in gathering requirements for the design of a new kind of mobile NSS including expert and user focus groups and an online survey aimed at determining the social acceptance of such a system. The focus groups were used to get a first impression of people's attitudes towards and functional wishes for mobile NSS. While we focused more on functional requirements in the expert focus groups (due to their negotiation expertise), the social acceptance in different use contexts became the main point in the user focus groups. The focus groups provided a lot of interesting qualitative data and gave first hints to which aspects were important for people and might lead to an acceptance of the system. We extracted 12 design guidelines for NSS from the qualitative data.

To support ideas from the focus groups and further investigate the concrete factors

leading to an acceptance we designed a questionnaire based on a NSS social acceptance model. We developed this model as a combination of the TAM and TPB models extended by a number of factors relevant specifically for NSS. With the help of the questionnaire we collected data from 120 respondents with little negotiation experience in an online survey. We learned that when designing NSS social issues cannot be neglected. Our survey shows that the use context of an NSS is an important factor influencing its social acceptance. The survey's respondents would not accept the use in face-to-face situations when the relationship to the opponent was important, i.e. with one's boss. However, when the relationship is less important, i.e. with a car dealer, it is more accepted. In situations in which the opponent is not aware of the NSS, e.g. on the phone, it is most accepted. Surprisingly, the subjective norm is the most dominant factor influencing the intention to use a mobile NSS. People value opinions of their close ones highly when deciding whether to use an NSS and also ask them for advice when negotiating. Some implications of these results were mentioned (section design implications). However, we believe that there is far more room for designers to address these aspects in their designs in diverse ways.

We were able to obtain our results by giving people a vision of how a new kind of mobile NSS could be used by the help of filmed scenarios. This enabled us to inform the design process of our envisioned system in an early stage before first decisions and implementations have been made.

Our current work involves implementing a first prototype of a mobile NSS following the guidelines named above. The main focus lies on a good preparation for the negotiation by offering a preference elicitation interface that adapts to the users' needs and cognitive skills, as well as an interactive training with a virtual agent. Support throughout the different negotiation phases will be provided by a virtual coach, who behaves like a knowledgeable friend and reacts to the current context of the user. Ideas for connecting users of the NSS and store negotiation data in databases to be accessed by every user are left for future iterations of prototype development.

After implementing the prototype we will be able to investigate more factors, which can only be perceived during the interaction with a running system, e.g. ease of use. Other aspects to be considered for future research are the influences of educational and cultural background of the user on attitudes towards negotiation and NSS.

Overall, when designing novel, mobile NSS we should aim for creating NSS not merely as tools but as social devices considering the use context and social networks.

3.11 Bibliography

Ajzen I (1991) The theory of planned behavior. Organzational Behavior and Human Decision Processes 50:179–211.

Ajzen I, Fishbein M (1980) Understanding Attitudes and predicting Social Behavior. Englewood Cliffs, Prentice-Hall.

Bruseberg A, McDonagh-Philp D (2002) Focus groups to support the industrial/product designer: a review based on current literature and designer's feedback. Applied Ergonomics 33(1):27–381.

Bui T (1994) Evaluating negotiation support systems: A conceptualization. In: HICSS 1994, IEEE Press.

Carey M (1995) Comment: concerns in the analysis of focus group data. Qualitative Health Research 5:487–495.

Carroll J (2000) Making Use: Secenrio-based Design of Human-Computer Interactions. MIT Press Cambridge.

Davis F (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 13(3):319–339.

Goodman J, Dickinson A, Syme A (2004) Gathering requirements for mobile devices using focus groups with older people. In: Designing a More Inclusive World, Proceedings of the 2nd Cambridge Workshop on Universal Access and Assistive Technology (CWUAAT), Springer.

Havard Business School Essentials (2003) Negotiation. Harvard Business School Publishing Corporation, Boston.

Hindriks K, Jonker C (2008) Creating human-machine synergy in negotiation support systems: Towards the pocket negotiator. In Proceedings of the First International Working Conference on Human Factors and Computational Models in Negotiation, HuCom'08, Delft, The NetherlandsHuCom'08.

ITU (2004) Itu internet reports: The portable internet,http://www.itu.int./portableinternet. Tech. rep., ITU.

Kersten G (1999) Negotiation support systems and negotiating agents. In Proceedings of the EKAW'02 Workshop on Knowledge Management through Corporate Semantic Webs.

Kersten G (2004) E-negotiation systems: Interaction of people and technologies to resolve conflicts. InterNeg Research Papers INR 08/04 pp 1–21.

Kersten G, Lo G (2003) Aspire: an integrated negotiation support system and software agents for ebusiness negotiation. International Journal of Internet and Enterprise Management 1(2):293–315.

Kersten GE, Lai H(2007) Negotiation support and e-negotiation systems: An overview. Group Decision and Negotiation 16:553–586.

Ling R (1997) One can talk about common manners!: the use of mobile telephones in inappropriate situations. in themes in mobile telephony. Tech. rep., COST 248 Home and Work group.

Love S, Perry M (2004) Dealing with mobile conversations in public places: Some implications for the design of socially intrusive technologies. In Proceedings of the ACM SIGCHI conference on Human Factors in Computing Systems, CHI'04, Vienna, Austria, 1195–1198.

Mallat N, Rossi M, Tuunainen V, Arni A (2009) The impact of use context on mobile services acceptance: The case of mobile ticketing. Information Management 46:190–195.

Neerincx M (2003) Cognitive task load design: model, methods and examples. In: E. Hollnagel (ed.), Handbook of Cognitive Task Design. Chapter 13. Mahwah, NJ: Lawrence Erlbaum Associates.

Nunamaker JF, Briggs RO, Mittleman D (1996) Lessons from a decade of group support systems. In Proceedings of the 29th Hawaii International Conference on System Sciences Volume 3: Collaboration Systems and Technology, IEEE Computer Society, Maui, Hawaii, 418–427.

Palen L, Salzman M, Youngs E (2001) Discovery and integration of mobile communications in everyday life. Personal Ubiquitous Computing 5(2):109–122.

Parameswaran M, Whinston A (2007) Social computing: An overview. Communications of the Association for Information Systems 19:762–780.

Pommeranz A, Brinkman WP, Wiggers P, Broekens J, Jonker CM (2009) Towards Design Guidelines for Negotiation Support Systems: An expert perspective using scenarios. In Proceedings of European Conference on Cognitive Ergonomics, ECCE'09, Helsinki, Finland, 27.

Pommeranz A, Brinkman WP, Wiggers P, Jonker CM (2010) Social Acceptance of Negotiation Support Systems. In Proceedings of the 6th Symposium of the WG HCI&UE of the Austrian Computer Society,USAB'10, 55–69.

Rangaswamy A, Shell G (1997) Using computers to realize joint gains in negotiations: Towards an electronic bargaining table. Management Science 43(8):1147–1163.

Schoop M, Jertila A, List T (2001) Negoisst: a negotiation support system for electronic business-to-business negotiations in e-commerce. Data and Knowledge Engineering 47(3):371–401.

Shih H (2004) Extended technology acceptance model of internet utilization behavior. Information & Management 41(6):719–729.

Sim J (2001) Collecting and analysing qualitative data: issues raised by the focus group. Journal of Advanced Nursing 28(2):345–352.

Smith J, Osborn M (2003) Interpretative phenomenological analysis. In: Qualitative Psychology: A Practical Guide to Methods. Sage, London.

Srivastava L (2005) Mobile phones and the evolution of social behavior. Behaviour and Information Technology 24(2):111–129.

Swaab R, Postmes T, P N (2004) Negotiation support systems: Communication and information as antecedents of negotiation settlement. International Negotiation 9:59–78.

Thomson L (2005) The Heart and Mind of the Negotiator. Pearson Prentice Hall.

Vetschera R, Kersten G, Koeszegi S (2006) User assessment of internet-based negotiation support systems: An exploratory study. Journal of Oganizational Computing and E-Commerce 16(2):123–132.

Wang W, Benbasat I (2005) Trust in and adoption of online recommendation agents. Journal of the Association for Information Systems 6(3):72–101.

Wixom BH, Todd PA (2005) A theoretical integration of user satisfaction and technology acceptance. Information Systems Research 16(1):85–102.

Yu JLC, Liu C, Yao JE (2003) Technology acceptance model for wireless internet. Internet Research 13(3):206–222.

CHAPTER 4

USER-CENTERED PREFERENCE ELICITATION

From the negotiation literature and the expert focus groups discussed in the previous chapter we learned that the preparation phase of a negotiation is crucial to the quality of the negotiation outcome. It is important that people reach a deep understanding of the negotiation domain, the other party's concerns and most importantly their own values and preferences. The current chapter, therefore, focuses on how to design for user-centered preference elicitation during the preparation for a negotiation.

A problem with regard to preference elicitation in current systems is that they are not suited to the constructive nature of human preferences and are often based on rational, quantitative models that do not match the mental models of people. In this chapter we present three studies dealing with (1) different ways of entering preferences (section 4.3), (2) factors influencing a user's motivation to enter preference details (section 4.4) and (3) participatory design of interfaces that support the human preference construction process (section 4.5).[1]

[1]This chapter is equivalent to: Alina Pommeranz, Joost Broekens, Pascal Wiggers, Willem-Paul Brinkman, Catholijn M Jonker. Designing Interfaces for Explicit Preference Elicitation: a user-centered investigation of preference representation and elicitation. User Modeling and User-Adapted Interaction, Volume 22, Numbers 4-5 (2012), 357–397.

Good acts are like good poems. One may easily get their drift,
but they are not rationally understood.
Albert Einstein (1879-1955)

4.1 Introduction

Web technology and computational intelligence enable the development of systems that assist users in tasks that are cognitively demanding. These smart systems are becoming essential tools for people to deal with information overload, huge search spaces and complex choice sets in different domains, such as product or service recommendations (Adomavicius and Tuzhilin, 2005) or decision support e.g. in health care, real estate, jobs or divorce negotiations (Johnson et al, 2005; Bellucini and Zeleznikow, 2006). Whereas a substantial amount of research in the field of recommender and decision support systems focuses on recommendation algorithms, formal representations and reasoning mechanisms, little research addresses the design of the user interfaces of these systems. Recently, see e.g. the work of (Knijnenburg et al, 2012), the importance of the user interface design and its effects on the user experience of recommender systems has been emphasized. The interface between the user and the system plays a major role in acceptance of the systems as well as the user's trust and satisfaction (Pu and Chen, 2007; Pu et al, 2012). In particular, the method and the interface designed to elicit user preferences influences decision accuracy and the intention to return (Chen and Pu, 2009).

Smart systems need accurate preference models to be able to give useful advice to the user. A preference elicitation interface needs to extract information from the user's mental representation of that preference and translate it into a representation the system can reason with. Preference modeling thus always involves three components: mental representation, elicitation and the system's preference representation. In this article, we focus on the middle part: the preference elicitation interface. Two key issues in preference elicitation are (1) a potential mismatch between the user's mental model of his or her preferences and the system's preference representation and (2) the influence of the elicitation process on the created preference profile. The first issue is a result of the discrepancy between the rational, quantitative models used in systems and the constructive, qualitative mental models of people. Whereas rational models assume that people have stable and coherent preferences that are always known to them, people rather construct their preferences as they go along in the decision task. In addition, it is difficult for people to express their preferences in numerical attribute weights and values as needed by automated systems, particularly if they are not experts in the domain. Whereas people might easily state that they prefer, e.g., more holidays to less holidays specifying this relation in concrete numbers, such as "I like 30 holidays 2.5 times as much as 28 holidays" is not intuitive.

The nature of how humans construct their preferences leads to the second issue,

namely that the method or process employed to extract preference information from the user influences the preferences the user constructs. Preference construction can be influenced by the decision context, the framing of the choice task and the way relevant information is presented. This has to do with psychological effects (Fischer et al, 1999; Johnson et al, 2005), e.g. loss aversion (the tendency of people to prefer avoiding losses to acquiring gains) or anchoring effects (relying too heavily on one piece of information), the emotions induced or earlier experiences retrieved from memory when the elicitation question is posed (Weber and Johnson, 2006). The fact that the process of eliciting preferences (e.g. giving information about choices and asking a number of valuation questions) influences the construction of preferences should be taken into account when designing a preference elicitation interface by actively supporting this process. Active involvement in the construction process and a user interface design leading to a positive user experience (Knijnenburg et al, 2012) during the elicitation as well as an understanding and trust in the system's output later on (Carenini and Poole, 2002) is important for the success of recommender systems.

Our work focuses on informing the design of preference elicitation interfaces from a user-centered point of view. In this paper we present three studies that explore how we can bridge the gap between users' mental models and a system's representation of the preferences and how the constructive nature of human preferences can be supported in an interface. We combined experimental as well as qualitative research involving users in the design process to be able to create a number of design guidelines for such interfaces.

In the first experiment we investigated input methods and elicitation process in a structured way. We presented users with different ways of entering preferences, including ratings (Likert scale rating), affective feedback, and sorting, both on an item (i.e., a complete holiday) as well as an attribute (i.e., beach, mountain, active, etc.) basis. Based on the results of this experiment we hypothesized that users are willing to spend more effort if the feedback mechanism (i.e. process and preference representation) enables them to be more expressive (e.g. by giving more dimensional feedback or navigating through the outcome space). We examined this hypothesis in two follow-up studies. In the second experiment we explored the trade-off between giving detailed preference feedback and effort. We investigated factors, such as content type, familiarity, ownership and directed opinion (positive or negative), that may influence this trade-off in an experimental setup. In a third study we explored how people prefer the preference elicitation process to be structured using hi-fi interface prototypes and a participatory design method. We looked at four fundamentally different processes of eliciting preferences based on different ways to process information. We used the

mind style theory by Gregorc (2006) which categorizes people based on perceptual and ordering preference. Perceiving information can be abstract (based reason and intuition) or concrete (using one's senses). The order of information processing can be sequential or random. Thus there are four types to process information: concrete sequential, concrete random, abstract sequential and abstract random. We built one interface prototype per style and evaluated the prototypes in individual user sessions followed by a creative design session with all users. Based on the results of all three experiments we constructed a number of design guidelines to support further development of preference elicitation interfaces.

The experiments will be discussed in sections 3 to 5. Section 6 discusses the results and presents the guidelines and section 7 concludes the paper. But first, we will give an extensive background on how people construct their preference, how current systems elicit them, and how well the theory matches to the practice.

4.2 Background

People's preferences have been the interest of researchers in many different fields including psychology, (behavioral) decision making, consumer research, e-commerce, intelligent systems as well as negotiation and decision support. We focus on topics relevant for designing user interfaces for preference elicitation for intelligent systems. In the following sections we give insights into (1) how people construct their preferences (the process we need to support with preference elicitation interfaces), (2) the state-of-the-art in preference elicitation interfaces and (3) how the latter take the human preference construction into account.

4.2.1 Constructive Preferences

A dominant model in contemporary economy is that of the rational consumer trying to always maximize his outcome. Preferences are seen as primitive, consistent and stable (McFadden, 1999). It also assumes that people know their preferences. Since the rational consumer tries to maximize the value outcome it is implied that he is able to compute the maximal outcome based on his preferences and make a rational choice. This computation can be represented in utility functions — a mathematical representation of a person's preferences.

Whereas these assumptions serve rational economic theories well, they are not always true for human behavior. More and more researchers gathered proof supporting a constructive view of human preferences. This view implies that people construct their

expressions of preferences at the time the valuation question is asked. Furthermore, the decision process itself and the context play a major role in the construction process (Payne et al, 1999).

There are different views on how people construct their preferences. (Simon et al, 2004) for instance found in their experiments that while people processed the decision task, their preferences of attributes in the option that was chosen increased and those for attributes of rejected options decreased. This is in line with achieving the meta goal of trying to maximize the ease of justifying a decision (Bettman et al, 1998). Similar effects have been found in negotiation settings reported by (Curhan et al, 2004).

(Fischer et al, 1999) focused on the goals of the decision task in relation to a prominence effect. This effect occurs when people prefer an alternative that is superior only on the most prominent, i.e. the most important, attribute. They confirmed in three studies that the prominent attribute will be more heavily weighted when the goal was making a choice between alternatives than when the goal was to arrive at a matching value.

(Weber and Johnson, 2006) state that people construct preferences from memory. The so-called PAM (preferences-as-memory) framework assumes that "decisions (or valuation judgments) are made by retrieving relevant knowledge (attitudes, attributes, previous preferences, episodes, or events) from memory in order to determine the best (or a good) action." Weber and Johnson emphasize that this is not an entirely cognitive view on preference construction since affect determines what the person recalls first. Information consistent with emotions is more available in memory. Johnson and colleagues (Johnson et al, 2005) found psychological effects, such as anchoring effects and effects occurring when complicated numbers or information are presented in the choice task. In their experiments different ways to measure preferences led to different results. To help people to construct their preferences in health care scenarios, Johnson and colleagues suggest to present default choices that lead to the best outcome for most patients and present information in a way that helps the patient to understand the outcomes of each choice. Consumer research looked at the interplay between affect and cognition on decision making (Shiv and Fedorikhin, 1999). In cases where people have only few cognitive resources available affective reactions tend to have a greater impact on choice, whereas with high availability of cognitive resources thoughts related to the consequences of the choice are more dominant. This finding can be influenced by personality and by the representation of the choice alternatives.

In summary, there is consensus in the literature discussed that people do not always have stable and consistent preferences but rather construct them when necessary. There are numerous views on how people might construct their preferences. An easy, ready-to-implement recipe for designing interfaces for this task has not yet been established. There have been few attempts to guide system developers in this difficult task. (Carenini and Poole, 2002) point to the problems of clustering and matching algorithms in relation to the constructive process humans go through. Since users may not have the chance to construct their preferences they might also not be able to understand the system's output. (Kramer, 2007) also found that consumers are more likely to choose a recommendation that matches their measured preferences when it is easy to see through the preference elicitation method and by that identify their expressed preferences. The gap between the user's mental model and the system's preference profile of the user can be bridged with explanations from the system (Carenini and Poole, 2002). More research is needed to design preference elicitation interfaces that elicit correct preference information from the user. In the following sections we will give an overview over current preference elicitation methods used in state-of-the-art interfaces.

4.2.2 Preference Elicitation Methods and Interfaces

Methods for acquiring user preferences range from implicit to explicit ones depending on the nature of the system. By implicit we refer to approaches in which the user is not "actively" involved in the elicitation task as it is the case in explicit. One example of implicit methods can be preference learning based on user behavior (e.g. items the user looked at or bought). Users may still be aware of the workings of implicit methods and expect an interpretation of their actions. By explicit we, however, refer to methods that require specific preference input such as ratings. The range from implicit to explicit is continuous, meaning that various degrees of user involvement are used in the methods explained below. In the following sections we will describe the methods typically used in recommender systems and decision support together with examples representing typical systems in the area. For more exhaustive reviews in the area see (Chen and Pu, 2004; Peintner et al, 2008).

Methods used in Recommender Systems

Recommender Systems (Adomavicius and Tuzhilin, 2005) are tools that provide personalized recommendations to people. They are integrated either in shopping websites, e.g. amazon.com or dedicated recommendation websites (Resnick et al, 1994;

Burke, 2000; Miller et al, 2003; Stolze and Ströbel, 2003). The interaction models employed to acquire preferences vary from implicit to explicit methods. Typically, preference elicitation is done through item-rating (and using filtering methods) or more conversational interaction using tweaks or critiques.

Rating-based recommender systems collect a number of initial ratings from a user and then try to estimate ratings for the yet unrated items. Based on a user's profile and estimations for unseen items they can recommend new products to the user that he or she could be interested in. Two methods are mainly used, collaborative filtering (Herlocker et al, 2004) based on similarities between users (as reflected by their ratings) and the content-based method (Pazzani and Billsus, 2007) based on item attributes instead. Most content-based recommenders employ machine-learning techniques to create a user profile. To recommend items the attributes of the item are compared to the user's profile to see which items would be of interest to the user. Systems using collaborative filtering are, for instance, MovieLens (Miller et al, 2003), or GroupLens (Resnick et al, 1994) (www.grouplens.org), but also commercial systems like Amazon.com; an example of a system using a content-based method is the book recommender system developed by (Mooney and Roy, 2000).

Recommender systems using collaborative filtering or the content-based method mostly focus on getting a substantial number of ratings from their users when they sign up. During use the explicit interaction between system and user is limited to recommendations from the system and voluntary ratings from the user. (Carenini et al, 2003) have instead proposed a more conversational and collaborative interaction. Key in the proposed interaction is that the system tries to elicit ratings or preferences when people are particularly motivated to give them, e.g. when the system cannot give a requested recommendation due to a lack of preference information or when the given rating puzzles the user. Methods developed based on this conversational model are 'Example Similarity and Tweaking' and 'Example-Critiquing Interaction'.

Most so-called FindMe systems (e.g. Car Navigator, PickAFlick, RentMe or Entrée (Burke et al, 1996; Burke, 2000, 2002)) employ example similarity and tweaking techniques. In the first step the user selects an item from the system's catalog and requests similar items. The system then retrieves a large number of alternative items from its database, sorts them according to similarity to the chosen item and returns a small number of alternatives with highest similarity to the user. In case the system offers tweaking the process is essentially the same with the only difference that the user gives a tweak in the first step, e.g. "show me similar, but cheaper items". The system only returns items to the user that satisfy the tweak, that are cheaper in this

case. A similar technique is example-critiquing or the candidate/critique model (Pu and Chen, 2008). In example-critiquing users are presented with a set of candidates they can critique. Candidates have to motivate users to state their preferences and the most preferred solution needs to be among the displayed candidates (Faltings et al, 2004). Several strategies have been proposed to select candidates, e.g. using extreme examples (Linden et al, 1997), diverse examples (Smyth and Mcginty, 2003) or so called Pareto-strategies (Viappiani et al, 2005). In an iterative process the system learns the users' preferences from their critiques and updates the user models. Critiques can be system-suggested or user-initiated. Work to enhance the critiquing has been done in the area of dynamic critiquing, in which compound critiques (critiques operating over multiple features) are generated on-the-fly (McCarthy et al, 2005). One of the first interfaces using example critiquing is the APT Decision Agent (Shearin and Lieberman, 2001). For details of other systems and algorithms in this area and specific design guidelines on how to develop example critiquing interaction, see (Chen and Pu, 2009).

Decision Support Systems

Decision Support Systems (DSS) are interactive systems that support users in taking decisions by eliciting preferences and offering analytical tools to scrutinize decisions. Unlike recommender systems that focus on finding the best outcome in a huge set of possible outcomes, decision support systems focus much more on the process of taking a decision and the role of preferences influencing that process.

Preferences are elicited explicitly because it is important that the user understands the relation between his or her preferences and the possible outcomes of the decision making process. Decisions that are supported by these systems are often of much higher importance than choosing to buy a book or see a movie, e.g., medical systems (Hunt et al, 1998; Johnson et al, 2005). It is, therefore, important to have a precise model of the users' preferences.

The majority of decision support systems are based on multi-attribute utility theory (MAUT) (Keeney and Raiffa, 1993) and, therefore, represent preferences in form of utility functions. In order to construct utility functions the system needs to elicit values and weights for the given attributes of an item. The two most popular preference elicitation techniques are *absolute measurement* and *pairwise comparison* (active elicitation) (Aloysius et al, 2006). Absolute measurement (e.g. salary scores 9 on a scale 1–10 of importance, whereas number of holidays score 5 out of 10) does not require the user to make explicit trade-off judgments. Pairwise comparison (e.g. salary

is more important than number of holidays) explicitly asks for a trade-off between attributes. Weights in most existing systems are entered by users on discrete scales by selecting a rating from a drop down list or using horizontally aligned radio buttons and on continuous scales by using a slider. Aloysius and colleagues (Aloysius et al, 2006) found an impact of the preference elicitation technique used on the user acceptance of DSS. Their study comparing absolute measurement and pairwise comparison showed that forcing the user to make explicit trade-off judgments has a negative effect on user acceptance of the system. Note, that this does not mean that the decision outcome will be worse. However, due to higher perceived effort and decisional conflict the user perceives the accuracy of the system to be lower.

Configuration Systems

Similar to recommender or decision support systems are configuration systems, which support the configuration of complex products and services. A growing demand for customer individual, configurable products also asks for improvements of configuration systems that usually have to deal with a wide variety of users. In this area Ardissono and colleagues (Ardissono et al, 2003) have developed the CAWICOMS workbench to develop configuration services. The way this workbench manages user models and personalizes the interaction between user and system by customizing the acquisition of requirements and information presentation is interesting. The system exploits user classes based on stereotypes that specify skills as well as interests. In the beginning of the interaction the system asks explicitly about background information of the user. This helps to define the user class and make estimates about the user's interests and skills based on the stereotypes. During the interaction the system observes the actions of the user to update the skills and interests continuously as they may evolve. The reasoning of the system is based on the rational assumption that the user always tries to maximize her own utility by setting item features to satisfy her needs. Depending on the system's assumptions it selects certain features as critical and others as less important which are presented as supplementary information.

4.2.3 Support of human preference construction in current methods

In order to build a system and in particular the interface of the system that is usable and supports the user in creating and entering his or her preferences system designers need to consider how human preferences develop. In detail, this means they have to support users in constructing their preferences in a cognitive as well as affective way and maybe look at underlying interests (or values) as a basis for selecting the right

attributes. But in how far do the methods presented above actually take these aspects into consideration?

Generally, the implicit methods do not actively involve users in constructing their preferences. Collaborative filtering methods base their choices on the assumptions that similar people like similar things. However, there is always the danger that the system creates an erroneous user model and the user gets confused about seemingly unrelated recommendations.

Explicit methods focus more on the user. Conversational methods allow the users to construct and reflect upon preferences. Example critiquing has been explicitly developed based on the constructive view of human preferences. This is reflected in the attempts to improve the algorithms to pick examples that will help the users to uncover hidden preferences. Whereas tweaking and example-critiquing are widely used for Recommenders, many decision support systems use active elicitation methods based on utility models. This requires users to enter values and weights in form of numbers. Other active elicitation methods like pairwise comparison do not require numbers but still assume that a user is able to compute which of the given options is better. With unknown items this is a difficult task.

A combination of explicit questions, in particular in the beginning of the interaction, to place the user in a certain class and continuous explicit updates of the preference model based on user behavior (Ardissono et al, 2003), can help to create an accurate preference model. Continuous updates of the model and adaptations of the interface support the constructive nature of human preferences as they may evolve during the interaction with the system.

Surprisingly few systems explicitly try to elicit underlying interests before deciding which attributes or items are worth looking at for preference elicitation, notable exceptions are (Fano and Kurth, 2003; Stolze and Ströbel, 2003). Affect is also under-explored in current preference elicitation methods. Only the movie recommenders by (Ono et al, 2007) and whattorent.com ask for input about emotions or moods. However, underlying interests and affect are important aspects of human preference construction that should be considered in a preference elicitation interface.

In summary, we can say that there are a few methods that are explicitly based on the constructive view of human preferences. There is much room for more explicit consideration of human preference construction also including values and affective aspects.

4.2.4 Related Work

Whereas most work described in the previous subsections focuses on different techniques in which a system can either explicitly or implicitly arrive at a preference model, our work focuses strongly on the design of preference elicitation interfaces, in particular for explicit preference elicitation. Naturally, the interface design is also determined by the interaction style or technique that is chosen. However, besides that, we believe that there is much room for improvement and greater support of the way in which humans construct their preferences. This constructive view has been acknowledged by others who also established a number of guidelines for preference elicitation (Payne et al, 1999; Pu et al, 2003; Pu and Chen, 2008; Pu et al, 2012). Our work is intended to build on this work by extending the number of guidelines in order to help other designers. However, our work differs as it focuses on a number of issues neglected in the current literature. These include: (1) an in-depth investigation of interface elements considered appropriate and preferred by people for entering preferences, (2) the match of the acquired input with input required by algorithms currently used to create a system representation of the preferences, (3) intrinsic motivational factors that lead people to spend more effort to give more detail about a preference and (4) ways to structure the process of preference construction by the design of an interface based on different information processing styles. Investigating these aspects required close interaction with target users. In our view, when designing new preference elicitation interfaces a participatory design process will lead to a greater understanding of interface aspects that would not be acquired with user evaluations of finished prototypes. The majority of the work presented above does not follow such an approach. Most closely related to our work is the work by Barneveld and Setten (Barneveld and Setten, 2004), who also involved users actively in the design phase of a TV recommender system with the means of brainstorming and interactive design sessions. Furthermore, they also investigated which interface widgets would be preferred by users to give preference input. Our work differs from theirs as we looked more at different types of input (rating, ordering, affective, navigation in the first study and higher level elements, e.g. a chat, in the third). In addition, whereas they focused on a design for one particular domain (TV) our work aims at a general understanding of how to design preference elicitation for different types of systems and different user groups. In the following sections we will describe our studies in detail.

4.3 Study 1: Investigating different ways of entering preferences

In this experiment we compared different ways of giving preference input (ranking, ordering, navigational) regarding perceived liking and effort and how well the extracted information serves as input for an outcome ranking algorithm. We agree with conclusions of (Knijnenburg et al, 2012), that the algorithms cannot be studied in isolation with end-users, but have to be investigated together with the preference input to fully understand the complete user experience. Therefore, these two aspects are combined in this study. Liking is only one aspect of the user experience. We decided to investigate effort as most preference elicitation tasks require some level of effort from the users even before they can actually judge the usefulness and efficiency of a system (e.g. the accuracy of recommendations provided after the initial preference elicitation)(see also (Pu et al, 2012)).

In this study, we considered ordering and rating tasks on both a property and outcome level. To investigate the effect of affective input we compared standard Likert-scale rating to affective ratings using the AffectButton (Broekens and Brinkman, 2009). This button (Figure 4.1) enables users to enter dynamic (i.e. graded) emotions. It renders a face that changes directly according to the mouse position and scroll wheel. The mouse-coordinates inside the button and the scroll wheel together define the values on the affective dimensions Pleasure, Dominance and Arousal (PAD) (Mehrabian, 1980) respectively. All three dimensions are represented by values on a scale from -1 to 1 (e.g. -1 displeasure to 1 pleasure and accordingly). The pleasure dimension indicates how pleasurable an emotion is, e.g. fear or anger are emotions that are not pleasant whereas joy or contentness are pleasant. Dominance indicates the nature of the emotion ranging from submissive (e.g. in fear) to dominant (e.g. in anger). The arousal dimension indicates the intensity of an emotion ranging from low to high. Whereas joy has a high intensity, contentness has a low intensity. By using the AffectButton the users select an affective triplet from the PAD space (as reflected by the emotional expression of the button itself; the PAD concept is not visible to the user).

Furthermore, we compared a *navigational input method*, inspired by guidelines proposed by (Pu et al, 2003) (i.e. any preference in any order and immediate visual feedback) to traditional ordering of properties. In the navigational input method users navigate through the outcome space by changing any one property at a time and receiving visual feedback for the new choice.

Figure 4.1: Example expressions: from left to right Happy (PAD=1,1,1), Afraid (-1,1,-1), Surprised (1,1,-1), Sad (PAD=-1,-1,-1), Angry (-1,1,1)

To see how the interaction between the input method and the system's computation influences the end result (ranked list of items) we used the preferences over properties obtained from different methods as input for the lexicographic ordering. The lexicographic ordering was chosen as it does not require numerical input from the participants (properties need to be ordered, but are not associated with a numerical weight) and by that allowed us to use ordering tasks in the experiment. It has also been argued that it is a natural and intuitive way to derive preferences over objects from an importance ranking of properties (Liu, 2008). This type of ordering compares two items according to the property that is rated most important. Other properties will only be considered if the value of the most important property is the same for both objects. So given a user prefers having a garage to a garden with his house, then an option A that has a garage is always better than an option B without a garage, even in cases where option B has many other attributes that the user also likes but finds less important than a garage. If option A and B contain a garage the algorithm will compare the options based on the next important attribute, e.g. whether they have a garden and so on.

4.3.1 Research Questions

Overall, we addressed three topics: (a) different preference input methods (interface), (b) the navigational input method and affective inputs and (c) the outcome ordering using a lexicographic algorithm with input from the property rating/ordering methods. In detail, we focused on the following research questions.

(*1*) *How do people perceive the different input methods in terms of liking and effort?*

(2) *Do users prefer the navigational input method to standard ordering and rating methods in terms of effort, intuitiveness, ease of use and liking? Can the navigational input method extract the same information as the property ordering method?*

Table 4.1: Overview of 8 preference elicitation tasks.

Task	Description
1A	Order 9 property values (given at the same time)
1B	Order 27 holidays
2A	Navigation through holidays
2B	Order 3x3 property values (given three at a time)
3A	Likert rating of holidays
3B	Affective rating of holidays
3C	Likert rating of properties
3D	Affective rating of properties

(*3*) *Do users prefer to give affective feedback? How does the user perceive the quality of the resulting outcome orderings?*

(*4*) *How similar are outcome lists generated with the lexicographic ordering to a list created by the user (baseline)?*

4.3.2 Study Setup

We ran an experiment consisting of 8 ordering/rating tasks (tasks will be numbered throughout the paper), 2 comparisons of results and a final questionnaire. An overview of the ordering/rating tasks is presented in Table 4.1 (each task will be discussed in more detail below). After execution of a task we asked participants to rate (on a 7-point Likert scale) how much effort the task cost and how much they liked the task. We chose holidays as our domain, since people can easily relate to holidays and have preferences about different aspects of holidays. Each holiday has the properties type, location and accommodation, with the respective alternative values relaxation, active and city trip, Mediterranean, Scandinavia and Alps, and hotel, camping and apartment (Table 4.2).

Material

The study material consisted of two sets of 9 cards, each showing one alternative value for a holiday property, one set with pictures (Figure 4.2b) and the other without pictures. Further, there were two sets with 27 cards showing complete holidays; one set with 4 pictures to give an orientation about what the holiday could look like (Figure 4.2c), and one set without pictures. Furthermore, we provided a computer interface

Table 4.2: Properties of holidays and the alternative values for each property used in the experiments.

Location	Accommodation	Type
Mediterranean	Apartment	Relaxation
Alps	Hotel	City trip
Scandinavia	Camping	Active

for participants to rate either holidays or alternatives for properties of holidays one at a time. Rating was done using either a 9-point Likert scale from like to dislike or the AffectButton.

Participants

We tested 32 participants, 10 female and 22 male, who were mainly students and researchers within the field of information technology aged between 21 and 31. Each participant had to do all tasks the experiment consisted of. The order of the tasks was counterbalanced to avoid carry-on effects. However, as the property space is kept small we expect people to know their preferences for the holiday preferences from the start or construct them easily. We do not see this as a problem as the focus of the study lies on different ways to enter a (possibly known) preference, not on constructing it.

Design

Effort and Liking of input methods After each input method we asked participants to fill in a short questionnaire rating how much they liked the method and how much effort it took them.

Standard input for lexicographic ordering and baseline In task 1A participants were asked to order all nine property values (see Table 4.2). This property ordering was later on used as input for the lexicographic ordering of holidays. Task 1B -ordering 27 cards showing complete holidays, each consisting of a combination of the three properties- was used as a baseline to compare holiday lists. Equally preferred holidays could be put on the same level. All cards had to be laid out on the table from most preferred to least preferred.

Navigational input method To test the effect of a navigation through the decision space, i.e. holidays to elicit preferences, two tasks were presented to the user. In the

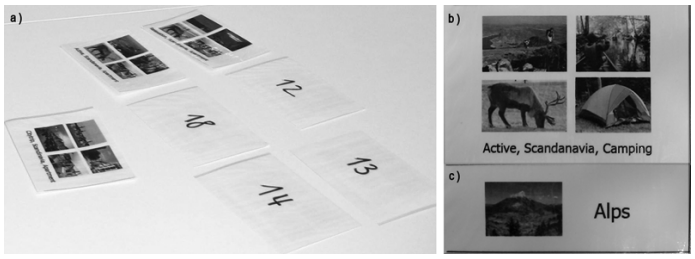


Figure 4.2: (A) Navigational task: Card in the third row presents current holiday, the participant can look at two other holidays at a time. (B) card representing a holiday. (C) card representing one property.

navigation task (2A), the participants were initially presented with a random card showing a complete holiday and asked to find their most preferred holiday by changing one property at a time to any of the two alternative values of that property, e.g. location could be changed from Scandinavia to either the Alps or the Mediterranean. As there were three properties (location, accommodation and type) that could be changed to two other values than the current, each holiday had six related holidays. The subjects could have a look at all six holidays (two at a time) related to the present one before deciding which one to navigate to (Figure 4.2a). The task was presented as a paper prototype. Once the subjects found their most preferred holiday the procedure was repeated for the least preferred holiday starting with the most preferred one. The cards showed three property values of a holiday and four pictures, which were used to give the participant an idea about the kind of holiday.

In the second task (2B), the subject had to order the alternative values of each of the three holiday properties (see table 4.2). Each property was presented on a card with a picture. Furthermore, the subject was asked to order the properties (type, location and accommodation) according to importance when searching for a holiday.

In addition to the effort and liking questionnaire, a questionnaire was presented to the user containing a number of questions about the intuitiveness and ease of use of the navigation (2A) and property ordering (2B) tasks.

Affective Feedback To study the effect of affective rating methods we used a 2x2 experimental setup. We had four different conditions: (1) 9-point Likert rating of nine holidays (3A), (2) affective rating of the same holidays (3B), (3) 9-point Likert rating of all nine property values (3C) and (4) affective rating of all nine properties

(3D). Holidays and properties were presented one by one and in random order. For each condition a simple algorithm generated an ordered list containing nine holidays based on the user input. In the first condition the list was ordered directly based on the user's holiday preference feedback. In the second condition feedback variables pleasure, arousal and dominance were summed and then used to order the list. In the third condition the weight of the property value entered by the user was used to calculate a sum for each holiday. This sum was used to order the list of holidays. In the fourth condition the pleasure, arousal, and dominance feedback was summed and then used to order the property values; from this property ordering an ordering of the nine holidays was derived. These algorithms resulted in four differently sorted lists, each containing the same holidays. After the rating and ordering tasks, users were asked to compare the four lists to their own holiday ordering.

Preference Ordering We used the information collected in tasks 1A, 2A, 3C and 3D (tasks based on holiday properties) as input for the lexicographic ordering to compute orderings of all 27 holidays (for details see (Pommeranz et al, 2008)). Besides an objective comparison, we asked participants to judge which list better reflected their preferences; the one they specified themselves in task 1B or the list generated with the lexicographic ordering method from the input from task 1A.

Procedure

The study was conducted during two weeks. Each experiment took about 45 minutes and consisted of eight tasks considering preference input, two comparisons of resulting lists and a final questionnaire. The ordering tasks and the navigational task was carried out using cards whereas for the rating tasks we used a computer interface. To compute the resulting lists with ordered holidays using the lexicographic algorithm one of the two present researchers entered the data from the ordering tasks into a computer program. This included the ordering of 27 holidays which we used to compute objective measures of proximity between the different lists. Before the tasks were explained and executed a general introduction was given about the goal of the experiment and the holiday domain. Furthermore, subjects were told that each task stands for itself, which means there is no need to remember anything between the tasks. The presentation of tasks to users was counter-balanced to avoid order of presentation effects.

4.3.3 Results and Discussion

Effort and Liking of Input Methods

After each task participants were asked to rate how much they liked it and how much effort it took them. A Multivariate analysis of variance (MANOVA) [1] with repeated measures was conducted to examine an effect for the ordering/rating style (independent within-subject variable) on the perceived effort and liking (dependent variables). We found a significant main effect for ordering/rating style ($F(14,18) = 10.71$; $p < 0.001$)[2], which was found again in the univariate analysis of the effort rating ($F(7, 217) = 27.91$; $p < 0.001$), and the liking rating ($F(7, 217) = 3.17$; $p = 0.003$). As expected, figure 4.3 shows that task 1B (ordering all 27 cards) clearly stands out as least preferred and highest in effort. Figure 4.3 also shows that more traditional individual property ordering (2B) or rating (3C) tasks were rated low on effort and relatively high on liking. This suggests that people appreciate the relative cognitive simplicity of these tasks; dealing only with a small part of the outcome space complexity. From the tasks that involved evaluating the complete holidays (1B, 2A, 3A, and 3B) it seems that the navigational input method (2A) is most preferred. Considering rating tasks it is interesting to notice that both tasks involving affective feedback are scored equally high in liking as Likert-scale ratings, although affective input is considerably more effortful.

Navigational Input Method

Besides considering liking and effort we also compared the navigational input method to the ordering of alternatives of holiday properties in terms of intuitiveness and ease of use. With a MANOVA with repeated measures (various ratings as dependent measures, and the task as independent within-subject variable) we found a significant main effect ($F(4,28) = 3.14$; $p. = 0.030$) for task, which was only found again in univariate analysis on effort ($F(1,31) = 9.02$; $p. = 0.005$) and intuitiveness rating ($F(1,31) = 4.64$, $p. = 0.039$). Examining the means shows that participants rated the navigational input method ($M = 3.0$, $SD = 1.65$) more effortful than ordering the property alternatives ($M = 2.0$, $SD = 1.16$) and less intuitive ($M = 4.9$, $SD = 1.48$) than the ordering ($M = 5.6$, $SD = 1.32$). This suggests that the more traditional ordering method is preferred.

[1] www.statsoft.com/textbook/anova-manova/

[2] F stands for F-statistic, p indicates significance

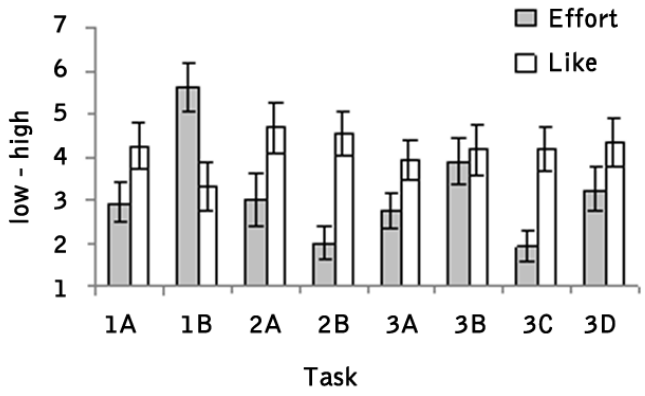


Figure 4.3: The mean liking and effort rating of ordering/rating tasks, including a 95% confidence interval

Studying the tasks in more detail revealed the navigational input method was the only method that enabled participants to enter dependencies between the alternatives of the holiday properties. For a considerable group of the participants (34%) the most and least preferred holiday had at least one equal value. One participant even had two equal values. This means two things. First, a property independent approach is not suitable for all people to describe their preferences. Second, the navigational input method might be an effective approach to determine whether for a specific individual preferences over properties are dependent.

Affective Rating

We analyzed the effect of affective rating using a MANOVA with repeated measures. It showed a main effect of affect versus Likert scale rating ($F(2,30) = 24.00$; $p. < 0.001$) and property versus whole holiday rating ($F(2,30) = 6.73$; $p. = 0.004$) with no significant interaction effect. These main effects were found again in the univariate analysis on effort for affect versus Likert scale rating ($F(1,31) = 46.32$; $p. < 0.001$) as well as for property versus holiday rating ($F(1,31) = 13.90$; $p. = 0.001$). This means that both affective-, as well as holiday-based feedback are associated with a higher perceived effort in preference elicitation (Table 4.3). With regards to the perceived quality of the resulting lists generated by the simple algorithms we found a significant main effect for affect versus Likert scale rating ($F(1,31) = 6.12$; $p = 0.019$).

Table 4.3: Summary of mean liking and effort scores for the tasks 3A-3D and generated lists.

Condition	Liking	Effort	Quality of outcome list
Likert & Holiday	M=3.938 SD=1.318	M=2.750 SD=1.191	M=6.188 SD=1.786
Affect & Holiday	M=4.188 SD=1.575	M=3.906 SD=1.594	M=5.500 SD=2.064
Likert & Property	M=4.188 SD=1.731	M=1.938 SD=1.014	M=6.031 SD=2.177

This suggests that the algorithmically-generated lists based on affective feedback matched the user's preferences less well than the lists that were generated based on Likert-scale feedback (see column 4 in table 4.3). This can be due to two reasons, either the participants did not understand the semantics of the AffectButton well and by that could not express their preferences correctly or the algorithm used to calculate the outcome lists did not work well given the input variables (pleasure, arousal, dominance). We exclude the first reason as previous research suggested that the AffectButton is a valid and reliable affective feedback device when used for rating the affective content of emotion words (Broekens and Brinkman, 2009) as well as film music (Broekens et al, 2010b).

A deeper analysis suggested that our way of mapping affective dimensions to algorithms that are intended for one dimensional preference values was too simplistic. To understand which factors are most important in predicting the holiday list created by the user (1B) we did a regression analysis given the Likert rating and pleasure, arousal, dominance ratings (stepwise) over all holidays. The same analysis was repeated for the property values (now predicting the property ranking of task 1A). The regression analysis predicting holiday-ranking resulted in a significant model (r=0.66; $F(2,285) = 110$; $p < 0.001$). The model included the Likert rating (β = -0.55; t = -9; $p < 0.001$) and pleasure rating (β = -0.15; t = -2.5; $p = 0.012$) as significant items. The regression analysis predicting property-ranking showed similar results, but included Likert rating and dominance as significant items. In both cases at least one affective factor was included in the predictive model. This suggests that affective feedback helps the user to express preferences.

Preference Ordering

We used the different methods of rating and ordering properties (see section 3.2.3) as input for the lexicographic ordering algorithm to investigate how well this algorithm can perform given a variety of inputs. These methods include affective rating (D3), 9-points rating (C3), ordering 9 property values (A1), ordering the properties and then 3x3 values (B2). The algorithm generated ordered lists for each user, and these lists were compared with the lists that the users specified themselves in the 27-card ordering task (B1). This is essentially a comparison between two rank-ordered lists containing the same items. The similarity between these lists is computed in two ways. Kendall's τ can be seen as a distance measure; it is based on the minimal number of switches between two adjacent items in one list that is needed to attain the second list. Spearman's ρ is another well-known rank correlation method. Both measures are normalized and range from -1 to 1, where 1 indicates that the lists are identical, 0 no relation at all, and -1 indicates reverts ordering. Figure 4.4 shows the correlation coefficients averaged over participants between the standard list (specified by the participant in task B1) and the lists generated with the lexicographic ordering method with different types of user input. All correlations are significant ($p < 0.001$), which indicates that the generated lists are much more similar to the standard list than random lists. This suggests that different input method combined with lexicographic ordering can result in "true" preference orderings. For more details on the analysis, please see (Pommeranz et al, 2008). As we cannot guarantee that the user-specified list (ranking of 27 holidays) is ideal, given that the task was tedious and little appreciated by the users, it is hard to say how close each generated list came to an ideal list of a person's preferences. Interesting to note, however, is a clear difference between the lists generated from affective feedback and non-affective feedback, whereas the lists generated with affective feedback are less similar to the user-generated lists. We believe this is due to a difficulty of translating the 3-dimensional affect feedback into a one-dimensional ranking, as explained in the previous subsection.

Summary of Results

The results showed three important aspects relevant for understanding the process of preference elicitation and the match between the user's mental representations and the system's model. First, the results confirmed that cognitively less demanding ordering or rating tasks were perceived as less effortful and liked most by users. So, liking and effort go hand in hand (see similar results of FT2 trial in (Knijnenburg et al, 2012)). Second, navigation through the outcome space (moving from item to item by

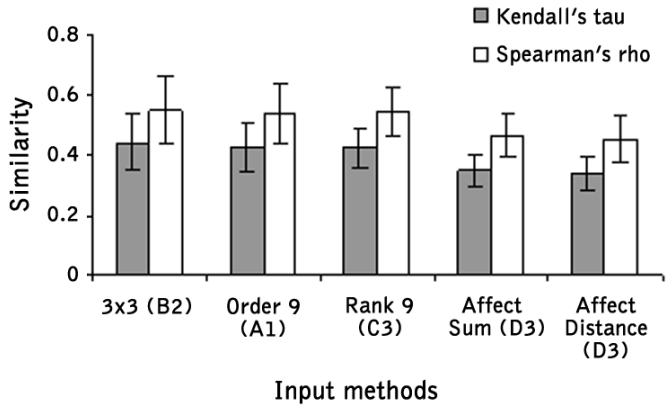


Figure 4.4: Similarity of lists generated from different input methods to standard list.

changing an attribute value at a time) enables users to express dependencies between attributes that were not revealed by other methods, and, affective feedback enables users to express preferences in other dimensions additional to liking. Third, effort is not an indicator of how much a method will be liked in these last two cases. Affective feedback and navigation were rated significantly higher in effort than other methods, but still high in liking. We hypothesize that this indicates that users are willing to spend more effort if the feedback mechanism (process and preference representation) enables them to be more expressive (or maybe more entertaining as mentioned by (Pu et al, 2012), which is good because it enables the system to extract more preference information and by that build a more accurate user model.

4.4 Study 2: Testing user motivation to give preference detail

To test the above mentioned hypothesis we investigated the trade-off between giving detailed preference feedback and effort. We examined factors (e.g. familiarity - also mentioned considering recommendations by (Sinha and Swearingen, 2002; Pu et al, 2012)) that can influence this trade-off in an experimental set-up. The focus of this study was on investigating the influential factors in a neutral set-up, i.e. without other motivational factors that could be present in a recommender set-up, such as giving detailed feedback to receive better recommendations or to serve the community

of users. Therefore, we chose for a simple content rating task. In order to make sure we did not introduce a bias in ratings by telling people that we investigate the level of detail people give, we instructed participants that the experiment was about creating an alternative top-40 list of famous people and popular music. While this is an incentive to take part in the study in general, there was no incentive to give more detailed preference feedback.

The study consisted of two follow-up online experiments. Since the second experiment was an enhanced version of the first one, we will only elaborate on the second experiment here. To see the details and results of the first experiment please refer to (Broekens et al, 2010a). The main improvements we made in the study are changes to the interface including using more familiar rating mechanisms such as thumbs and stars, an option to replay a song and the omission of given emotional tags as an input level. Furthermore, we gave a more detailed explanation of the interface (screenshot seen in Figure 4.6), let people try all levels before starting the experiment and always showed the following input level during the interaction. By this we reduced problems with the ratings that may have occurred in the first study due to misunderstanding the interface.

4.4.1 Research Hypothesis

Our hypothesis for the following experiment was that *the level of detail persons are willing to give in their feedback depends on content type of the item, familiarity with the item, ownership of the item and opinion about the item.*

4.4.2 Study Setup

We set up a content-rating experiment with content type, familiarity, ownership and opinion as factors and detail as dependent variable. The content was preselected by the experimenters. We purposefully chose two types of content, music and pictures of famous people, each allowing for different ways to form a preference. We hypothesized that people would be able to form preferences for music spontaneously and give detailed feedback (also in form of affect feedback) while listening to the music even if the song was unknown to them. In the case of famous people we did not expect this behavior as a picture alone does not allow for getting to know the person better and by that would lead to a formation of a directed (positive/negative) opinion. The other experimental factors familiarity, ownership and opinion for each content item were indicated by the subjects during the study.

Figure 4.5: AffectButton: the cross indicates the position of the mouse cursor inside the button, the face changes accordingly.

Material and Procedure

Participants received an email with the invitation to participate including a link to the application needed for the study. The study was done online, so subjects did the experiment at a place and time of their own choice. The email contained detailed instructions about how to use the application including a screenshot of the interface (see Figure 4.6). After the participants started the application, they were asked to fill in demographic information (age, gender, and education). Then they were presented with an example picture (Donald Duck) that had to be rated using all levels of detail to ensure that all participants were familiar with the interface. After that, the application presented 30 songs and 30 pictures of famous people (one at the time, at random). The pictures were labeled with the famous person's name. Songs were presented as audio samples without an indication of the title or artist name. They could be played as often as the participant wanted. For each picture/song they were asked to fill in their familiarity with the song or person (6-point scale: 0 and 1 was interpreted as not knowing the item and 2–5 was interpreted as knowing the item) and whether or not they owned the song or media concerning the person (yes/no) (see right side of the window in Figure 4.6). Then they were asked to give their opinion about the picture/song. The four levels of feedback detail were:

(*1*) *Thumbs-down/neutral/thumbs-up.* All subjects had to rate their opinion about each item using this input level. This is the minimum level of detail that can be given on one dimension (liking) including a neutral position.

(*2*) *A 6-point scale (represented by 6 stars, one star being the minimum).* This is the usual form of giving more detailed feedback, as used on many websites. It introduces the possibility to give a higher resolution of detail but still on one dimension (liking).

(*3*) *Affective feedback* using the AffectButton (see Figure 4.4) an interactive button that can be used to give affective (emotional) feedback based on three dimensions: pleasure, arousal and dominance. It is a dynamically changing selectable

emotion expression. This introduces the possibility to give fine grained feedback on 2 extra dimensions (arousal and dominance) in addition to the liking dimension.

(*4*) *Free text input.* This option enables subjects to tag the item. We assume this to be the most fine-grained and high dimensional kind of feedback, as essentially users can use any tag they want. People were instructed to use any words that express their opinion about the item.

For each stimulus they had to give at least a thumbs-down/ neutral/ thumbs-up opinion (3-point scale) rating. Neutral was interpreted as no opinion, thumbs-down and thumbs-up were interpreted as having an opinion. After that they had the choice to enter more detail to their opinion or go to the next picture/song. There were 4 levels of detail and each level had to be filled in before the participant could go to the next to make sure the user takes an active decision in whether to give more feedback or not. The user could always see the following level of detail. At every level, subjects could stop giving feedback and go to the next stimulus, except at the obligatory first level.

Participants

A broad range of people, in total 41, participated in the online experiment of which 13 female and 28 male, aged between 11 and 58 (M=31, SD=10). Participants have different cultural backgrounds as well as nationalities (including Dutch, German, Swedish, and Chinese) and education level (education level ranged between high school (with an exception of children aged 11 and 13) and post master level, Median=Bachelor).

4.4.3 Results

Before analyzing the data in detail we checked for any effects of the experimental setup. First, we found that items rated in the last half had an average level of detail equal to 1.8, while in the first half this was equal to 1.9. This indicates that participants gave less feedback later in the experiment, which can be attributed to the time it took (30 min) to rate the 60 stimuli. However, as the effect is rather small, this poses no problems for interpreting those items rated later. Second, we found a healthy distribution of thumbs-based feedback about items (26%, 40% and 34% of the cases were rated as bad, neutral or good respectively). The fact that 40% were rated as neutral and 60% with a positive or negative opinion allowed us to use opinion as factor in the analysis. Third, we found positive correlations (all correlations significant, $r >$

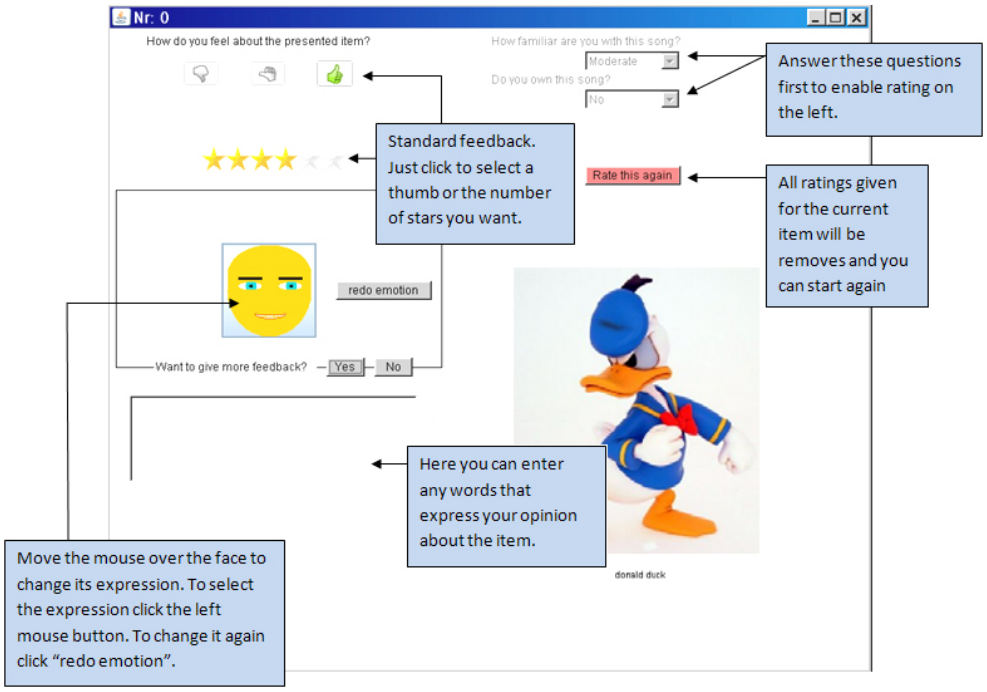


Figure 4.6: Instructions for the interface used for testing the level of detail people are willing to give.

0.7) between the ratings entered in levels 1–3 indicating that users were consistent when rating an item with different input methods (thumbs, stars or AffectButton).

In the further analysis we focused on main effects, as familiarity, ownership and opinion are not experimental controlled variables. Ratings were aggregated per subjectXfactor, averaging over the rated levels of detail, resulting in 41 paired measurements per main effect analysis. We interpreted familiarity ratings < 2 as unfamiliar and $>= 2$ as familiar. For the factor opinion we differentiated between directed opinions (thumbs-up and thumbs-down ratings) and neutral.

Our hypotheses were confirmed with respect to the influence of ownership, opinion and familiarity, and to a lesser extend the influence of content.

Subjects rated familiar items (M=2.12, SD=0.87) with more detail (paired $t(40)$=-5.19, $p<0.001$) than unfamiliar items (M=1.72, SD=0.75). Items that are owned are

rated (M=2.19, SD=0.88) with more detail (paired $t(37)$=-4.12, $p<0.001$) than items that are not owned (M=1.83, SD=0.78). These two effects might influence each other, owned items have a much higher chance of also being familiar. For these two factors we checked the interdependency using a 2x2 repeated measured ANOVA, and this showed indeed that when taking both factors into account, only ownership remained a significant factor ($F(1, 37)$=25.2, $p<0.001$), and familiarity did not ($F(1, 37)$=2.78, p=.104). The interaction effect was not significant ($F(1, 37)$=0.20, p=ns). When subjects had a positive or negative opinion (M=1.98, SD=0.79), they rated with more detail (paired $t(40)$=-5.77, $p<0.001$) than when they had no opinion (neutral) (M=1.60, SD=0.75)). Further analysis revealed that a positive opinion was related to rating with the highest amount of detail (M=2.10, SD=0.81), followed by negative opinion (M=1.85, SD=0.77), and no opinion having the lowest detail (M=1.57, SD=0.73). All differences were significant in paired t-tests at the level of $p<0.01$.

We did find a significant effect of content type (paired t(40)=-2.58, p=.014). However, the difference was small. Music (M=1.91, SD=0.73) was scored with only a little bit more detail than images (M=1.78, SD=0.80). This means that, although the effect of type of content was significant, the effect was relatively small compared to the effects of the other three factors (a difference in means of about 0.13 compared to around .40 for the other factors). The tendency to give detail seems to be a factor that should be explained from within the subject, an important finding in light of preference elicitation.

Finally, we show the distribution of the level of detail used to rate cases in Figure 4.7 and Figure 4.8. Each bar in Figure 4.7 represents the number of cases rated with a level of feedback (so, if a user stopped at level 3, he/she rated the item with level 1, 2 and 3; explaining why 100% of the cases was scored with at least level 1, as this was obligatory) split between having an opinion (positive or negative) or not having an opinion. Figure 4.7 shows an overall trend for using more feedback when a positive or negative opinion is present. Most notably, in about 40% of the cases where a positive or negative opinion is present, affective feedback (level 3) was used to express more detail. Figure 4.8 shows the distribution of highest level of detail used. Each bar represents the number of cases at which a user stopped giving feedback (so, if a user stopped at level 3, it is counted under level 3 only). In general, Figure 4.8 shows that the majority of cases was scored using only thumbs-based feedback. Interestingly, more cases ended with affective feedback then with stars-feedback, indicating that when more feedback is given, a preference exists for giving multidimensional affective feedback (although this difference was not statistically significant in a paired T-test comparing the number of times raters stopped at level 2 versus level 3, (paired $t(40)$=-

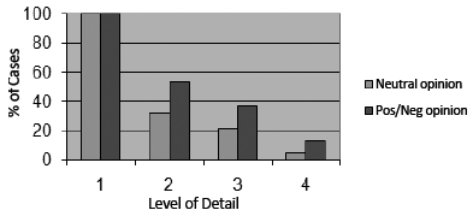


Figure 4.7: Cumulative feedback detail distribution

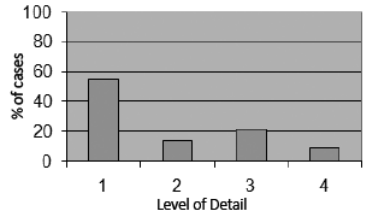


Figure 4.8: Overall feedback detail distribution

1.15, $p<0.256$)). Finally, our results show that text input was used least often by the participants. As described by others, e.g. (Ames and Naaman, 2007), incentives for tagging often have a social basis, e.g. to help others in a community to find content. The lack of community-based motivators in our study may be one of the reasons for low tagging responses. Therefore, the outcomes with regard to tagging as a feedback level may be less representative for preference feedback in general.

Summary of results

The data analysis showed that familiarity, ownership and having an opinion about that item are the main factors in influencing the preference detail people are willing to give, and thus the amount of effort they are willing to put into giving feedback. As we found only a small difference in detail for pictures versus music, we can at least tentatively conclude that the willingness to give feedback is not so much triggered by content types but more so by what one thinks or knows about content. Although both content types in our study could be owned in form of media (music, books, film)

we hypothesized that people would generally be able to give more detailed feedback for music items as they could form a preference by listening to the music during the study. This hypothesis was not confirmed and follow up studies should be done to investigate different content types.

Our results also show that multidimensional affective feedback is used when people have the choice to do so. Moreover, people in general prefer to give more feedback in the form of multidimensional affective feedback (at least when they can use the AffectButton) than to give more feedback using a finer grained one dimensional method (stars). This suggests that a preference elicitation interface — trying to adapt the amount of feedback detail it extracts from a user — should either give thumbs or stars as first level, after which the next level of detail should be affective or at least add a new feedback dimension. As the study set-up (no post-questionnaire or interview) did not allow a deeper analysis of people's reasons for giving a certain level of feedback, we do not know exactly why people who gave detailed feedback stopped more often at the AffectButton level than the stars level. One explanation would be that they just liked the AffectButton. However, we can exclude this reason because the effort that people spent to go all the way to the affective feedback level varied under the experimental conditions. Especially in the case where people had a positive or negative opinion they went to this level. If they just liked the AffectButton or wanted to try it due to its novelty they would have used it equally across all conditions. Our interpretation of why participants used the AffectButton is that due to its multiple dimensions (pleasure, dominance, arousal) the expressive power is enhanced. Whereas the stars only offer a finer grained scale on the liking dimension compared to the thumbs, the AffectButton allows people to express their attitude towards an item in two additional dimensions dominance and arousal. These dimension could be more applicable if people have a strong opinion about an item and feel the need to express this opinion. This has to be confirmed in future studies.

The obtained results give interesting insights for the design of preference elicitation interfaces used in different systems including recommenders, especially with regard to adaptive preference elicitation (also suggested by (Pu et al, 2012)). Although users of recommenders may have extrinsic motivations to give detailed preference feedback in general, it is important to know in which cases they are able or willing to give more details and in which form. Knowing that a user has a directed opinion (obtained by simple thumbs input) or is familiar with an item can be used to ask the user for more detailed and multi-dimensional input. In the case of a neutral opinion a system asking the user to spend effort of giving more details that she may be able to provide can be perceived as annoying and should be avoided. Note also the difference between

knowing an item or having a directed opinion. Often in recommender systems people do not know the items, however, by providing samples (e.g. music, book excerpts) the user can still form an opinion and by that be motivated to give more detailed feedback.

4.5 Study 3: Exploring the preference elicitation process with interface prototypes

Until now we have looked at different ways to enter a preference including rating, ranking and navigating in the first experiment, and different detail levels of rating in the second experiment. Besides motivation to spend effort, the design of the elicitation process is a second factor we would like to investigate. As (Pu et al, 2003) pointed out "stating preferences is a process rather than a one time enumeration of preferences that do not change over time". Therefore, it is important to explore how to facilitate the human preference construction by the means of preference elicitation interfaces that are intuitive for users and allow as well as motivate them to be expressive. This can only be done by involving the user in the design process. We addressed the interface design in our third study. In specific, we explored different ways of structuring the process of preference construction in an interface. Next, we elaborate on an exploratory study in which we investigated four fundamentally different processes of eliciting preferences represented in four hi-fi preference elicitation interface prototypes. Similar to the suggestion of (Pu et al, 2012) of comparing systems or interfaces side-by-side in user experiments we presented the four prototypes to each participant. In addition, we also allowed participants in a creative participatory design session to construct new ways of eliciting preferences based on (elements of) the four interfaces.

4.5.1 User-Centered Prototype Design

We created four interface prototypes for eliciting preferences. To be able to include decision context into the interface we chose to elicit preferences for a certain domain. The domain in this case was jobs, which allowed us to show example job offers as decision context. Different from the holiday domain we used in the first study where it was important to arrive at a ranking of outcomes, we wanted to support people in this study in (1) constructing their job preferences and (2) getting an idea of the resulting preference profile. Choosing to negotiate for a new job is different from picking the next holiday destination as it has a bigger impact on people's lives. This is also why we focused more on underlying interests which are stable over a longer time period and influence one's preferences. The navigational input we used in the first study was

not applied in the following prototypes because it focuses more on finding the best outcome than giving people an understanding of their preferences.

To design the prototypes we first compiled a set of design guidelines from the relevant literature. Please refer to our previously published work (Pommeranz et al, 2010) for the detailed guidelines.

Given the set of design guidelines we selected appropriate existing interface elements (e.g. ValueCharts (Carenini and Loyd, 2004), a virtual job agent) and created new ones (e.g. job offer clusters, post-it notes with preference information). Next, we combined these elements into the four interfaces. There are, of course, many combinations of elements possible, which would lead to an exponential number of prototypes. Instead of creating this high number of prototypes we combined the elements in a way that each prototype differs in how it structures the elicitation process and how it interacts with its users. Structuring the process in different ways can be linked to how people process information. Therefore, we created different ways of user-system interaction, each supporting one thinking style based on the theory by Gregorc (2006). The mind styles theory categorizes people based on perceptual and ordering preference. Perceiving information can be abstract (based reason and intuition) and concrete (using one's senses). The order of information processing can be sequential or random. This leaves us with four types: concrete sequential, concrete random, abstract sequential and abstract random. Concrete sequential thinkers like order and logical sequence and learn best in a structured environment. Concrete random thinkers like experimenting to find answers, using intuition and therefore, learn best when they are able to use trial-and-error approaches. Abstract sequential thinkers like analyzing situations before making a decision or acting and applying logic in solving or finding solutions to problems. Abstract random thinkers like to listen to others and establishing healthy relationships with others. They focus on the issues at hand and learn best best in a personalized environment. Based on these different characteristics we, first, chose an overall way of interaction, that would fit a mind style, e.g. a structured, step-wise approach for the concrete sequential thinker. Second, we identified which elements could be combined to achieve such an interaction, e.g. in the step-wise approach first a simple selection of values, then *ValueCharts* (Carenini and Loyd, 2004) showing links between values and fit of job offers, then tables with details for one offer and last an overview/summary showing the elicited preferences.

Following this approach we could create meaningful combinations of the elements. However, people do not perfectly fit into one style but have a unique combination of characteristics. In the evaluations we did not try to find the best prototype to choose

and develop further, but rather evaluate the different design elements used. In the following creative session we then gave the participants the chance to combine them in different ways that they preferred and found more usable. We implemented the designs as hi-fi prototypes because this was the best way to ensure that the users get a feeling for the interaction with the system. In the following sections we describe the four prototypical interfaces highlighting the interface elements used (*italic font*).

4.5.2 Conversation: Abstract-Random Style

This prototype (Figure4.8) focuses mainly on a collaborative interaction style, in particular the natural interaction, between the user and the system employing mixed-initiative. A natural way of building a preference model is being questioned by an expert, who can understand what you want by asking the right questions. In real life this could be a job agent. Since this is a known and intuitive way for people to express their preferences we designed a very simple interface based on a conversation with a *virtual agent*. Another design criterion used in this prototype is system transparency. We tried to reach transparency by two means: the affective state of the agent and the "thoughts" of the agent regarding the user's preferences. In the first simple version there are three states of the agent implemented, speaking with positive expression, thinking and confused. The second feature is a *thought bubble* above the agent's head. In the beginning of the conversation it is empty. It gets filled with tags (forming a *tag cloud*) whenever the agent could retrieve an interest or issue from the chat that seems to be important to the user. To ensure natural interaction during the evaluation sessions the prototype was implemented as a client-server application for a Wizard-of-Oz testing, i.e. the role of the agent was taken by a real person who was invisible to our participants.

4.5.3 Post-its: Concrete-Random Style

This prototype focuses on supporting the constructive nature of human preferences. Two things inspired the interface shown in Figure 4.9. First, preferences are rather unstructured to begin with. They are not necessarily linked to each other. Second, preferences change dependent on the context.

We used *post-it notes* as a real-world metaphor for organizing thoughts. The interface allows dragging as many post-it notes onto the so-called preference view as the users want. They can then write the important issues on the notes, add a value and specify whether they like, want, dislike or do not want these issues. At any time they

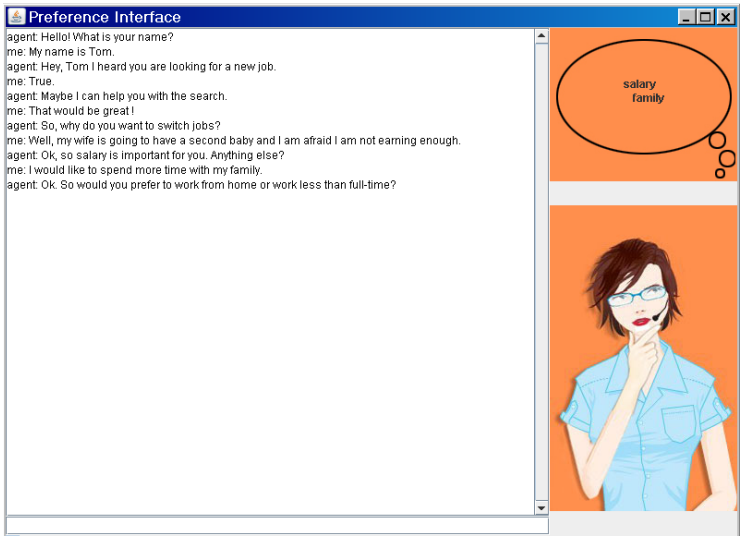


Figure 4.9: User interface for conversation with intelligent agent

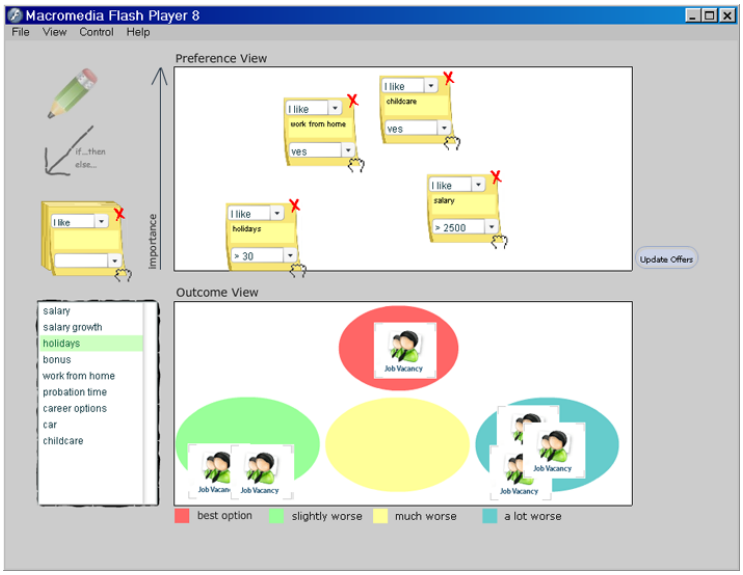


Figure 4.10: Visual construction of preference profile.

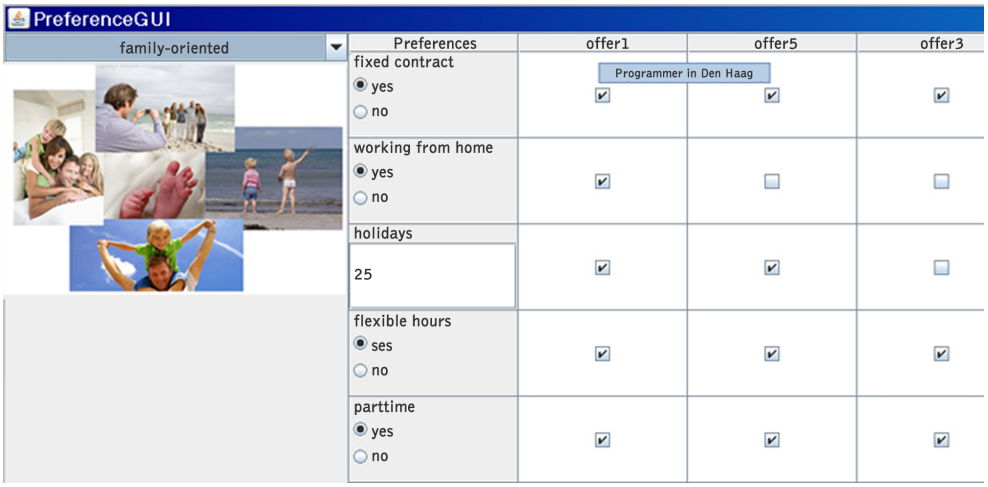


Figure 4.11: Choosing and adjusting a default profile.

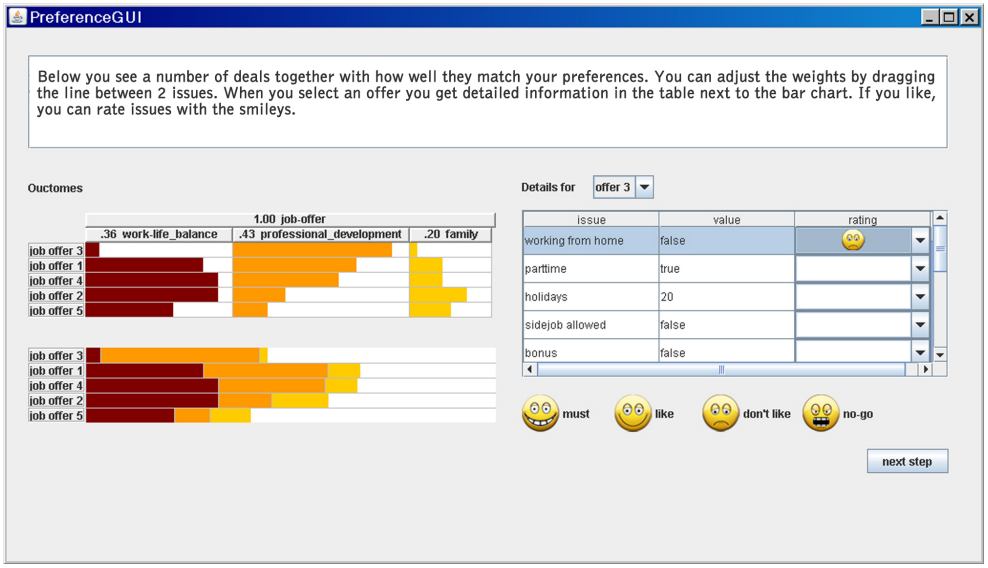


Figure 4.12: Preference elicitation using ValueCharts and affective feedback.

can remove, add or drag around the post-its to structure their profile. More important issues can be dragged further up and less important ones down.

At the same time we provide the users with the needed context to make their choices of how to structure the notes. The context is a number of job offers in the *outcome view* that get arranged into *clusters* according to good fit to the current preference profile. This could be done in real-time while the user is interacting with the notes to give immediate visual feedback. For simplicity reasons the arrangement takes place after pressing the "update offers" button. In the evaluation we discussed both options.

4.5.4 Comparison: Abstract-Sequential Style

In this prototype (figure 4.10), based on the value-focused thinking approach the user chooses from a list of *interest profiles*: family-oriented, money-oriented, career-oriented, or self-fulfillment. We chose these profiles because they represent life goals that are linked closely to jobs. In a real system this needs to be scientifically proven. In order to help people choose a profile we added a visual stimulus to each profile. We chose a moodboard-like collection of images as often used in advertising to convey a certain feeling or style. Each moodboard consists of a collection of images that represent the particular profile at a glance. The selection of images aimed at giving a diverse view of the profile (e.g. career profile: doctor, model, business man etc.) in order to avoid that users focus too much on a particular image. In the second step, the user received a filled-in list of preferences that fit the chosen profile. To give the user decision context to understand their preferences and refine the preselected ones we present a list of job offers.

The data is presented in form of a *decision matrix* similar to the ones often used on product comparison websites. Both the preferences and the offers are ordered by importance, from top to bottom and left to right respectively. By hovering over the job offer with the mouse the user gets a description of the jobs. Since we are not expecting that people fit perfectly into a profile the users have the chance to adjust the preference values as well as the ordering. As soon as they enter a new value or drag and drop the rows around the job offers get ordered based on the new input to give visual feedback of the consequences. We use a lexicographic ordering, since it delivered good results in our first study. During the evaluations we also discussed the possibility for the user to drag the job offers, which will result in adapted preferences.

4.5.5 Stepwise: Concrete-Sequential Style

In the fourth prototype (figure 4.11) the interaction is similar to the APT Decision agent (Shearin and Lieberman, 2001) following three steps: (a) letting the user give only a small number of preferences, (b) then receiving a list of offers to compare and (c) giving feedback to attributes that appear in the offers. We adapted this approach and ask the users in the first stage about their three most important interests (e.g. work-life balance or professional development) instead of negotiable issues. By that we follow the value-focused thinking approach (Keeney, 1992). After choosing the interests the user enters the interface depicted in Figure 4.11. The interface aims at helping the user explore several job offers (decision context) with regard to the user's interests and by that construct his preference profile. To compare the offers we used *ValueCharts* (Carenini and Loyd, 2004). The user can adjust the (initially equal) importance of the interests. He receives immediate visual feedback on how well the job offers match his interests, while adjusting the importance by growing or shrinking of the job offer bars. By double clicking on an interest the job offers get ordered according to good fit. The interface also offers the possibility to critique any attribute of a job offer. Once the user chooses to look at a job offer in more detail the table on the right gets filled with all values for existing attributes in the job offer. The users are free to give affective feedback on any issue-value pair they want, but are not forced to rate all of them. We included "musts" and "no-goes" as hard constraints in the system, i.e. a job that does not comply with either will not be an option to the users. When the user is done exploring his options, the interface reveals an overview over elicited preference profile, which supports the transparency of the system.

4.5.6 Exploratory User Study

In order to understand in depth how we can support the human process of preference construction with an adequate interface we did an exploratory user study. By collecting large amounts of qualitative data we aimed at informing the design process of preference elicitation interfaces. Our prototypes served as a means to discuss relevant issues to the participants and foster a creative process rather than finding usability problems of the prototypes. We specifically aimed at receiving feedback on the different interface elements used and how they can be combined in an optimal way to support the process of constructing one's preferences. In the following sections we elaborate on the set-up of the study and its outcomes.

Figure 4.13: Interface Elements for Creative Session.

Material

We used the four hi-fi prototypes elaborated above in this user study. Furthermore, we created paper versions of all interface elements we had used in the four hi-fi prototypes (Figure 4.12), e.g. the virtual agent, the post-its, the value charts or the tag cloud, as well as standard interface elements such as text fields, check boxes, sliders, comboboxes, etc. Additionally, we had a number of blank papers, pens and scissors to give the participants the chance to create their own interface elements. These materials were used by the participants in the second part of the session to design their own preference elicitation interfaces.

Participants

We included 5 male and 3 female participants. The participants were people with different backgrounds, i.e. artificial intelligence, affective computing, design, linguistic and visual perception. We intended to have a mixture of people with diverse backgrounds in order to get different views on the interfaces.

Procedure

The study was divided into two parts: eight individual sessions with one participant at a time and a collaborative creative session with all eight participants.

The sessions were carried out in a lab setting. Participants were first briefed about the background of the study and the intention. We emphasized that we would like to receive constructive feedback on the different elements of the prototypes to inform

future designs of preference elicitation interfaces. After the briefing we provided the participants with a scenario describing a 35 year-old family father who would like to switch jobs. We chose using a scenario rather than the participants' real job preferences for two reasons. The first is of practical nature: Since our interfaces were limited regarding their domain knowledge, we wanted to make sure that the issues and interests people want to express preferences over were available in the system. The second reason was trying to get participants to use the interfaces in a similar way to be able to compare the feedback. The participants then interacted with each prototype for about 10 minutes on average. The order of prototypes was changed per participant to avoid ordering effects. Their task was to fill in job preferences that would fit the person described in the scenario. During the interaction the participants were asked to think aloud. All actions and voices of the participants were recorded by the help of the Camtasia Studio software (http://www.techsmith.com/camtasia.asp). Each prototype saved the preferences to a log file. The person leading the evaluation intervened whenever participants seemed to be lost, asked for help or forgot to think aloud. Often the evaluator and the participant already got into discussions about new ideas and problems with the interfaces during the interaction. After interacting with the prototypes we interviewed the participants informally to get a grasp of their experiences, constructive critique and new ideas. We used printed screenshots of the interfaces as reminders. Together with the evaluator new ideas were developed and discussed and drawn onto the printed screenshots.

The individual sessions were followed by a creative session with all eight participants. Goal of this session was to explore new ways to structure the elicitation process in the interface from the users' point of view. The session consisted of two parts, a group discussion and participatory design session aimed at creating new paper prototypes. After a short introduction to the meeting including a reminder of all four interfaces and the agenda, we started a general discussion about the interface elements. The discussion took part with the whole group for about 20 minutes. After that we split the participants into two groups of four participants each. Each group was provided with the same set of materials described above and instructed to use the material to create their own version of a preference elicitation interface. They were encouraged not only to combine the elements existing in the four presented prototypes but also create new ones. This part of the creative session was planned for about 30 minutes. However, since both groups were not done within that time frame, the session took about 1 hour. The creative session was concluded with a presentation of the two groups' results to each other. During the presentation new discussions arose about design decisions.

4.5.7 Results

During the individual sessions, the informal interviews and the group discussions we gained detailed feedback on the four prototypical interfaces as well as new ideas, including tips and new combinations of the interface elements. In order to extract the feedback from the collected data we annotated the recordings from the individual sessions using NVivo (www.qsrinternational.com). Based on the annotations we created a table with feedback on each prototype per participant. In addition, we made a list of observations of how users used the prototypes and a list of new ideas that were discussed in the individual and the collaborative session. Next, we will elaborate on the main findings that are relevant for designing preference elicitation interfaces. Table 4.4 shows the positive and negative comments per element. For a more detailed description of the feedback per interface element, we refer to our previously published work (Pommeranz et al, 2010).

Some of the interface elements had obvious usability issues, e.g. the checkboxes in the decision matrix which were not interactive. These were due to programming difficulties or time constraints during the creation of the prototypes. As we already anticipated some of these issues before conducting the study, a researcher was present during the study to clarify such issues whenever a participant seemed to have a problem. We asked the participants to focus on the fit of the different elements for entering preferences. We also asked for constructive feedback on improving and combining the elements once the users understood how the elements worked. Regardless of the interface elements used, an important aspect for our participants was the ability to explore the link between their preference input and the desirability of outcomes (in this case job offers). An element that the participants found highly useful for this exploration were the ValueCharts, because they give immediate visual feedback while keeping details about the selected interests/issues. Another well-liked element supporting the construction of preferences was the post-it note. Furthermore, using default profiles was anticipated since it gets the elicitation process started more easily than starting from scratch. Based on a given profile a number of common preferences can already be displayed. Carefulness needs to applied with designing the interface in this case. Most people had trouble fitting themselves into one of the four given profiles. Therefore, a more flexible input of the separate interests should be possible. During the collaborative session several ideas were mentioned to create a more flexible input of interests, e.g. using a questionnaire, pictures combined with sliders for importance or the virtual agent.

This feedback was also reflected in the new preference elicitation interfaces that

Table 4.4: Feedback per interface element.

Element	Positive	Negative
virtual agent	engaging, straightforward, natural way to enter preferences, easy to use, no constraints	low feasibility, too slow, vague, profile not clear, no comparison of jobs, depends on how good the agent is
tagcloud/ thought bubble	gave users a hint of what the system is "thinking"	-
post-its	liked by most participants	too difficult to operate, too many hidden things
Outcome view/ clusters of offers	participants liked seeing and exploring job offers, tie between preferences and consequences	offers were not draggable
Interest Profiling	most users liked it as a starting point, efficient, less effort, use of pictures	trouble deciding on one fitting profile (preferences should already be visible when choosing)
Decision matrix	similar to product comparison websites	problems with visualization: difficult to understand that offers are ordered and draggable
ValueChart	gives an overview of how the job offers fit the profile but without losing the detailed information of how well each interest/issue scores in an offer,immediate visual feedback	no link between the ValueChart and the table with issue ratings
Affective feedback	natural	must-have smiley was not interpreted as a hard constraint
Preference Summary	was liked, gives overview, clarifies preference profile	should appear while you are adjusting your preferences, missing interactivity

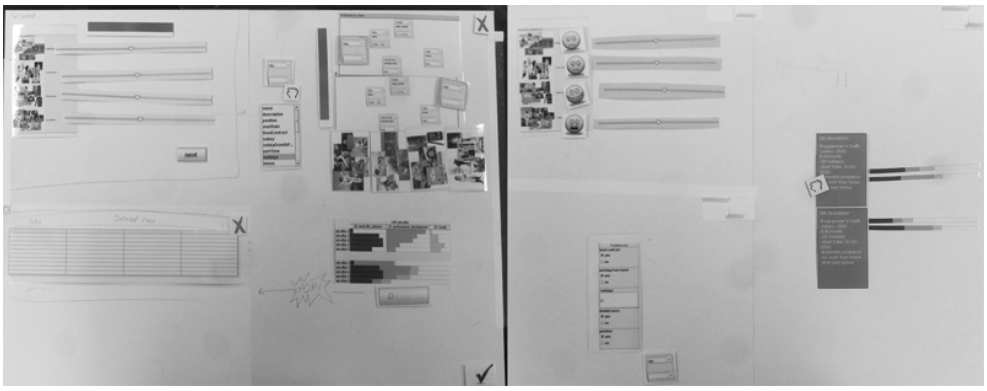


Figure 4.14: Design proposal group 1 (left) and group 2 (right).

the two groups designed in the second part of the collaborative session. The results are depicted in figure 4.13. Both groups were in favor of having three views on their preferences, i.e. the underlying interest profile, the issue preferences including an importance ranking and values for each issue, and a number of job offers representing the decision context. Whereas group 1 left it all up to the user where to start in the interface and which views to maximize/minimize, group 2 focused on (stable) underlying interests in the first step before giving a number of preferences in the context of example job offers.

Regarding our hypothesis (users are willing to spent more effort if the feedback mechanism (process and preference representation) enables them to be more expressive) we can conclude that people are indeed willing to spend more time on investigating the links between their interests, issue preferences and outcomes (jobs). This was mentioned by the participants and observed by the researcher during the study. The participants emphasized that it allows them to be more in control of creating their own preference profile, which will then be used by the system. Having that level of control and understanding of the system's model was anticipated by the participants (see similar results on user control in (Pu et al, 2012)). We believe, this shows the importance of supporting this constructive process in order to make the outcome of the system comprehensible and trustworthy. However, participants also expect the system to support them during this exploration of links where possible, e.g. by offering default preferences based on profiles and by giving immediate visual feedback while adjusting the different elements. They are not willing to spend much time on cognitively demanding tasks that do not seem necessary (e.g. creating every post-it).

4.6 Discussion and Design Guidelines

In the three studies we presented, we tackled the problem of designing user interfaces for explicit preference elicitation. Two important aspects to consider when designing such interfaces are: matching the mental models of users' preferences to the representations of the system and supporting the process of human preference construction so that "true" preferences can be elicited by the system. In our first study we investigated both aspects by studying (a) different ways of giving preference feedback (process) and (b) what kind of information the methods deliver and how the outcomes (ranked holiday lists) compare to a baseline created by the participant. We learned that effort generally goes hand in hand with liking when comparing tasks that are similar with regard to the process and type of input (e.g. rating with Likert scale or ordering attributes). However, in cases where the process (navigation) and the type of feedback (affective) was more sophisticated in terms of expressive power and understanding of one's own preferences, participants rated the methods high in liking even though the results show a substantial increase in perceived effort or are less easy to use. Therefore, we hypothesized that people are willing to spend more effort if the feedback mechanism enables them to be more expressive. In the two following studies we tested this hypothesis.

The following online rating experiment focused on the motivation people have to give feedback in a neutral setting (by that we mean that they are not motivated e.g. by social aspects as it is often the case in recommender systems) and which factors influence that motivation. The main factors we found were familiarity of an item (also predicted by ownership) and whether people already have a formed opinion about the content. Furthermore, we could conclude that once people decided to give more levels of feedback they went more often all the way to the affective feedback level than just the 6-point star rating. While we can conclude safely that an interface should offer motivated users the possibility to enter more detailed feedback (guideline 1), we do not know the exact reasons for people to enter affective feedback (with the AffectButton). Given the fact that the star based rating offers only a finer grained one-dimensional feedback (liking) compared to the thumbs, whereas the AffectButton offers two additional dimensions (dominance, arousal), we believe it offers more expressive power. In the case of people having a defined (positive or negative) opinion on an item they might feel the need to express this opinion with more detail and on more dimensions. Based on the fact that people liked giving affective feedback despite increased effort (compared to traditional methods) in the first study and more participants in the second study stopped at this level than at the stars level we can say

that affective feedback should be considered when detailed preference feedback is needed (guideline 2).

After studying motivation in this structured way, we took a more exploratory approach in the third study to understand how to design the process of preference elicitation interfaces from a users point of view. By actively involving the participants in the design process we were able to understand how they prefer an interface to be designed. We learned that an important aspect of the process is that it allows people to understand their own preferences and that people feel in charge of creating their profile as opposed to just answering questions that are used by the system to build the profile. In particular, being able to explore their preferences from different angles including underlying interests and consequences (in form of rankings of decision outcomes) within the same interface supported people's process of constructing their preferences. Participants liked design elements that supported this exploration in a natural way that allowed immediate visual feedback. Whereas design guidelines established earlier (Pu and Chen, 2008) already point to giving decision context and immediate visual feedback, we would like to add the importance of exploring *interests, preferences* and *outcomes in the same physical space.* This enables the user to receive feedback on three related concepts at the same time while adjusting one of the views, which is not the case in interfaces proposed by Pu and her colleagues. As participants were in favor of this kind of interaction and view of their preferences we believe there is a basis for a new guideline (see guideline 3).

Furthermore, the study supported results from our first study regarding the effort people would like to spend. People preferred using interest profiles as a first step and getting preference suggestions from the system (on an attribute basis). The comments of the participants indicated that they considered starting *from scratch* (i.e. filling in values for every attribute themselves) as an effortful task that seems to be redundant if the system is able to give suggestions based on the interest profile (guideline 4).

Given the results from the three studies we established the four following design guidelines for preference elicitation interfaces:

(*1*) *As motivated users are willing to spend more effort, users should be given the option to express more detail if they feel the need to do so.*

(*2*) *Affective feedback should be considered as a way for specifying detailed preference feedback with multiple dimensions.*

(*3*) *The user must be able to explore his/her interests, preferences and outcomes in the same physical space in a way that gives immediate feedback on the links between the three concepts.*

(*4*) *Profile/interest selection serves as an easy (i.e. reduced effort) starting point for showing default preferences that can subsequently be adapted by the users.*

These specific guidelines are meant to extend the more general existing guidelines from the literature (e.g. giving immediate visual feedback, context in form of example outcomes, focusing on values, any preference in any order etc., see (Pu et al, 2003) and (Pommeranz et al, 2010)) instead of being an exhaustive list by themselves.

4.6.1 Limitations and further investigations

Our goal was to inform the design of preference elicitation interfaces in general. The results should therefore not be restricted to specific tasks or systems. We believe that they are generally valid for preference elicitation done for recommender systems as well as decision support systems. However, the research questions we investigated had an influence on the choice of domain and type of tasks for each experiment. We chose holidays in the first study with the assumption that most people either have holiday preferences or are able to construct them easily. The focus of the study was on preference input mechanisms in connection to their use in an algorithm that computes a preference ranking over outcomes. In order to compare the different outcoming lists to a baseline we asked people to give their own ordering of the items in the outcome space. This limited the size of our domain to a great extend, as with nine (3 properties times 3 alternative values) property values the number of holidays that could be created was already 27. We thought that sorting an even higher number of holidays would be an overwhelming task for the participants, and the effort ratings confirm this. The limitations to the value space of the properties, however, poses difficulties to transfer the results of the study to other domains as most real-world applications of recommender systems deal with a high number of values, properties and outcomes. Especially, the navigational task would not be feasible in the same way as in study 1 if the number of values and properties was higher than three. It would have to be adapted by using an intelligent algorithm showing only a small portion of the outcome space at a time. If people still liked the task in other scenarios would have to be retested. Another aspect of study 1 that leads to a limitation of the results is the fact that we tested only the lexicographic algorithm to generate outcome rankings. To generalize the results connected to the liking and similarity of outcome lists other

algorithms should be employed and a detailed investigation needed to be done of how to map the 3-dimensional feedback obtained by the AffectButton into a 1-dimensional outcome ranking.

Considering the second study two things need further investigation. One is the relation of the strength of an opinion to the need to give feedback. In the current set-up this was not possible. Second, we need to further investigate why people preferred to give more detailed feedback in form of affective feedback with the AffectButton and how to use this multi-dimensional feedback as additional information on the user's preferences.

Based on the results from the third study questions about the design of the interest profiling arose. Interesting work that we will consider for this aspect has been done by (Kay, 2000), who focused on the scrutable student models in learning environments. Scrutable stereotypes are used to support learners in tuning their student models. By scrutinizing the models the user can also understand what the system believes about them and what these beliefs are based on.

With regard to the guidelines, it has to be noted that whereas the first guideline is applicable to any preference elicitation task, guideline 2 is more suited for domains in which the user is either familiar with the items or can easily form an opinion about an item (e.g. music or book recommenders). Guidelines 3 and 4 are focused more on domains in which users have to construct preferences (due to being a novice or changing preferences). Guideline 3 is especially helpful for negotiation/decision support systems or recommenders that advice users in important decision-making tasks (real estate, financial advice, job negotiations etc.). Guideline 4 is useful in domains where the number of properties are very high (e.g. cameras or other electronic devices with many features).

4.7 Conclusion

The importance of preference models for intelligent systems of different sorts (e.g. recommender systems, decision support systems) has long been acknowledged by researchers. However, focus within the area of preference modeling has been mainly on algorithms for computing preferences (elicited in form of numbers and weights) and system representations. A group of researchers has lately focused on designing methods for preference elicitation from a user's point of view, that are in accordance with behavioral decision making theories (constructive preferences). More research in this direction is needed to give researchers and practitioners a good understanding of

how to design trustworthy preference elicitation interfaces, that involve users in the process of constructing their own profile that reflects true preferences. We have pointed to two main difficulties that still exist, namely matching people's mental models and the influence of the elicitation process on the elicitation outcome. Furthermore, our studies showed that affective factors are important to consider in preference elicitation. The results suggest that more research in this direction seems worthwhile. With the studies presented in this paper we have only done a first step towards an optimal design of preference elicitation interfaces. However, we believe that the results we obtained from tackling the problem in different ways (with structured experiments and exploratory, participatory research) help in advancing the research on interface design for preference elicitation and encourage others in the field to follow that route. Our own research agenda includes usability testing of the interface design obtained from the third study as well as investigating the links between underlying values and attribute preferences.

4.8 Bibliography

Adomavicius G, Tuzhilin A (2005) Towards the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering 17(6):734–749.

Aloysius JA, Davis FD, Wilson DD, Taylor AR, Kottemann JE (2006) User acceptance of multi-criteria decision support systems: The impact of preference elicitation techniques. European Journal of Operational Research 169:273–285.

Ames M, Naaman M (2007) Why we tag: motivations for annotation in mobile and online media. In Proceedings of the SIGCHI conference on Human factors in computing systems, San Jose, CA, 971–980.

Ardissono L, Felfernig A, Friedrich G, Goy A, Jannach D, Petrone G, Schäfer R, Zanker M (2003) A framework for the development of personalized, distributed web-based configuration systems. AI Magazine 24:93–108.

Barneveld J, Setten M (2004) Designing usable interfaces for tv recommender systems. In: Personalized Digital Television, Human-Computer Interaction Series, vol 6, Springer Netherlands, 259–285.

Bellucini E, Zeleznikow J (2006) Developing negotiation decision support systems that support mediators: a case study of the familywinner system. Journal of Artificial Intelligence and Law 13(2):233–271.

Bettman JR, Luce MF, Payne JW (1998) Constructive consumer choice processes. Journal of Consumer Research 25(3):187–217.

Broekens J, Brinkman WP (2009) Affectbutton: Towards a standard for dynamic affective user feedback. In: Mühl C, Heylen D, Nijholt A (eds) Proceedings of Affective Computing and Intelligent Interaction (ACII), IEEE Computer Society Press, Amsterdam, The Netherlands, 1–8.

Broekens J, Pommeranz A, Wiggers P, Jonker CM (2010a) Factors influencing user motivation for giving preference feedback. In: 5th Multidisciplinary Workshop on Advances in Preference Handling in Conjunction with ECAI 2010, Lisbon, Portugal, 19–24.

Broekens J, Pronker A, Neuteboom M (2010b) Real time labeling of affect in music using the affectbutton. In Proceedings of the 3rd international workshop on Affective interaction in natural environments, AFFINE '10, Firenze, Italy, 21–26.

Burke R (2000) Knowledge-based recommender systems. In: Kent A (ed) Encyclopedia of Library and Information Systems, vol 69, Marcel Dekker, New York.

Burke R (2002) Hybrid recommender systems: Survey and experiments. User Modeling and User-Adapted Interaction 12(4):331–370.

Burke RD, Hammond KJ, Young BC (1996) Knowledge-based navigation of complex information spaces. In: 13th National Conference on Artificial Intelligence, AAAI Press, Portland, OR,462–468.

Carenini G, Loyd J (2004) Valuecharts: analyzing linear models expressing preferences and evaluations. In: AVI '04: Proceedings of the working conference on Advanced visual interfaces, ACM, New York, NY, 150–157.

Carenini G, Poole D (2002) Constructed preferences and value-focused thinking: Implications for ai research on preference elicitation. In: AAAI-02 Workshop on Preferences in AI and CP: symbolic approaches, AAAI, Edmonton, Canada, 1–10.

Carenini G, Smith J, Poole D (2003) Towards more conversational and collaborative recommender systems. In: 8th international conference on intelligent user interfaces, ACM, Miami, FL, 12–18.

Chen L, Pu P (2004) Survey of preference elicitation methods. Tech. rep., Swiss Federal Institute of Technology In Lausanne (EPFL), Lausanne.

Chen L, Pu P (2009) Interaction design guidelines on critiquing-based recommender systems. User Modeling and User-Adapted Interaction Journal (UMUAI) 19(3):167–206.

Curhan JR, Neale MA, Ross L (2004) Dynamic valuation: Preference changes in the context of face-to-face negotiation. Journal of Experimental Social Psychology 40(2):142 – 151.

Faltings B, Pu P, Torrens M, Viappiani P (2004) Designing example-critiquing interaction. In: IUI '04: Proceedings of the 9th international conference on Intelligent user interfaces, ACM, New York, NY, 22–29.

Fano A, Kurth SW (2003) Personal choice point: helping users visualize what it means to buy a bmw. In: IUI '03: Proceedings of the 8th international conference on Intelligent user interfaces, ACM, New York, NY, 46–52.

Fischer GW, Carmon Z, Ariely D, Zauberman G (1999) Goal-based construction of preferences: Task goals and the prominence effect. Management Science 45(8):1057–1075.

Gregorc A (2006) The Mind Styles Model: Theory, Principles, and Practice. AFG.

Herlocker JL, Konstan JA, Terveen LG, Riedl JT (2004) Evaluating collaborative filtering recommender systems. ACM Trans Inf Syst 22(1):5–53.

Hunt D, Haynes R, Hanna S, Smith K (1998) Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review. Journal of the American Medical Association 280(15):1339–46.

Johnson E, Steffel M, Goldstein D (2005) Making better decisions: from measuring to constructing preferences. Health Psychology 24(8):17–22.

Kay J (2000) Stereotypes, student models and scrutability. In: Gauthier G, Frasson C, VanLehn K (eds) Intelligent Tutoring Systems, Lecture Notes in Computer Science, vol 1839, Springer Berlin / Heidelberg, 19–30.

Keeney R (1992) Value-Focused Thinking: A Path to Creative Decision Making. Harvard University Press.

Keeney R, Raiffa H (1993) Decisions with Multiple Objectives: Preferences and Value Trade-Offs. Cambridge University Press.

Knijnenburg B, Willemsen M, Ganter Z, Soncu H, Newell C (2012) Explaining the user experience of recommender systems. User Modeling and User-Adapted Interaction 22(4-5), 441–504.

Kramer T (2007) The effect of measurement task transparency on preference construction and evaluations of personalized recommendations. Journal of Marketing Research 44(2):224–233.

Linden G, Hanks S, Lesh N (1997) Interactive assessment of user preference models: The automated travel assistant. In: User Modeling: Proceedings of the Sixth International Conference, Vienna, Austria, 67–78.

Liu F (2008) Changing for the better: Preference dynamics and agent diversity. PhD thesis, University of Amsterdam.

McCarthy KJ, Reilly K, McGinty L, Smyth B (2005) Experiments in dynamic critiquing. In: Proceedings of the 10th International Conference on Intelligent User Interfaces, IUI'05, San Diego, CA, 175–182.

McFadden D (1999) Rationality for economists? Journal of Risk and Uncertainty 19(1):73–105.

Mehrabian A (1980) Basic dimensions for a general psychological theory: Implications for personality, social, environmental, and developmental studies. Cambridge, MA.

Miller BN, Albert I, Lam SK, Konstan JA, Riedl J (2003) Movielens unplugged: experiences with an occasionally connected recommender system. In: Proceedings of the 8th international conference on Intelligent user interfaces, ACM, New York, NY, USA, IUI '03, pp 263–266.

Mooney R, Roy L (2000) Content-based book recommending using learning for text categorizations. In Proceedings of the Fifth ACM Conference on Digital Libraries, ACM, San Antonio, TX, 195–204.

Ono C, Kurokawa M, Motomura Y, Asoh H (2007) A context-aware movie preference model using a bayesian network for recommendation and promotion. In: Conati C, McCoy K, Paliouras G (eds) User Modeling 2007, Lecture Notes in Computer Science, vol 4511, Springer Berlin / Heidelberg, 247–257.

Payne JW, Bettman JR, Schkade DA (1999) Measuring constructed preferences: Towards a building code. Journal of Risk and Uncertainty 19(1-3):243–270.

Pazzani MJ, Billsus D (2007) Content-based recommendation systems. In: Brusilovsky P, Kobsa A, Nejdl W (eds) The adaptive web: methods and strategies of web personalization, Springer-Verlag, Berlin, Heidelberg, pp 325–341.

Peintner B, Viappiani P, Yorke-Smith N (2008) Preferences in interactive systems: Technical challenges and case studies. AI magazine 29(4):13–24.

Pommeranz A, Broekens J, Visser W, Brinkman WP, Wiggers P, Jonker C (2008) Multi-angle view on preference elicitation for negotiation support systems. In: Brinkman WP, Hindriks K (eds) Proceedings of First International Working Conference on Human Factors and Computational Models in Negotiation (HuCom08), Delft University of Technology, Mediamatica, Delft, The Netherlands, pp 19–26.

Pommeranz A, Wiggers P, Jonker C (2010) User-centered design of preference elicitation interfaces for decision support. In: Leitner G, Hitz M, Holzinger A (eds) USAB2010 - HCI in Work & Learning, Life & Leisure, Springer, Klagenfurt, Austria, Lecture Notes in Computer Science, vol Vol. 6389, pp 14–33.

Pu P, Chen L (2007) Trust-inspiring explanation interfaces for recommender systems. Knowledge-Based Systems 20(6):542 – 556.

Pu P, Chen L (2008) User-involved preference elicitation for product search and recommender systems. AI Magazine 29(4):93–103.

Pu P, Faltings B, Torrens M (2003) User-involved preference elicitation. In: Workshop notes of the Workshop on Configuration, the Eighteenth International Joint Conference on Artificial Intelligence, IJCAI'03, 56–63.

Pu P, Chen L, Hu R (2012) Evaluating recommender systems from the user's perspective: Survey of the state of the art. User Modeling and User-Adapted Interaction 22(4-5), 317–355.

Resnick P, Iacovou N, Sushak M, Bergstrom P, Riedl J (1994) GroupLens: An open architecture for collaborative filtering of netnews. In: Smith JB, Smith FD, Malone TW (eds) ACM Conference on Computer Supported Collaborative Work Conference, Association of Computing Machinery, ACM Press, Chapel Hill, NC, 175–186.

Shearin S, Lieberman H (2001) Intelligent profiling by example. In: Sidner C, Moore J (eds) IUI '01: Proceedings of the 6th international conference on Intelligent user interfaces, ACM, New York, NY, USA, pp 145–151.

Shiv B, Fedorikhin A (1999) Heart and mind in conflict: the interplay of affect and cognition in consumer decision making. Journal of Consumer Research 26(3):278–292.

Simon D, Krawczyk DC, Holyoak KJ (2004) Construction of preferences by constraint satisfaction. Psychological Science 15(5):331–336.

Sinha R, Swearingen K (2002) The role of transparency in recommender systems. CHI '02 extended abstracts on Human factors in computing systems, ACM, New York, NY, USA, 830–831.

Smyth B, McGinty L (2003) The power of suggestion. In: Gottlob G, Walsh T (eds) IJCAI-03, Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence, Morgan Kauffman, Acapulco, Mexico,127–132.

Stolze M, Ströbel M (2003) Dealing with learning in ecommerce product navigation and decision support: the teaching salesman problem. In: Piller FT, Reichwald R, Tseng M (eds) In Proceedings of the Second Interdisciplinary World Congress on Mass Customization and Personalization, TUM, Munchen, Working Paper Series of the Department for General and Industrial Management, TUM Business School, Technische Universität München.

Viappiani P, Faltings B, Zuber VS, Pu P (2005) Stimulating preference expression using suggestions. In: Aha DW, Tecuci G (eds) Proceedings of the AAAI Fall Symposium on Mixed-Initiative Problem Solving Assistants (FS-05-07), AAAI, Washington, DC, 128–133.

Weber EU, Johnson EJ (2006) Constructing Preferences from Memory. Cambridge University Press.

CHAPTER 5

SELF-REFLECTION ON PERSONAL VALUES

A good starting point for the construction of preferences, as presented in the previous chapter, are underlying values of a user. From the decision making literature we have seen that underlying values are more stable than preferences over longer periods of time and a major factor in taking decisions. Furthermore, sharing underlying values between negotiation partners can lead to trade-offs suited to both parties and thus better negotiation outcomes. We have already used value profiles in the preference elicitation interface prototypes discussed in the previous chapter. However, we learned from user feedback that people have complex individual value systems that cannot be easily captured in a predefined profile. Currently there are no digital tools available to support people in defining their value profiles with regard to decisions. This chapter deals with the design of such a tool for the job negotiation domain. We elaborate on the design of a prototype for value-reflection (section 5.4), which was used in subsequent user studies (section 5.5) and an online survey (5.6). We conclude with a list of design guidelines for value elicitation tools.[1]

[1]This chapter is equivalent to: Alina Pommeranz, Pascal Wiggers, Catholijn M Jonker. Designing for self-reflection on values for improved life decisions. Interacting with Computers (submitted in March 2012).

Life is sacred, that is to say, it is the supreme value, to which
all other values are subordinate.
Albert Einstein (1879-1955)

5.1 Introduction

In today's globalized world taking major life decisions is a complex task. For example, people do not or cannot simply consider their local environment when deciding where to live, and occupational choices are vast (as opposed to earlier generations who often followed into the steps of their fathers). Take, e.g. a young family in which both mother and father have a job and take care of their children. The decision of one parent to apply for a new job depends, among other things, on the number of alternatives in the current job market, the person's needs and wishes in terms of career opportunities, tasks, involvement with other people, salary etc. In addition, the decision requires consideration of the family, e.g. the time left for taking care of the children; whether they would have to move, which will in turn affect the spouse's work and life situation; and maybe in some cases even the distance to other relatives. This type of life-decision making requires balancing one's own needs and those of other stakeholders, such as family members or negotiation partners, e.g. employers. Furthermore, consequences of these decisions are long-lasting.

As a process, decision making requires assessing an often vast set of alternatives according to one's preferences in order to find a suited outcome. Economic theories, e.g. (Coleman and Fararo, 1992), assume this to be a matter of rationally calculating each option's utility based on given, stable preferences. However, this view does not represent real life decision making. Especially untrained decision makers (i.e. people who are not familiar with decision-theories, the domain or decision making as a major part in their daily work) often follow an adaptive model (Payne et al., 1993).

According to this adaptive model individuals simplify decision making through applying choice heuristics as a response to their own limited cognitive processing abilities in complex decision tasks (Payne et al., 1993, p. 2). As partially reflected in the example above, three major types of factors influence the choice of a decision strategy or heuristic, the problem characteristics (e.g. the task variables), the decision maker's characteristics (cognitive ability and prior knowledge) and the social context (e.g. accountability). This decision making behavior can lead to choices that are not optimal. Especially in complex situations as outlined above, people 'zoom in' quickly on a small set of alternatives, find the best among these and then try to justify their choice to others, e.g. by adapting their original preferences. Keeney (1992) has entitled this behavior alternative-focused thinking and argued that instead a value-focused thinking approach would help people improve their decision making. He suggests seeing decision making as a creation of decision opportunities rather than decision problems. This is possible by first identifying and analyzing one's values and

then creating or looking for alternatives that would fit these. Other studies (Arvai and McDaniels, 2001) have shown that this approach leads to better decision making. In particular, people felt more comfortable with their decisions, more knowledgeable about relevant issues to make an informed decision and considered a wider array of decision-relevant issues.

While value-focused thinking seems promising, it requires time and effort to gain experience and understanding of one's values. Values are abstract concepts that are often hard for people to understand and articulate (LeDantec et al., 2009). This has been acknowledged by Keeney (1992) ("It may initially be difficult to articulate, review and revise your objectives.", p 549). While Keeney suggests a dialogic approach, in which an advisor supports a client to identify values by asking many open questions, we believe, technology can be designed to support people in self-reflection on personal values. A mobile tool, e.g. in form of a website or mobile application, would provide an additional resource for decision makers to reflect whenever they feel like it and not just in dialog with a coach. Also situated reflection would be possible, i.e. whenever a significant situation has occurred or generally in the context relevant to the decision making context, e.g. thinking about work values at work.

As part of a 4-years research program (Hindriks and Jonker, 2008) aiming at computer-supported decision and negotiation support of non-expert decision makers, we are designing new technologies for improved decision making. As opposed to current decision support system (DSS) design which focuses on technical solutions implementing economic theories (Carenini and Poole, 2002), we propose a human-centered approach taking into account the cognitive and emotional characteristics of human-decision making in the design process. The work presented in this paper focuses on the first step of a value-focused approach to decision support, i.e. supporting decision makers with digital tools to understand their values and the relative importances in the given decision context. Our main research question is 'How can we design digital tools to support people in reflecting on and assessing their values?' In the following we elaborate on our work's theoretical framework of decision making behavior, which we sketched only shortly above. Further, we outline our approach to the design of value-reflection tools to support informed decision making and define concepts of awareness and self-reflection. Next, we describe our design-related work, including the creation of a prototype for value-reflection in the job context. We will present a user study of the prototype and a user workshop to gain new design insights. Given these insights we sketched design ideas that were tested with a sample of 82 people in an online survey. Based on the analysis of respondent's preferences for the designs we created five design guidelines provided in the discussion section.

5.2 Background

5.2.1 Decision making

Research on decision making is vast and coming from several disciplines, such as economics, psychology and organizational science. As a comprehensive overview is outside the scope of this paper, we will concentrate in the following on decision making theories that have been most prominent for DSS research and the most relevant to our focus on values.

One prevailing decision making theory in economics is rational choice theory (Coleman and Fararo, 1992). This theory describes rationality as acting to maximize personal benefits while minimizing costs. All models belonging to this theory assume people's choice of the best action according to consistent and stable preferences which reside in the decision maker's head. Other preference assumptions are completeness (alternatives can be ranked in order of preference) and transitivity (whenever A is preferred over B and B is preferred over C, then also A is preferred over C). Preferences are often expressed in utility functions.

Rational choice theory has been subject to criticism from behavioral economics and decision making, two areas of research that take into account psychological aspects of decision making. In reality decision problems are often ill-defined which means that the desired goal state is only clarified throughout the decision process and the decision makers are often not rational in the way assumed by rational choice theory. According to Payne et al. (1993) adaptive decision making model, people adopt strategies or heuristics to simplify the decision making process in complex situations with many alternatives. Such heuristics are often much more selective in the use of information from the decision context than economic models suggest and, thus, appear less rational.The authors state that based on cognitive ability and prior knowledge decision makers have several decision strategies at their disposal that they use in an adaptive manner to solve a decision problem with reduced cognitive effort. While choosing a strategy a decision maker follows four meta-goals, i.e. maximizing decision accuracy, reducing cognitive effort, minimize negative and maximize positive emotions and maximizing the ease of justifying their decisions.

Which strategy is used depends on the characteristics of the person, the decision problem (e.g. task variables, such as the number of alternatives or time pressure) and the social context (e.g. accountability to family members). Even small changes in the task environment may lead people to adapt their strategy. Payne and colleagues argue that this adaptation results from people's limitation of cognitive power to process

information in complex decision situations. Thus, while some strategies they identified like the weighted additive rule (which considers the values of each alternative on all attributes and relative weights of the attributes) come close to the rational model outlined above, other strategies such as lexicographic heuristics (which selects the alternative with the best value on the most important attribute) or the elimination-by-aspect heuristic (which eliminates alternatives based on predefined cut-off values for attributes) focus only on limited information of the decision context. Overall, strategies differ on several properties, e.g. whether they make trade-offs among attributes (compensatory) or not (non-compensatory), whether they use quantitative (e.g. summing of values) or qualitative reasoning, or whether information processing is consistent (i.e. the same for each alternative) or selective.

While this adaptive behavior can be seen as an intelligent way to deal with decision complexity by balancing cognitive effort and decision accuracy, it can also lead to problems. A non-compensatory strategy (e.g. lexicographic heuristic) can lead to an elimination of a potentially good alternative early in the decision process. Furthermore, it has been shown that adaptive behavior can lead to preference reversals or changes depending on how a task is stated, e.g. either as a choice or a matching response (Tversky and Slovic, 1988), or in which order elements of a choice set are considered (Tversky and Sattah, 1979). More focus on effects on preferences are discussed in the following subsection.

Within HCI, and to be specific human factors research, another strand of decision making research emerged in the 1980s as an opposition to rational choice theory, i.e. naturalistic decision making (Klein, 1997). The label "naturalistic" reflects that researchers in this field focused on decision making in real life situations. Often these situations are marked by time pressure, uncertainty, vague goals, high stakes, team and organizational constraints, changing conditions, and varying amounts of experience. One dominant model proposed by Klein (1997) is the recognition-primed decision model. This model suggests that people use their prior experience as a repertoire of patterns, that each new situation can be matched with to find a course of action. A distinction is made between this type of intuitive processing and analytical processing in the form of simulating how a course of action would play out in the current decision situation. While the balance between intuition and analysis is interesting, it seems to be a decision model more applicable to experts acting in highly complex, time sensitive environments (e.g. aviation) than to our current focus on supporting non-experts in taking major life decisions (which are less based on split second actions).

5.2.2 Preferences

As mentioned above there is a difference in assumptions about people's preferences being made by different theories. Preferences are statements about a desired condition on an attribute. Rational choice theory assumes that preferences are stable, consistent and known to the decision maker. However, as an implication of the opportunistic and task-based use of heuristics in the decision making process, preferences in the adaptive model are seen as constructive, i.e. attention to information and methods used to combine the information vary with different tasks. In this sense preferences can be seen as labile (Fischhoff and Lichtenstein, 1980). Besides psychological effects such as anchoring or framing elaborated in (Payne et al., 1993), other constructive processes of preferences have been suggested, e.g. the preferences-as-memory framework, which assumes that "decisions (or valuation judgments) are made by retrieving relevant knowledge (attitudes, attributes, previous preferences, episodes, or events) from memory in order to determine the best (or a good) action." (Weber and Johnson, 2006). In line with the ease of justifying a decision, one meta-goal of decision making, Simon et al. (2004) found in their experiments, that people's preferences changed in favor of a chosen option.

With respect to designing computer-supported decision support, it is important to take into consideration that preferences may be labile and subject to effects related to information display. We have tackled challenges related to preference construction in our earlier work on designing preference elicitation interfaces (Pommeranz et al., 2012). In that work the importance of values (or underlying interests) has already been pointed out.

5.2.3 Values

"The term 'values' has been used variously to refer to interests, pleasures, likes, preferences, duties, moral obligations, desires, wants, goals, needs, aversions and attractions, and many other kinds of selective orientations" (Williams, 1979, p. 16). This variety of terminology coming from different disciplines led to confusions about the nature of values. Accordingly, in rational, economic theories (e.g. as mentioned above) a distinction between preferences and values is unclear and often the two concepts are used interchangeably. However, especially in psychology, there is a clear distinction between preferences and values (Shiell et al., 1997). In this view values are defined over fundamental aspects of life. This includes, e.g., health, happiness or prestige. Values are generally seen as more stable than preferences.

Overall, and important for our work on decision support, "the view that values motivate and explain individual decision-making has [now] been widely accepted." (Cheng and Fleischmann, 2010) According to Keeney (1996), who proposed value-focused thinking, "values are fundamental to all we do; and thus, values should be the driving force for our decision-making." Unfortunately, this is not considered in the majority of DSS. Before we discuss the details of Keeney's proposed framework in the next section, we give a short definition of the concept of value that will be used throughout this paper.

One general distinction made in ethics literature, is between intrinsic and instrumental value. While an intrinsic value is the value an object carries in itself and is an end in itself, an instrumental value is a means to an end. For example, consider someone valuing happiness (intrinsic value), for whom the love (instrumental value) to his wife is a path to happiness. Another term used for intrinsic is terminal and was coined by social psychologist Rokeach (1973). In the Roakeach Value Survey he classified 18 terminal values (e.g. happiness, equality, freedom, social recognition) and 18 instrumental values (e.g. ambition, love, courage, honesty, independence). Schwartz and Bilsky (1990), in his effort towards a universal inventory of values, specified 56 basic human values grouped into 10 value types (achievement, benevolence, conformity, hedonism, power, security, self-direction, stimulation, tradition, universalism). Besides such classifications of values with respect to value relationships, Rescher (1969) has proposed five other principles as criteria for classifying values reflecting the many different perspectives one can use to approach the concept of values. One example showing the diversity of values, is the type of benefits at issue, according to which values can be classified into material and physical, economic, moral, social, political, aesthetic, religious (spiritual), intellectual, professional, and sentimental. For an in-depth discussion of value definitions, classifications and inventories, please see (Cheng and Fleischmann, 2010).

Borrowing the value definition of Cheng and Fleischmann (2010), which is based on other definitions including Roakeach's and Schwartz', we conceptualize values "as guiding principles of what people consider important in life." An important aspect to point out, is that values should not be seen in isolation. People have value systems consisting of a complicated web of values and as Schwartz (1996) pointed out, "attitudes and behavior are guided not by the priority given to a single value but by trade-offs among competing values that are implicated simultaneously." In particular, complex decisions may promote some values while violating others cherished by the decision maker. This leads to conflicts often avoided by people. Some aspects of our work were related to more narrow value inventories specific to the context.

5.2.4 Value-focused thinking

As a basis for our focus on explicating values relevant to a decision making context, we take Keeney's proposed framework of value-focused thinking Keeney and Raiffa (1993). A major aspect of this framework in line with our goal of empowering users to make informed decisions is its proactive stance. By suggesting to consider decision making as a creation of new opportunities rather than problem solving it puts the decision maker in control over the situation she has to face. To be more specific, Keeney posits that typical approaches to decision making, namely first concentrating on a given set of alternatives and then evaluating them according to one's values, leave the decision maker in a position which is merely reactive to a given situation. By actively approaching decision situations through, first, focusing on one's values and then choosing or creating alternatives that suit these values, the decision maker can instead channel her thinking efforts in order to achieve better decisions.

In detail, value-focused thinking proposes to make values explicit in the first stage of decision making. Simply listing values or objectives – statements of what one wants to achieve in a decision context – however, is not sufficient. Often people confuse ends with means. It is, therefore, important to identify which are the means objectives that ultimately lead to fundamental objectives, i.e. the ends one wants to achieve in a decision context. Important to note is, that means and ends are context dependent. If your decision context, e.g. concerns investing your available funds for retirement, the amount of money you receive at the start of your pension may be a fundamental objective. If you are, however, deciding how to achieve a good life during retirement, the amount of money may serve as a means objective.

Overall, value-focused thinking is structured into the following steps: (1) identifying objectives through hard thinking and creativity, (2) structuring objectives into mean and fundamental ones (3) creating alternatives beyond merely obvious ones by e.g. thinking about how to better achieve one's objectives and (4) evaluating in how far alternatives promote or trade-off fundamental objectives.

As Keeney acknowledges, articulating and revising values is difficult. People may feel that they are merely reflecting on their values as a philosophical exercise and may not see the immediate advantage with respect to decision making. However, over time they will gain experience and coherent value patterns will emerge that can be instantiated to many decision situations.

That value-focused thinking indeed leads to improvements in decision situations has been confirmed by Arvai and McDaniels (2001), who compared value-focused

thinking to alternative-focused thinking in group workshops dealing with management of risk in environmental decisions. The outcomes of their experiment showed that people in the value-focused condition felt more comfortable with their decisions and more satisfied that choices reflected their values. Furthermore, they considered a wider array of decision-relevant issues and felt more knowledgeable with regard to being able to make an informed decision. In line with Keeney's prediction of effort to think hard about values, the authors also report that participants in the value-focused condition considered the decision task more mentally difficult than participants in the alternative-focused condition.

To summarize, value-focused thinking promises several benefits to informed and thoughtful decision making. However, it requires substantial mental effort from decision makers to articulate and structure their values. People should, therefore, be supported in the process of thinking hard about their values as well as articulating them and relating them to a given decision context. Value-focused thinking suggests a dialog-based approach assessing one's (fundamental and means) objectives, i.e. one's values in a given context. Given the job domain this would be comparable to a career coach who asks many open questions to probe a client's values and support the client to understand her own values. While we are not planning to make such coaches obsolete by introducing a tool that can take over these tasks, we believe that a digital tool can provide additional help in value-reflection.

5.3 Our approach: A design stance aiming at user deliberation

Given the suggested and proven benefits of a value-focused approach to informed, thoughtful decision making and its applicability to taking complex and important life decisions, we use this approach as a basis for our digital decision support. In particular, we focus in the following on supporting people in value-reflection as this is the most difficult and effortful step in value-focused thinking. While Keeney suggests methods for assessing values which are strongly tied to dialog with experts, we focus on (additional) computer supported methods. Our main research question is, thus, 'How can we design digital tools to support people in reflecting on and assessing their values?' We aim to answer it by providing a set of design guidelines for such tools.

According to the user-centered design stance, which we follow in our work, we put the user in the focus of our design work. However, while in general we favor participatory design work, asking users early on in the design process about how a

value-reflection tool should be designed, did not seem to be an appropriate approach given that the literature suggests that values are abstract concepts, and that "there is an inherent difficulty in talking about values." (LeDantec et al., 2009) Therefore, the overall approach we took in the current work is rather design-led. By that we mean that the focus of our efforts lay on designing and implementing prototypes to convey our ideas about how value-reflection can be supported and to trigger feedback from users. This design process was continuously informed by expert knowledge provided by three counselors through interviews, brainstorming and feedback on the prototype. We would like to emphasize that the prototype is not to be seen as an early product to be iteratively improved through user feedback, but instead as a means to trigger users' critical deliberation and dialogic engagement with us as the designers and researchers. In order to trigger critical reflection some design choices we made (and discuss below) are rather provocative, e.g the choice of predefined values. Dialogic engagement was mainly fostered through the set-up of our user studies described in section 5.

Before we continue with the details of the prototype design, we define the concepts central to our work, i.e. self-reflection and awareness. Again, given that values are not naturally in the focus of people's thoughts and conversations, we believe that a digital tool should aim at reaching a state of awareness in the user, i.e. a user needs to become aware, on the one hand, that values are important to assess when making a decision and, on the other hand, what her values actually are and how they differ in importance. While the importance of values for decision making can be explained, getting the user to be aware of her values is the core of support we need to deliver to the user. Awareness is defined as "having or showing realization, perception or knowledge" (Merriam-Webster). Thus realizing or knowing one's values also leads to self-knowledge (i.e. the understanding of one's own capabilities, character, feelings, or motivations). This state of awareness can be achieved through self-reflection, which is the "examination of one's own thoughts and feelings" (Merriam-Webster). As stated by (Sengers et al., 2005), reflection is "bringing unconscious aspects of experience to conscious awareness, thereby making them available for conscious choice". In this sense bringing unconscious values to conscious awareness makes them available for value-focused decision making. According to Sas and Dix (2009) "reflection on experience has the potential to improve learning and practice, through enabling understandings gained from one's experience and consequently better future choices. Reflective skills, when properly applied can help people notice patterns of behavior (more or less effective), together with the underlying values and beliefs". In the following we will elaborate on how we designed a prototype that aims at supporting users in reflective practices to reach awareness of their values used in decisions.

5.4 Prototype Design

This section describes the design process of the so-called Reflections website which was used in several studies (section 5) to engage in dialog with users about how to support value-reflections with digital tools. Several experts were involved throughout the design process, most prominently in the early phase to provide insights into their practice to support people in value-reflection.

5.4.1 Expert Sessions

We conducted two semi-structured interview/brainstorm sessions, one with a job coach and one with two life counselors. The choice of participants was motivated by the fact that supporting people in reflection on their lives (or in the first case their career) is a major aspect of these experts' work practices. Participants were, first, briefed with the overall aim of our project (decision support) and the details of our current work (support for value-reflection). The remainder of each session was structured into three parts, (1) focusing on typical work practice and sharing of experiences of the experts, (2) focusing on methods used to support people's value-reflection and (3) brainstorming ideas for computer-supported value-reflection. Both interviews were audio-recorded and transcribed. To analyze the data two researchers annotated the data separately using the following predefined codes: cases (work example), question (expert questions used in coaching), method (specific methods and tools the experts use), aim (purpose of various aspects of the experts' work), assumption (underlying the experts' work), application (anything related to what a digital tool could or should do or be like). After separate coding the researchers discussed the data, in specific the few conflicts in the annotations, to reach a shared understanding of the themes discussed during the sessions. Summaries of our findings were sent back to the experts for feedback and validation. In the following we present the data relevant to the design of digital tools to support value-reflection including the dominant themes that arose from the discussions and an inventory of reflection methods used in practice.

Themes

The main theme that arose from the discussions was the *uniqueness* of the client. The experts described in depth how different every client and, thus, every session is. Therefore, the counselor has to adapt to each client by trying and using different methods to help the client reflect on his (work-)life. In this respect, the counselors were also skeptical towards digital tools as they felt that only a human could adapt

to the client in an optimal way. This aspect of uniqueness was also the result of our previous work on comparing methods for value elicitation (Pommeranz et al., 2011) which showed strong personal differences in ways people prefer (or are able to) express their values.

The job coach pointed out, that the exact methods to be employed with the clients may be less important, but the fact that they enter the reflection process on an emotional level is important. Thus, emotional triggers, such as art or poetry, are useful in making people reflect. Overall, however, the experts agreed that despite their uniqueness, most people need *guidance* to get from concrete reflection on experiences to more abstract values and the formulation of life goals. According to the experts it is hard for many people to think of abstract values. Therefore, the reflection process needs to be divided into several small steps. Specific questions of the form "Why is this important to you or what does this mean to you?" can be used to support reflection.

Another overarching theme that was brought up in both sessions was the role of *trust* between the counselor and the client. It is often difficult for people to open up and discuss their intimate experiences. This difficulty can be reduced through the built-up of a trusting relationship between the counselor and client and a setting producing comfort for the client. Part of this is the adaption to the unique features of the client as explained above. In addition, the experts in the first session emphasized that one should not judge the expressions of the clients and their interpretations. Thus, one should not point out that what a client says now contradicts something mentioned by the client previously. The experts labeled this aspect as *individual truth* held by each person. While something may seem contradictory to us, it may make sense for the client.

Another less prominent theme, but still discussed, was the role of *group* therapy. In the counselors' 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.

Reflection Methods

Experts use several methods to support people to reflect including visual, metaphorical or storytelling ones. Visual methods provide the clients with visual stimuli, e.g. pictures or paintings. The association card method is an example of this type. The counselor lets the client choose a card with an image from a set of so-called association cards (e.g. card sets used in psychological therapies or a set of images preselected by

the counselor) that appeals to her. This card is then used for reflection, starting with the concrete content, i.e. what is shown on the card to reasons for picking the card, experiences that the card triggers and their importance leading to more abstract values. The same reflection process can be used with other triggers, e.g. paintings (e.g. in a museum) or photographs. Metaphors are often used due to their figural nature, i.e. a person does not need to talk about personal aspects directly, but can pick figures that represent these aspects. One method mentioned by the experts is asking people: "if you had to describe yourself as an animal, which animal would you pick?"

Storytelling is an aspect of many methods, but can also be used as a starting point, asking clients to tell a story of a previous experience. Storytelling can also be triggered by the use of concrete questions from the counselor, e.g. "What does friendship mean to you?" According to the experts these questions should be formulated as starting with Why? or What? to trigger reflection instead of How-questions. For concrete links between reflection and values, one expert explained the use of value tables, i.e. lists of values given to the client to pick the ones relevant to the reflection discussed with the counselor.

5.4.2 Design of the Reflections website

The expert interview data informed the development of our prototype for value-reflection support that was used a means to communicate first design ideas and elicit user views. We would like to emphasize that the prototype is not necessarily meant as a first design to be iteratively improved to a final product. Instead, similar to technology probes (Hutchinson et al., 2003), we see the prototype as a trigger to foster deliberation and discussion among users and designers about how value-reflection can be supported with technology.

The prototype in form of a website was developed for the work domain, in particular to prepare for a job negotiation. The website's structure is based on a tab layout, including seven tabs: introduction, tools, reflections, values, preferences, competencies and friends. We chose this layout as it is common in browsers and other software. Furthermore, it provides guidance to the user (navigation from left to right tab), but without being too strict (as e.g. a wizard style interaction). Users can go forth and back as they wish. The introduction tab contains a text explaining the importance of value-reflection for constructing preferences, evaluating alternatives, and making a decision. In addition, it outlines the basic interaction with the prototype. The tools tab (Fig. 5.1 top left screen) offers several ways to reflect based on the expert sessions: association cards, storytelling, reflection questions, symbolic thinking and uploading

personal pictures. The interaction steps for each reflection tool are the same: When a user clicks on a tool icon (Fig. 5.1 (1)), a pop-up opens allowing the user to fill in a so-called *reflection* (to avoid misunderstandings we will continue using italics for the word reflection whenever we are referring to the concrete inputs in the prototype - as opposed to reflection as discussed in the paper elsewhere).

Once the means to reflect has been chosen (e.g. a photograph) or entered (e.g. a story), the user is asked to describe its content, what important things it does remind the user of, an emotion, and a value that is related to the reflection (Fig. 5.1, top right screen). This strict design has been chosen for two reasons, (1) due to practicality, as it simplifies saving to the database and processing the data in a later stage and (2) due to the experts' advice on supporting users to get from the concrete experience, e.g. what happened in a story or what do you see in a picture, to the abstract reflection on values related to the experience. Similar to value tables that the experts use to label written reflections of their clients with concrete values, we defined a list of work-related values in the system that was shown to the user in the last step of a *reflection*. We used the career anchors (functional competence, managerial competence, service, security/stability, entrepreneurial creativity, challenge, autonomy/independence and lifestyle) introduced by Schein (1990) as we considered them most fit for reflection on career choices. While using predefined values may limit a user's uniqueness, we took this choice consciously. On the one hand to probe user reaction to such a preset of values and on the other as it allows the system to give advice on preferences based on these given values.

Once a user saves a *reflection*, it is entered into a database (Fig. 5.1 (2)). All *reflections* can be reviewed on the reflections tab (Fig. 5.1, bottom left). At this point users can still add more values to their *reflections* or share them with other users (see friends tab). The same functionality is available on the values tab, however, the focus here lies on analyzing the frequency or reflecting on situations sharing a value. This is conveyed through a system-generated tag cloud of all values (Fig. 5.1 (3)) based on their frequency of occurrence in different user reflections. If a user adds e.g. lifestyle to the majority of *reflections* this will be the most prominent word in the tag cloud. The user can click on each value in the cloud to filter all reflections according to that value. As the website is meant for supporting job choices and negotiations, it also offers functionality with regard to creating a preference profile (preferences tab) and reflecting on core qualities (competencies tab). The preference construction interface (Fig. 5.1, bottom right) is adapted from our previous work (Pommeranz et al., 2012), and offers a holistic view including preferences, job offers, and a suggestion by the system based on the user's most prominent value (Fig. 5.1 (4)). In the future we aim

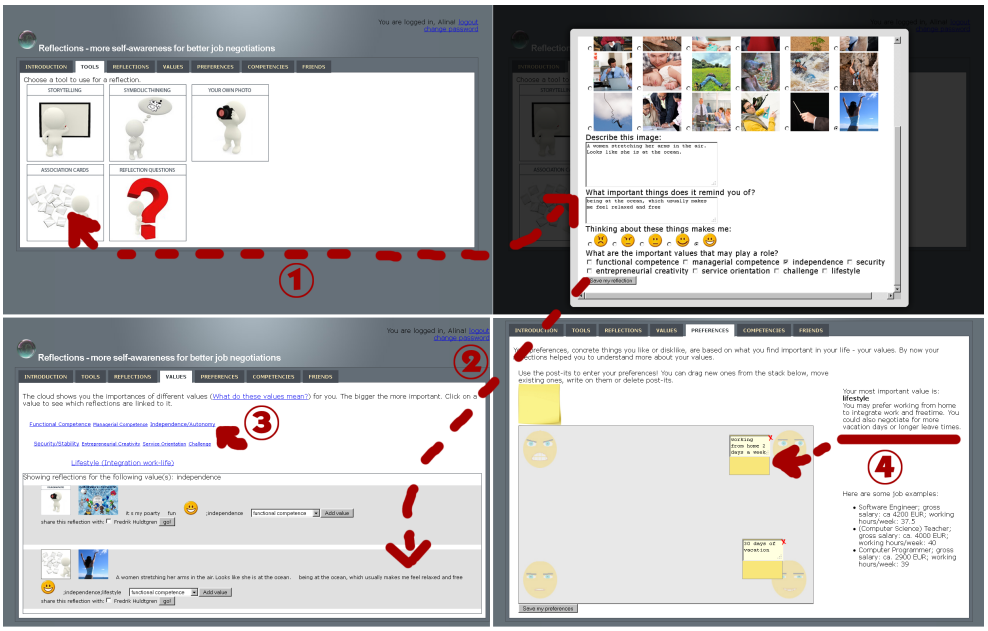


Figure 5.1: Sequence of screenshots to enter a reflection, look at the importance of values and enter preferences (the arrows are not part of the interface).

to build this advice on more complex webs of values. The competencies tab offers a simple way to reflect on one's core qualities following the core-quadrant method introduced by Ofman (2006). The last tab (friends) offers functionality to connect to other people using the system. Once befriended users can write direct messages and share *reflections* with each other. Based on the experts' experience that group discussions can help some people to get a deeper understanding of their values, we included this feature.

5.5 User studies

The prototype described in the previous section was used as trigger for communication and deliberation between users and the authors of this article of how value-reflection can be supported by technology. We conducted three sequential user studies. The first study aimed at getting direct feedback on the prototype as a way to engage in discussions of what is good or bad and how aspects of the design could be changed

to better adapt to the users. To support engagement between the designer of the prototype (first author) and the users the study was set-up in an interactive manner, which allowed the users to give direct feedback through a message function in the prototype (which is also part of its functionality as explained before) and the designer to reply in the same manner. The second study was a participatory design study with subset of the participants of the first study. In the set-up similar to Future Workshops (Kensing and Halskov-Madsen, 1991) participants were engaged in inventing new design solutions tailored better to their needs. The prototype was again used as a means for communication, in specific as a trigger for critique in the workshop's first phase. The workshop resulted in a set of considerations of import for designing technology to support value-reflection. These considerations were finally made explicit in design sketches and tested in an online survey with a large user sample.

5.5.1 Interactive User Study

Participants and Set-up

Eleven participants (3 female, 8 male) aged between 21 and 50 (M=30.4, SD=8.9) from six different countries were recruited through the university group's network to take part. Three participants dropped out due to a self-stated lack of time, eight people remained to finish the study. Each person used the Reflections prototype for a minimal period of one week and maximal 4 weeks. Participants were instructed to use the website and its mobile version as often as they want, but if possible at least once per day to enter so-called reflections, or think about your job preferences or competencies. They were further instructed to try all functions at least once. We asked for any kind of feedback, including bugs, aesthetic, functional and conceptual aspects. Participants were connected to the designer through the system's friend functionality to be able to send feedback messages while using the system. Another option was sending a traditional email to the designer. After using the prototype over the given time period participants were asked to fill in an online questionnaire consisting of the items shown in table 1.

Collected data and user feedback

Throughout the study we received feedback from seven participants: four messages through the system, six emails and one chat conversation reporting between one and five problems each. The majority of responses considered bugs (9) that were fixed directly. Three comments regarded aesthetic representation (e.g. of the tag cloud) and three comments were on a conceptual level regarding the understanding of values and

Table 5.1: Questionnaire items

	open questions
1	List the three most negative aspect(s).
2	List the three most positive aspect(s).
3	Do you have anything to add considering the usability of Reflections (website and/or mobile version)?
	Rated on a 7-point Likert scale (1=strongly disagree,7=strongly agree):
1	More awareness about my values/interests will be beneficial in a job negotiation.
2	More awareness about my preferences will be beneficial in a job negotiation.
3	More awareness about my competencies will be beneficial in a job negotiation.
4	The platform helped me to reflect on my values.
5	The platform helped me to reflect on my preferences.
6	The platform helped me to reflect on my competencies.
7	Creating reflections on-the-go with a mobile version is important for self-reflection.
	Multiple choice:
1	Imagine an upcoming job negotiation. After using the tool I feel ...self-confident, well-prepared, know exactly what I want, know what I already know.
	Open questions:
1	What changes do you suggest to be done to the platform, so that it helps you more or in a better way to reflect on your values, preferences or competencies?
2	Would you like to suggest other values that you were missing in the tool that are important in a job context?

attaching values to a *reflection*. The latter led to discussions between the users and the designer, but no changes were administered to avoid influencing the experience of other participants. A major aspect was the choice of work related values that were difficult to attach to *reflections*, which often focused on private aspects of life. One participant suggested having a layer using more personal values that are only later related to work values. In total we collected 47 *reflections* (one to seven per participant) 11 preferences and nine competencies. This shows that all participants used the tool to reflect on values, but less than half of the participants to reflect on preferences or competencies.

Eight participants (2 female, 6 male) filled in the questionnaire. As shown in figure 5.2, the majority of participants (except P8, who scored neutral) thought that "more awareness about my values will be beneficial in a job negotiation". Similar ratings were obtained for the benefit of awareness about competencies. Interestingly, two participants did not agree with the statement "more awareness about my preferences will be beneficial in a job negotiation." Given that many people focus on preferences in job negotiations, we were surprised that two participants did not consider awareness of preferences beneficial. However, it needs to be said that the study and the tool point out the importance of values and possible dangers of focusing to quickly on

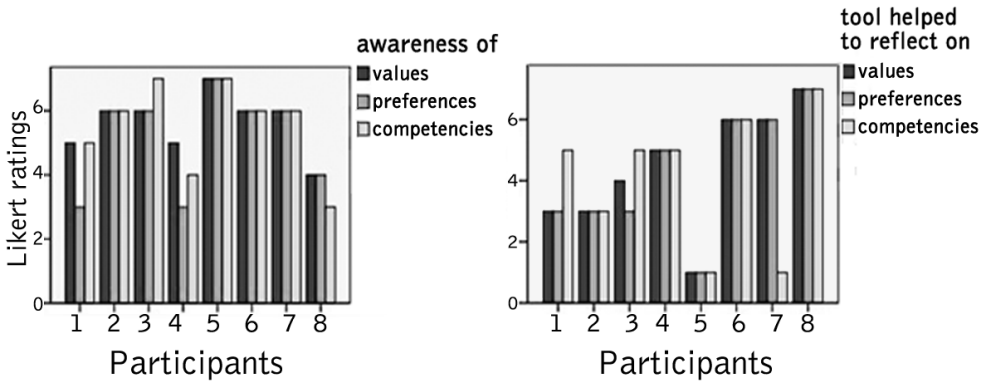


Figure 5.2: Likert scale (1-7) ratings per participant for benefit of awareness (on the left) and in how far the tool supported reflection (on the right).

preferences which may have influenced people's responses. Answers to statements relating to the extent to which the website helped them to reflect on their values, preferences and competencies were more diverse. Clearly, P5 did not believe, the tool was helpful in this respect. The rest of the participants were divided into a group of people that considered the tool helpful (P4, P6, P7 and P8), and a group that considered it less helpful, but saw some potential for reflection on competencies (P1 and P3). "Creating reflections on-the-go with a mobile version is important for self-reflection" was only agreed upon by two people. Considering the effect of using the website with regard to an upcoming job negotiation, three participants answered that after using it they feel self-confident, three claimed to know exactly what they want, one thought he would know what he already knows and one did not answer.

Table 2 gives an overview of the positive and negative aspects of the website as mentioned by the participants. Overall, from the positive comments we can see that participants liked the diversity of the reflection tools, i.e. that there are several ways to reflect. Furthermore, they said that the website was easy to use and learn. Major problems that were identified were the abstract nature of some tools, which leads to a lack of clear links between the (rather personal) *reflections* and work-values. Several participants mentioned that it was difficult to link a value to their *reflection*. A related aspect is that participants had trouble finding values that they hold and to understand the predefined ones. This was further elaborated by P1 in a general comment: "... maybe I want to add my own values and not the predetermined ones. When an interface prescribes certain steps (tell a story, tell what's important about

Table 5.2: Positive and negative aspects of Reflections

Participant	Positive	Negative
P1	1. It gives a way to capture daily reflections, which feels good - like you telling someone your deep thoughts and feelings which you might rarely do with actual human beings. 2. There are multiple ways to capture reflections. 3. It's easy to use.	1. List of values feels forced upon me and limited and doesn't always seem to match [the reflection] i choose one as the closest approximation of the values i feel are associated with a reflection, but that particular value might not feel like mine. 2. Unclear how preferences and competencies are connected to my reflections/value. 3. I'm missing a more free form tool, where i just collect some thoughts
P2	1. The tools are nice 2. It is a nice application for people without self-knowledge or self-reflection.	1. The faces do not look very natural 2. The website is not very user friendly yet and could have some nicer graphics and layout. 3. You cannot fill in 1 competence, it will not be saved.
P3	1. The core qualities are a very good idea, and are presented well. They make a good addition to the system. 2. I like the preferences page. Using a board to hang post-its on is a fun idea. Also, the tip on your most important value is a good idea. Maybe this could be expanded in the future. 3. The idea of making a link between your lifestyle and your work-style is a very good goal.	1. It needs to fill the gap between the tools (like a photo of my life) and the conclusions about your way of working. 2. The questions asked in the tools could be sharpened. They are standard and sometimes not relevant. 3. The system could introduce me more to the ... different kinds of values. Perhaps a new functionality could be added to make me explore these values more so I know what is relevant for me.
P4	1. quick and easy to learn 2. It kept giving me consistent important values 3. unique experience, I never used anything like this before that showed me what i care about most	1. It is still buggy 2. The values it listed did not include all the values i cared about. 3. It looks unpolished.
P5	1. Association cards: helpful, but still too abstract. 2. It is a good idea to make values and competences explicit.	1. [some tools] are too abstract. 2. Not all values i am seeking [are] there 3. Once i add a value i cannot remove it.
P6	1. Clear interface 2. Easy navigation 3. Many options	1. Definitions of some values not clear 2. Difficult to attach a value to certain activities (everything seems to be lifestyle). 3. Overview page is cluttered and difficult to read.
P7	1. More than one way to know yourself. 2. Consistency. 3. Seems to be a scientific approach.	1. A bit long instruction. 2. The size of the pop-up window does not fit the content. 3. No delete function after submit
P8	1. It help to you to think about things that you aren't used to. 2. You can see you hard points easily. 3. You can see all your reflections together.	1. You don't obtain any feedback of your own reflexions. 2. It's difficult to see the relation between the work and your reflections. 3. I don't think that a photo or an story can help you.

it, select some values) in some ways I feel that I'm not able to express myself as I see fit - it feels like I'm behaving how someone else wants me to behave." Other general comments with regard to the guidance of the system were made by P2: "At present, the system more or less prompts you to make up your values, preferences and competencies by yourself and then input them. I believe the system needs to become one that actually helps you discover them." and P5: "the tools are a good start, but need better guidance." P8 takes it even a step further by saying "when you chose a photo or a story the platform should tell you which value or competencies are you talking about." Additional job-related values mentioned by the participants were: cooperation, participation, contribution, empathy, motivation, innovation, change, flexibility, independence, respect, appreciation, money, pleasure, teamwork, being around other people, friendliness and, generally, social aspects of work.

In summary, we can conclude that participants generally believe in the benefit of awareness of their values, preferences and competencies and that a computerized tool could lead them to know what they want and start a job negotiation self-confidently. However, the concrete implementation needs further improvement, and first of all more thought. While many ways to reflect are appreciated, more guidance could be provided by the system to link *reflections* to personal values. To get more in depth insights we conducted a participatory workshop with a subset of the participants, as explained next.

5.5.2 Future workshop

To gather qualitative feedback and creative ideas for a tool to support people's value-reflections we held a 1.5 hour brainstorm session inspired by Future Workshops (Kensing and Halskov-Madsen, 1991), a Scandinavian method for IT design supporting participatory engagement.

Participants and Set-up

We invited all 11 participants of the interactive user study as well as two experts from the previous study to the session. Four participants (no experts, all male) participated. We used the Reflections prototype as a trigger to gather feedback on and problems with value-reflection. Furthermore, we used materials, such as colored pens, post-its and paper, for noting critiques and ideas as well as sketching new design proposals. Borrowing from the Future Workshop method, our session had three main parts (1) a critique phase, (2) a fantasy phase and (3) an implementation phase. We structured the phases so that the largest time portion was reserved for the generation of ideas. In

a short briefing, the method was introduced; steps and rules were written down on a white-board which were visible during the whole session.

Critique phase (25 minutes): Participants were instructed to critique the Reflections website on any possible level (bugs, aesthetic, functional or conceptual). Each participant was restricted to 30 seconds talking time per critique statement. The short speaking time allowed every participant to voice himself and not be ignored or interrupted by a dominant person. The statements were then written on post-its and put up on a wall labeled 'problems'. Last, post-its were grouped into categories by the participants.

Fantasy phase (40 minutes): To overcome difficulties of non-designers to be creative we started with the generation of metaphors, inspired by Kensing and Halskov-Madsen (1991). For instance, besides being a website, the prototype could also be seen as a diary. Metaphors help people think outside the box and see artifacts from new perspectives. Next, participants were given time to sketch ideas on paper. To open up for creativity participants were instructed to come up with utopian ideas and did not have to think about technical constraints. After about 15 minutes of sketching, we started an open brainstorm similar to the first phase. People stated ideas in 30 seconds and noted them on post-its that were stuck on the wall labeled 'ideas'. Three out of four participants actually created sketches of designs, while one participant wrote down his ideas in words. The participants all shared their ideas. After collecting ideas each participant had three votes (green stickers that could be stuck directly on the ideas) for their three most preferred ideas.

Implementation phase (15 minutes): We ended the session with a discussion of practicability and implementation of different ideas.

Results

Problems identified in the first phase were categorized by the participants into technical, graphical, user experience and concept of reflection. The latter two categories often overlapped. Overall, more problems related to the concept of reflection and user experience than technical and graphical problems. Examples were lack of motivation to use the tool, too much time is needed to use the tool, unclear preference input, gap between work and life values, lack of guidance, lack of private feel, and no option to add own values. The metaphor generation in the second phase resulted in: therapy, meditation, consultant, career advisor, diary, stress relief, dream, conflict resolver and dating site. These metaphors highlight the different functions a value-reflection

tool could fulfill. While consultant/advisor suggest an expert role in the sense that the tool advises the user, dream, diary or meditation are personal ways to reflect and therapy or stress relief hint to a medical perspective. Participants used the metaphors as inspiration for the sketches without being instructed to do so. Ideas ranged from concrete design suggestions (mind-mapping as a reflection tool, scrapbook or diary with handwriting for personal feel) created during the sketching to utopian ideas, e.g. a pill dispenser (for therapy) and abstract thoughts, e.g. on adding life goals and links to values. In the last phase it was agreed upon that development should focus on providing simple, but highly usable functionality, as a bad implementation could hinder people to use the tool at all.

Besides the concrete ideas that were mentioned several other design considerations were the focus of discussion. These included the user's motivation, personality, privacy, guidance and advice. Personality was mentioned as an important aspect of a value-reflection tool. The participants discussed how important it is that the tool provides a personal and secure feeling. Similar to a diary that you use to note intimate feelings or experiences. One participant suggested a diary-like interface. Furthermore, participants liked that the website provided several ways to reflect which allowed to pick the one that suits the user's personality best.

A discussion arose regarding the free-form diary style and guidance that should be provided by the system. Participants mentioned that guidance is needed to get from the concrete (images, stories) to the abstract (values). One participant thought it would be a motivational factor, as with less guidance it can be unclear for people what the benefit is of using such a website. However, other participants thought that too much of it can also impede the use as reflecting is an activity that is more free and personal. Too much structure would feel like the system's values would be imposed on the user. Overall, all participants agreed that reflections should first be tagged with personal values, that a user could add to the system, but these then had to be matched to the work-values predefined in the system. The latter were, however, hard for participants to grasp and definitions would be needed. One participant suggested a system-led dialog with the user to give the user a deeper understanding about a certain value and find out whether this work-value relates to the user.

Guidance is also related to considerations of individual truth and in how far the system should judge its users. This aspect was identified by the experts previously. While the participants did not use the same notation, they discussed in how far the system should provide new insights to the user, to surprise the user and make her learn something new about herself. A designer needs to consider to what extent the system

would make assumptions about a user's input and in how far the system could still promote an individual truth.

In any case participants preferred the system to have a certain level of transparency. To achieve transparency they suggested that the system offers more explanations, e.g. for the benefits of each reflection tool (e.g. why should I write a story?). Especially, for the rather general reflection questions participants were unsure how reflecting on those would give them insights to work-related values. Thus, one suggestion was to have more personal and context-sensitive questions. In addition, one participant suggested to give immediate feedback from the system on the importance of the different values while putting in a *reflection*. This idea was sketched by a participant who suggested to have a pie chart in which each piece represents a value which grows as soon as the value is added to a *reflection*. This sketch was voted as one of the best ideas in the workshop (remember that each participant had three votes).

Last, participants discussed an integration of the website with social networks like Facebook. One participant suggested that besides inviting Facebook friends to Reflections, profile information and even status updates could be used to create a value-reflection profile automatically. Other participants preferred to keep Reflections a private tool that does not connect to other websites. This was also seen as a matter of trust. The more private and secure the tool feels the more a user can trust it, and as someone may want to enter intimate reflections, trust is very important.

In summary, several considerations were discussed and often they intertwined. These considerations do not yet provide clear guidance to a designer of such tools. While discussions between the four participants were insightful and provided first ideas for concrete implementations, we still needed to test whether the ideas would hold for a larger population. Therefore, we conducted the online survey described in the following section to test several concrete design ideas.

5.6 Online Survey

The themes from the expert sessions as well as the discussions during the user studies presented in the previous sections led to important design considerations. However, as they were extracted from a small sample of people, it still had to be seen in how far they would be generalizable over a larger part of the population and how they could be translated to concrete designs. To this end we created an online survey to test different design ideas and relate them to people's personal characteristics of self-reflection as well as their attitudes towards awareness of values and value-reflection tools.

5.6.1 Participants

We recruited participants through our personal networks and further snowball sampling. From 119 people who approached the study, 82 completed the survey, 35 female, 45 male and two with unknown gender. As our to be designed website is mainly aimed at non-expert decision makers to make job choices, we tried to recruit young people that are about to or have recently finished their education. Respondents were aged between 22 and 64 (M=31.71, STD = 7.095), which shows the majority to be what we call young professionals (ca. 25-35 years old). The majority of respondents, ca.75%, held a university degree. Participants came from 13 different countries, the majority from the Netherlands (30), the United States of America (15), Germany (11), China (7) and Sweden (6). All 82 participants were included in the analysis.

5.6.2 Materials: Design sketches

Based on the input of our expert and user studies we created a set of design sketches to be tested with a large sample. As not all themes/considerations were easily representable in sketches, in particular trust and emotional triggers, we decided to test these with a set of statements to be rated on Likert scales (see part 3B in Appendix B). We decided to create two opposing versions for each design idea in order to ask people to make a clear choice. We are aware that people may have more nuanced preferences and may, e.g., prefer a little bit of guidance over no guidance or too much guidance. However, when implementing interfaces clear choices have to made. And while a system can be adaptable, such an adaption would still be discrete and not continuous. In the survey we emphasized that we are showing extremes of each idea in order to understand what users may prefer. The concrete implementations may differ from the sketches. In total, we created 10 sketch pairs each representing one design consideration with two extreme opposite ways to implement. All sketches are to be found in Appendix A. Table 3 gives an overview of the design pairs, including the theme/consideration it represents, and the two extreme options.

To ensure that each pair represented the theme we wanted to test well enough, we asked 10 people to rate the extent to which this was the case. In particular, we presented each person with the 10 pairs together with the theme the pair was supposed to represent and asked them to rate on a 7-point Likert scale how far the contrast given in theme was represented. The first five people we asked were previously involved in the studies (either as an expert (2) or participant of the user studies (3)). This was motivated by the fact that our representations were created based on these people's input. The second five were explicitly chosen as not to be related to our work and

Table 5.3: Overview of design sketches

Design pair (theme)	design A	design B
1 (uniqueness)	same look & feel for all users	personal look & feel (customizable)
2 (uniqueness)	only given values	possibility to add new values
3 (uniqueness, guidance)	structured form-style	free-form (scrapbook/ diary-style)
4 (guidance)	no explanation for a tool	explanation for each tool
5 (uniqueness)	general reflection questions	personal reflection questions
6 (guidance)	system shows inputs (user can adjust them if(s)he thinks something is inconsistent)	system points out possible inconsistencies of inputs
7 (transparency)	value chart on separate tab after entering reflections	value chart adjusts while entering reflections
8 (guidance)	definitions for values given in the system	find out the meaning of a given value and whether it fits you through questions
9 (guidance)	link a reflection directly to predefined work-values (for job choices)	link a reflection first to personal values & then to predefined work-values (for job choices)
10 (guidance)	overview of reflections without explanation	overview of reflections with explanation

Table 5.4: Validation results

design pair	M (rating)	STD	Design changes administered in accordance with the raters
1	6.2	0.79	
2	6.2	0.79	Button highlighted
3	6.3	0.82	
4	5.9	0.88	
5	4.1	1.60	The set of general questions was considered personal, and was therefore adjusted to be more general.
6	6.2	0.79	
7	5.2	1.55	Reflections on the tab layout were removed as they were confusing
8	5.8	1.48	Highlighting added to make dialog more clear.
9	6.2	0.79	
10	5.4	1.33	highlighting of the important part in the interface

with different backgrounds (in terms of education level and expertise) to reflect the target group of the study. As shown in table 4, the mean ratings were all above the neutral point of the scale and standard deviations were low. Small changes were administered to the sketches in accordance with the raters to improve the distinction between designs A and B.

5.6.3 Set-up

The survey was divided into four parts: (1) demographic information including age, gender, level of education, occupation and frequency of writing a diary, (2) reflection-rumination questionnaire (RRQ) (Trapnell and Campbell, 1999), (3) design ideas for value-reflection (separated into binary choice of sketches (3a) and a set of questions probing links to social networks, emotional triggers and trust (3b)), and (4) a set of statements probing respondent's attitudes towards self-reflection, awareness and decision making.

Besides testing the sketches and other design considerations, we hypothesized that there maybe differences in what type of interfaces people prefer depending on their current level of self-reflection, e.g. that more reflective people may want less guidance. Therefore, we used the RRQ in part 2 to retrieve a score for people's level of reflection and rumination (a less healthy and often insecure way to think about oneself over and over.). The survey was implemented as an online website. In the third part, we presented the sketch pairs together with their titles (see table 2). The sketches were shown from first to last in the same order for all participants, however, the position of sketch A and B was randomized between left and right. Respondents could select the preferred sketch by clicking on it. They could change their selection until they proceeded to the following sketch by clicking a next button. For each sketch pair respondents had the possibility to enter a comment.

Part 3b consisted of a question about the integration with Facebook (see appendix A) as well as 10 items aimed at measuring other design concepts suggested by the experts such as promoting individual truth (item3.1-item3.3), emotional triggers (item3.4-item3.7) and the role of trust (item3.8-item3.10).

Part 4 consisted of 11 items, aimed at measuring the constructs understanding of the value concept (item4.1 and item4.2), self-knowledge of values and their relation to decision making (item4.3-item4.6), perception of benefits of value awareness (item4.7 and item4.8), attitude towards a digital value-reflection tool (item4.9 and item4.10) and perceived effect of a tool (item4.11).

The survey was first checked by two experts (one on value-reflection and one researcher) to ensure face validity. We then ran a pilot test with six participants to ensure everything was working and comprehensible.

5.6.4 Data Analysis and Results

Measurements of Constructs

To measure the reliability of constructs uniqueness and guidance which were obtained as dichotomous data from the sketches we calculated Cronbach's alpha. In both cases the alpha values were very low (.180 and .405 respectively) suggesting that the items do not measure the same construct. Therefore, we decided to regard the 10 design pairs separately in the following analysis. For part 3a, Cronbach's alpha for the construct individual truth was .684,when leaving out item3.2. Thus for further analysis we used an aggregated measure (item3.1 and item3.3). Alpha values obtained for the constructs of emotional triggers (.472) and the role of trust (.477) were too low to suggest reliable measures of the same constructs. For the emotional triggers this is not surprising as people may use different means (e.g. visual, audio etc.) to reflect. In the further analysis we present this data as separate entities.

For part 4 acceptable alpha values were obtained for the multi-item constructs understanding of the value concept (.636), value relation to decision making (.837), perception of benefits of value awareness (.672) and attitude towards a digital value-reflection tool (.790).

User characteristics and attitudes

From an analysis of correlations between the different user characteristics measured, we found a significant positive correlation, $r(80) = .272$, $p < .05$, between the frequency of writing a diary and the level of self-reflection (obtained from the RRQ). Already predicted by the experts, emotional triggers play a big role in people's reflection. As shown in figure 5.3 all mean values are clearly above the neutral point of the scale (3). Especially photographs, songs, and artworks can inspire reflection. The lower value on diary entries may be related to the fact that not every participant writes diaries on a regular basis. Furthermore, we found a significant negative correlation, $r(80) = .223$, $p < .01$, between the age of the participants and their level of rumination, which suggests that young people ruminate more than older people. This could also explain the significant positive correlation, $r(80) = .315$, $p < .01$, we found between age and the construct of individual truth. Whereas older people prefer more to the promotion of their individual truth, younger people prefer that the system teaches them something new about their values.

With regard to the constructs measured in part 4 of the survey we found a significant positive correlation, $r(80) = .322$, $p < .01$, between the level of understanding of the

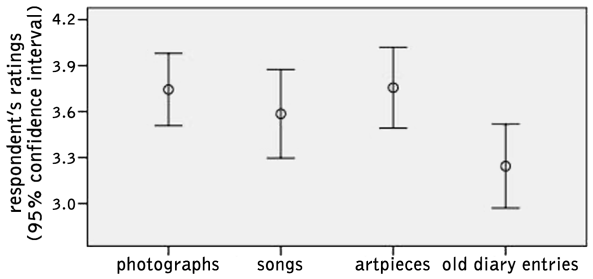


Figure 5.3: Average ratings of agreement of use of different emotional triggers for reflection.

value concept and people's understanding of the relation of values to decision making. In addition, we found a significant positive correlation, $r(80) = .442$, $p < .01$, between the perceived benefits of value awareness for career choices and a positive attitude towards using a value-reflection tool.

Attitude towards computer supported value-reflection

As shown in figure 5.4, the majority of respondents (more than 50%) would feel well-prepared for an upcoming job negotiation after using a value-reflection tool. Almost a third of the respondents would (also) be more self-confident and know exactly what they want. However, at the same time about a third of the participants think they would know what they already know. In addition, an analysis of responses to items 4.9 and 4.10 showed that 40.2 % would use a digital tool to self-reflect and 56.1 % believe, it would help many people make better decisions. In both cases about a third of the respondents were indifferent, which leaves a minority of respondents with negative attitudes towards computer supported value-reflection. The indifference of some respondents may be based on the fact that without having actually used the tool it is hard to answer this hypothetical question. We think, that using such a tool would help each user to form a directed (positive or negative) opinion. However, this overall positive attitude supports the motivation of our work and the creation of digital tools for self-reflection.

User preferences for designs

As shown in figure 5.5, clear preferences were found for design idea 2, 5, 7 and 8 (see table 2 for descriptions). In particular, almost 80 % of the respondents prefer to be

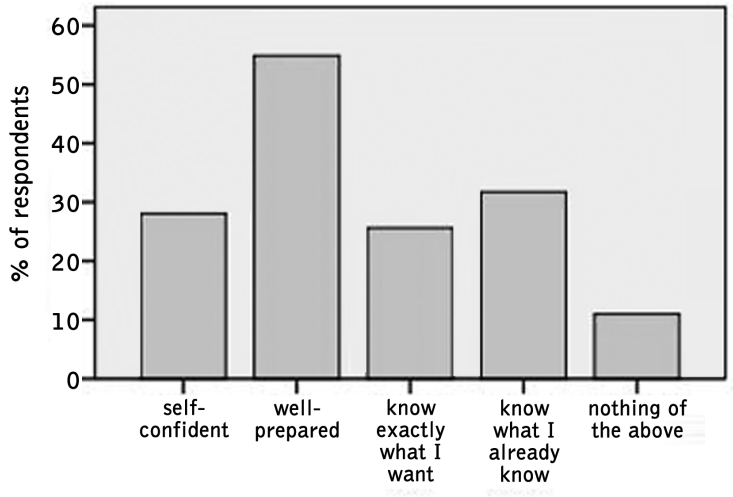


Figure 5.4: Percentages of respondents for each level of preparedness after using the tool.

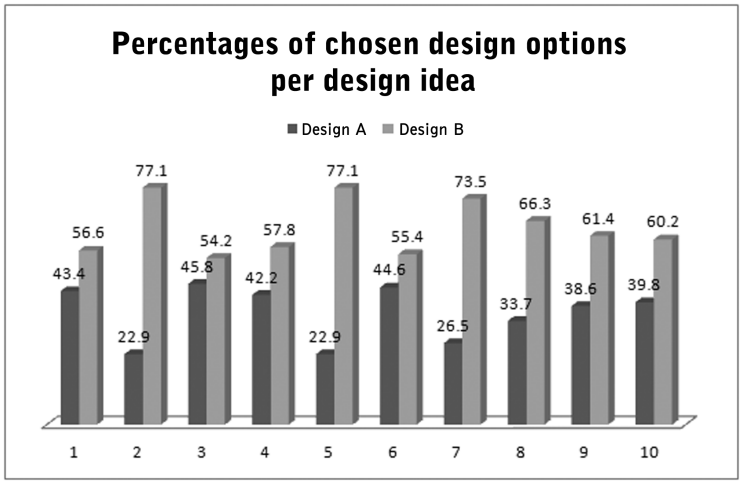


Figure 5.5: For each design idea (1-10, see table above) the percentages of people who chose design A (left column) or design B (right column) is shown.

able to add their own values to the system and about 80 % prefer personal questions for reflection to more general ones. Over 70 % of the respondents would like to receive immediate, visible feedback about their value frequencies (i.e. how often they reflect on each value) while filling in new a *reflection*. A correlation analysis between the level of self-reflection and the different design options further revealed a significant negative correlation, r(80) = -.329, $p < .01$, between self-reflection and design 7, meaning that less self-reflective people prefer immediate feedback more, while for highly self-reflective people the information could be shown in separate screens. Given that our tool aims at making people more reflective over time, such a preference should be considered in the design of the tool. About 66 % of the respondents prefer a dialog with the system (or a coach in the system) to understand the definitions of predefined values and whether they fit to them.

To investigate if we can predict any of the design choices based on personal traits we used a binary logistic regression (forward stepwise) analysis to predict each design pair with age, gender, reflection, rumination, perceived self-knowledge of values, understanding of the value concept and relation between values and decision making as covariates. For design pair 5 (general vs. personal reflection questions) gender and people's understanding of the relation of values to decision making were significant ($p < .05$) variables predicting design choice. About 95% of all females chose design B (personal reflection questions), whereas only 70 % of the males chose this design. In addition, the more people understand how their values relate to decision making, the more they prefer general questions. For design pair 7 (separate or combined visualization of reflection input and value importance) level of reflection (as explained above) and people's understanding of the relation of values to decision making were significant ($p < .05$) variables predicting design choice. In a separate correlation analysis of the latter construct and design 7 only a very weak correlation was found. Thus, this factor can be neglected. For design pair 8 frequency of diary writing was significant ($p < .05$) in the prediction model. A positive correlation, r(80) = .251, $p < .05$, was found, meaning that people who write their diary more frequently prefer a dialog approach to finding out the meaning of a value and whether the value fits them.

User preferences for different levels of integration with Facebook are shown in figure 5.6. More than half the users would not want to integrate Reflections with Facebook (or other social network) data. One fourth of the respondents would, however, use it for importing their Facebook profile to Reflections and about a fifth would use it to create automatic *reflections* based on their status updates or to connect to their friends. From the participant's comments it becomes clear that integration possibilities should be provided, but users want to be able to opt out. One participant

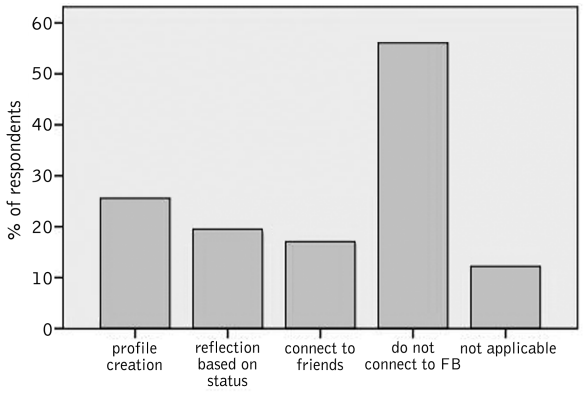


Figure 5.6: Percentage of people who chose each Facebook integration option.

stated “linking to Facebook should be optional, with multiple privacy options (settings to share your activity with others, or ability to only share with certain people, or linked but kept completely private-only accessing your status etc. for your personal insight...).” Another said, “I’d note that Facebook integration is probably important for some of the intended users. That said, I would not want to use it. I should be able to opt out.”

User comments

Participants were able to enter comments for each design sketch and after having completed part 3.Two participants commented on the usefulness of customizing the tool to a personal look & feel (sketch pair 1). P48 said, “The question is - is this a tool solely for accomplishing its goal: reflect and learn about yourself, or is it a place for one to sit down, reflect but also enjoy the process of reflection? In the latter case customization would be more useful, as it is also an indirect way to reflect yourself. However, as reflection is something not really tied to a tool, and can be done without it, I see little incentive to spend time customizing the look and feel in cases when you would use the tool only occasionally.” P77 pointed to the importance of the content. “Though customizable sounds good, I guess what’s important is the content (the pictures/notes you uploaded). I’d think a fancy look and feel would distract me from the content.” Several participants commented on design choice 3 that they would like to be able to get explanations when they ask for it, but pop-ups or other forced types of explanations should be avoided. E.g. “Explanation is good, but only if I ask for it please.”(P10) Similar comments were made on the explanation offered for the

overview (design choice 10), e.g. “I might want to see it the first time but have the option to dismiss and/or never show again.” (P47) Another related aspect was the way the system judges the input of a user (design choice 6). Similarly to the idea of individual truth mentioned by the experts, one participant pointed out, that “part of reflection includes also conflicting thoughts and emotions that can be perfectly valid - but it is up to user to decide what does and what does not make sense, not for the tool to divide reports into premade boxes of possible and impossible value combinations.” (P48) Although we agree with this view, over 40 % of the respondents liked that the system pointed out inconsistencies in the values linked to a *reflection*.

With regard to the level of personality of the questions, four participants reported that they would prefer a mix of general and personal questions, e.g. P47: “I chose personal because I think it will elicit more concrete writing, but I think overall I’d prefer a mix of both types of questions.”

Overall comments from the users referred to the importance of trust, e.g. as P1 stated, “it has to feel intimate and trustworthy”. Another respondent commented that, “Trust is a big issue, not just about privacy but also in the guidance the system provides. It also changes what people enter in the system. Some systems use a humanoid avatar designed as a person the target user can relate to.” (P17) The importance of trust to users is also reflected in the answers to item3.8. (“I only share reflections on myself with people I trust. The same is true for a digital system.”), which was agreed to by 79.3% of the participants (54.9% strongly agree).

Two additional ideas for functionality mentioned in the overall comments were the “evaluation of my past critical decisions” (P14) and allowing “others who can access my page to rate my values from their perspectives.(What they think my values are).” (P23) One participant raised the concern of a possible lack of continuous motivation to use the tool and mentioned several interesting questions for further research:“What would promote continued use of such a system? Beyond distinguishing what values a person currently holds important, why would such a system help a person develop more positive values, or values that help the person grow or adapt? Is that an objective, or would the social pressure of wanting to evolve past one’s current state be of value to a system such as this?” (P34) While our focus is more on identifying existing values to make better decisions, these questions are relevant for the general design of value-reflection tools.

5.7 Discussion

5.7.1 Design Guidelines

Based on the work presented in this paper we compiled a list of the following five design guidelines. Based on the survey we suggest concrete design ideas where possible.

GL1: Consider the uniqueness of the user by offering means for personalization. The theme of uniqueness was first brought up by the expert counselors during the interviews and was referred to as people using unique ways to reflect. Thus one important aspect of value-reflection tools is to offer many different ways to reflect and leave it open to the user to choose the way that suits her best. That users liked the option to choose from several reflection tools was confirmed in the first user study. Further, it was found that users consider a personal feel important. Ideas that were brought up were being able to customize the tool, adding one's own values and having a more diary or scrapbook style approach where users can create a personal (art-)piece of reflection. The survey confirmed that there is a strong preference among the large user group to be able to add their own values. One participant explained that "trying to fit very personal values in predefined boxes seems very forceful, and even condescending, dismissive of the values that might be very important, but are not on the list." (P48) Preferences for the other two design ideas were distributed among the two design options. Thus, we can conclude that customization and diary style is based on a user's personal preferences. They could be used as additional functionality in a value-reflection tool, but focus needs to be on providing several ways to reflect and the option to add personal values to the tool. In addition, regarding reflection questions, a mix of personal and general questions could be offered for males, while females could be provided mainly with personal questions, as more than 90 % preferred this type of questions.

GL2: Aim for a trustworthy design of the tool through careful implementation of privacy, dialog with the user and transparency. Reflecting on personal experiences and values is an intimate action, and thus when shared with others a level of trust needs to be established first. The counselors emphasized the importance of trust between counselor and client and a comfortable atmosphere that is necessary for the creation of trust. That this aspect can be transferred also to computerized systems was confirmed by the participants of the survey (ca. 80% agreed on the importance of trust to a value-reflection tool). From the participants' comments (in the first study and also the survey) we learned that the establishment of trust is based on the feeling

the user has when using the tool as well as the privacy and the type of interaction a tool provides. Especially when implementing group functionality, it is important to ensure that the user can set for each *reflection* who is allowed to see it or keep it completely private. We also propose that increasing a trustworthy feel could be achieved through transparency and user-system dialog. One example of transparency would be the immediate feedback from the system when a user enters a *reflection*, as this visualizes to the user how the system's value profile was adapted in real time. In the survey more than 70% of the respondents were in favor of this option. In addition, the majority of respondents favored a dialog with the system to clarify a value and see whether the value suits them.

GL3: Consider adjustable levels of guidance to get from concrete reflections to abstract values. The level of guidance that the system offers to the user was a theme that was mentioned first in the expert interviews and was a prominent consideration emerging from the discussions in the user workshop. While the experts deem it an important aspect of helping people to get from concrete experiences to abstract value concepts, similarly, the users emphasized that it is important for them to understand how to get from a *reflection* to the related work value. While we hypothesized that the level of preferred guidance could be dependent on a user's level of reflection or rumination this could not be confirmed through the survey. However, we found that younger respondents tended to prefer that the system teaches them something new about their values, thus, guides them in their exploration of values. Furthermore, although not the majority, a substantial amount of respondents liked the system to point out inconsistencies ($>$40%) and a structured approach to entering *reflections* ($>$45%). While our data did not provide clear preferences with respect to guidance, the diversity shows that balancing guidance with the open nature of reflection is difficult and needs to be considered carefully. Furthermore, several respondents stated that explanations should be available on demand. Based on these findings we propose to implement different levels of guidance in a value-reflection tool which can be set by each user at runtime.

GL4: Use emotional triggers to enter a reflection process. Mentioned first by the experts and confirmed by the survey, people use emotional triggers to begin a reflective process. Concrete examples are visual stimuli such as preselected (e.g. value-laden) images, paintings or personally owned photographs, audio stimuli such as music, written pieces, e.g. old diary entries or poems, or other art pieces. In line with GL1 it is important to design a tool so that each user can select her personally preferred trigger. Whereas some people like writing or looking at art, others reflect through listening to a song or the lyrics of a song. In any case, as supported by the

experts, emotional triggers are more useful than asking direct questions about people's values. Due to their abstract nature, many people have difficulties answering such questions.

GL5: Integrating a value-reflection tool with social network functionality should be optional and nuanced to allow for privacy. The survey revealed a clear division of opinion towards the use of social networks together with a value-reflection tool. A bit more than half of the respondents clearly stated that they would not want such an integration at all and about 10 % of the respondents do not use social networks. The remaining people preferred different levels of integration from just importing friend connections to using status updates as *reflections*. Based on this data we propose that designers provide functionality for integration with social networks, but leave it completely up to the user whether to use it or on what level. In addition, a social function could also be implemented in the value-tool itself (as in our prototype), and, e.g., provide functionality for assessment and discussion of other user's values. Again, users should be able to set the privacy level of such functionality by themselves.

5.7.2 Limitations

Overall, we believe that the studies showed a number of factors that are relevant for the design of value-reflection tools. However, the overarching themes – as represented in the set of guidelines – were mainly derived from the workshop held with a small set of participants. While this small number of people allowed us to consider ideas in more depth, it may have limited us in finding all relevant factors for such tools. As a result we cannot claim the list of guidelines to be exhaustive. Several other workshops or in-depth user interviews could reveal more factors that may be of relevance. Another limitation was that the design sketches presented in the survey were all static and thus the exact interaction was up for the respondent to imagine. We do not believe, this led to any major problems as the focus lay more on testing extreme opposites instead of concrete implementations. However, animated examples of ideas or even a new set of interactive prototypes could provide more nuanced user preferences. Therefore, future design guided by our work needs to test different implementations iteratively. Furthermore, we have focused on value-reflection in the presented work. Further work is needed to investigate how the retrieved value profiles can be related to concrete decisions.

5.8 Related work in HCI

Several strands of research relate to our work on self-reflection and awareness. While the importance of reflection has long been acknowledged in education, most prominent research on self-reflection within HCI has been done in the area of affective systems building on Boehner et al. (2005) interactional model of emotion. An example in this area is the Affective Diary (Lindström et al., 2006), which augments traditional diary keeping with sensor technologies. While such systems lead the user to gain emotional awareness, self-reflection can lead to awareness on different levels. Sas and Dix stated in their 2009 workshop abstract on "Designing for reflection on experience" that "self-reflection can be focused on thoughts and behaviors, and in particular on the identification of values, beliefs and assumptions that motivate such behaviors." (Sas and Dix, 2009) However, contributions at this workshop lacked the specific focus on reflection on values and decision making that we aim at in our work.

Recent trends in HCI on monitoring and improving people's behavior, which are more related to decision making in life, are persuasive systems (Torning and Oinas-Kukkonen., 2009) and personal informatics (Li and Forlizzi, 2010). The latter has a strong focus on supporting self-knowledge through collecting personal data and analyzing it. Therefore, in its focus on gaining self-knowledge and awareness personal informatics systems are similar to our approach. However, first, they do rely heavily on quantitative data often even sensed automatically through sensors and second, the goal is not always improved decision making. Take e.g. the quantified self movement (http://quantifiedself.com/) with many participants who are simply curious about their own data but do not use it explicitly for given decision situations.

The focus of persuasive systems lies less on self-knowledge, but motivating people to change their behavior. These systems focus often on choices regarding one's health or environmental choices. While the focus on (behavioral) choices relates to our work, we do not agree with the approach taken in persuasive systems. It seems that these systems mainly embed the designer's values and notions of what good and bad behavior is. Often it is even unclear whether these values are made explicit to the users and whether the user can then critique or adapt them. In our view, this approach does not lead to knowledge about how decisions may affect the user's values and what consequences the decisions have. Little emphasis is placed on self-reflection and awareness of one's own values.

Critique on persuasive systems has also come from Purpura et al. (2011), who aim at provoking "discussion of the conceptual and ethical limits of persuasive computing".

"When designers make decisions about the 'one right way' that should drive suggestions to influence the 'flawed' user, it removes agency from the individual." They pose the question whether such an approach to technology design is respectful of the user. Similar questions are also focus of Reflective Design introduced by Sengers et al. (2005). They argue that "reflection on unconscious values embedded in computing and the practices that it supports can and should be a core principle of technology design." In this sense reflective design is an overall approach to technology design and it appeals to both designers and users to reflect critically on the technology they build/use. In our work we borrow merely parts of the principles and strategies to apply them to decision support, e.g. openness of a decision support system to allow for user appropriation and skepticism with regard to decision advice.

Another related research agenda in HCI is that of Slow Technology, coined by Hallnäs and Redström (2001). This area has recently gained more attention as an opposition to our fast-past way of living augmented through ubiquitous computing leaving people with an omnipresent need to be constantly efficient and connected to others through technology. Slow technology aims to balance these aspects of people's lives with moments for reflection, mental rest and solitude. One could view our design efforts as part of this agenda in respect to slowing down the decision making process, taking time to reflect on one's values and considering alternatives in depth. Slow technology can, however, also be seen as promoting long-term thinking similar to Friedman and Nathan (2010) call for a multi-lifespan perspective of information technology. While our work is not aimed at considering multiple generations in the design, decision making supported through technology for value-reflection promotes thinking about the longer term consequences of a decision instead of fast, in-the-moment choices.

Besides these links to other HCI works, we believe our work is highly relevant to value sensitive design in the sense that it delivers tools that can help people assess and express their values, which is a first step to finding design trade-offs that manage all stakeholders' value considerations.

5.9 Conclusions and future research

According to the best of our knowledge there are currently no digital tools dedicated to value-reflection at this moment. Furthermore, decision support systems research has not yet focused on the integrating value-reflection in existing tools. However, we

[1] http://www.willodom.com/slowtechnology/

argue based on Keeney's model of value-focused thinking that supporting people in value-reflections, in particular with mobile tools, is an important direction to enhance people's decision making on major life choices. We have presented our ongoing design work including the development of a value-reflection prototype and several design sketches based on expert and user feedback. Designs have been validated with a large sample in an online survey. Based on our results we have compiled a set of five design guidelines to be used by other researchers/designers who intend to create value-reflection tools. We also contribute concrete design ideas of how the guidelines can be implemented.

Designing digital tools for value-reflection is a new area to investigate, which may not only be adding to decision support but also to HCI research strands of value sensitive design or slow technology. Therefore, we believe our work is a first contribution that other researchers of HCI and related fields can build upon. Future work should include the design and implementation of concrete tools to be tested in longitudinal studies. It has to be seen over time in how far these tools can motivate users to achieve awareness, enhance decision making or lead to behavioral changes.

5.10 Bibliography

Arvai JL, Gregory R, McDaniels TL (2001) Testing a structured decision approach: Value-focused thinking for deliberative risk communication. Risk Analysis 21(6):1065–76.

Boehner K,Depaula R, Dourish P, Sengers P (2005) Affect: from information to interaction. In Proceedings of the 4th decennial conference on Critical computing: between sense and sensibility, CC '05. ACM, New York, NY. 59–68.

Carenini G, Poole D (2002) Constructed preferences and value-focused thinking: Implications for AI research on preference elicitation. Technical report, American Association for Artificial Intelligence.

Cheng A-S, Fleischmann KR (2010) Developing a meta-inventory of human values. In Proceedings of the American Society for Information Science and Technology 47(1).

Coleman JS, Fararo TJ (1992) Rational Choice Theory: Advocacy and Critique, chapter Introduction. Newbury Park: Sage.

Fischhoff SP, Slovic P, Lichtenstein S.(1980) Cognitive processes in choice and decision behavior. Knowing what you want: Measuring labile values,Hillsdale, NJ:Erlbaum, 117–141.

Friedman B, Nathan LP (2010) Multi-lifespan information system design: a research initiative for the hci community. In Proceedings of the 28th international conference on Human factors in computing systems, CHI '10, 2243–2246.

Hallnäs L, Redström J (2001) Slow technology – designing for reflection. Personal Ubiquitous Computing 5(3):201–212.

Hindriks K, Jonker C (2008) Creating human-machine synergy in negotiation support systems: Towards the pocket negotiator. In Proceedings of the First International Working Conference on Human Factors and Computational Models in Negotiation, HuCom 2008, Delft, The Netherlands, 47–54.

Hutchinson H, Mackay W, Westerlund B, Bederson BB, Druin A, Plaisant C, Beaudouin-Lafon M, Conversy S, Evans H, Hansen H, Roussel N, Eiderbäck B (2003) Technology probes: inspiring design for and with families. In Proceedings of the SIGCHI conference on Human factors in computing systems, CHI '03, Fort Lauderdale, FL, 17–24.

Keeney R (1992) Value-Focused Thinking: A Path to Creative Decision Making. Harvard University Press.

Keeney R (1996) Value-focused thinking: Identifying decision opportunities and creating alternatives.European Joumal of Operational Research 92:537.

Keeney R, Raiffa H (1993) Decision with Multiple Objectives: Preference and Value Tradeoffs. Cambridge University Press.

Kensing F, Halskov-Madsen (1991) Generating visions: Future workshops and metaphorical design. In Design at Work: Cooperative Design of Computer Systems. L awrence Erlbaum, Hillsdale , N.J, 155–168.

Klein G (1997) Decision Making Under Stress: Emerging Themes and Applications, chapter The Current Status of the Naturalistic Decision Making Framework, Aldershot, UK: Ashgate Publishing Ltd.

LeDantec CA, Poole ES, Wyche SP (2009) Values as lived experience: evolving value sensitive design in support of value discovery. In Proceedings of the SIGCHI

conference on Human factors in computing systems,CHI'09, Boston, MA, 1141–1150.

Li I, Forlizzi ADJ (2010) A stage-based model of personal informatics systems. In Proceedings of the SIGCHI conference on Human factors in computing systems,CHI'10, Atlanta, GA, 557–566.

Lindström M, Stahl A, Sundström P, Höök K, Laaksolahti J, Combetto MJ, Taylor A, Bresin R (2006) Affective diary: designing for bodily expressiveness and self-reflection. In CHI '06 extended abstracts on Human factors in computing systems, Quebec, Canada, 1037–1042.

Ofman D (2006)Bezieling en kwaliteit in organisaties. Servire.

Payne J, Bettman J, Johnson E (1993) The Adaptive Decision Maker. Cambridge University Press.

Pommeranz A, Detweiler C, Wiggers P, Jonker CM (2011) Elicitation of situated values: need for tools to help stakeholders and designers to reflect and communicate. Ethics and Information Technology, 1–19.

Pommeranz A, Broekens JD, Wiggers P, Brinkman W-P, Jonker CM (2012) Designing interfaces for explicit preference elicitation: a user-centered investigation of preference representation and elicitation process. User Modeling and User-Adapted Interaction 22(4-5):357–397.

Purpura S, Schwanda V, Williams K, Stubler W, Sengers P (2011) Fit4life: the design of a persuasive technology promoting healthy behavior and ideal weight. In Proceedings of the 2011 annual conference on Human factors in computing systems, CHI '11, Vancouver, Canada, 423–432.

Rescher N (1969) Introduction to value theory. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Rokeach M (1973) The nature of human values, New York: Free Press3.

Sas C, Dix A (2009) Designing for reflection on experience. In Proceedings of the 27th international conference extended abstracts on Human factors in computing systems, CHI EA '09, Boston, MA, 4741–4744.

Schein EH (1990) Career Anchors (discovering your real values). Jossey-Bass Pfeiffer, San Francisco.

Schwartz S (1996) Value priorities and behavior: Applying a theory of integrated value systems. The Psychology of Values: The Ontario Symposium, 8. Hillsdale, NJ: Lawrence Erlbaum.

Schwartz SH, 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.

Sengers P, Boehner K, David S, Kaye J (2005) Reflective design. In Proceedings of the 4th decennial conference on Critical computing: between sense and sensibility (CC '05), ACM, New York, NY, USA, 49–58.

Shiell A, Hawe P, Seymor J (1997) Values and preferences are not necessarily the same. Health Economics 6:515.

Simon D, Krawczyk DC, Holyoak KJ (2004) Construction of preferences by constraint satisfaction. Psychological Science, 15: 331.

Torning K, Oinas-Kukkonen H (2009) Persuasive system design: state of the art and future directions. In Proceedings of the 4th International Conference on Persuasive Technology, PERSUASIVE '09, Claremont , CA, 30.

Trapnel P, Campbell J (1999) Private self-consciousness and the five-factor model of personality: distinguishing rumination from reflection. Journal of Personality and Social Psychology 76:284.

Tversky A, Sattah S (1979) Preference trees. Technical rept. no. 1, Stanford Univ. Calif. Dept. of Psychology.

Tversky A, Sattath S, Slovic P (1988) Contigent weighing in judgement and choice. Psychological Review 95, 371.

Weber EU, Johnson EJ (2006) Constructing Preferences from Memory. Cambridge University Press.

Williams RMJ (1979) Understanding human values, chapter Change and stability in values and value systems: A sociological perspective. New York: The Free Press, 15–46.

CHAPTER 6

METHODOLOGICAL REFLECTION

The previous chapters have dealt with concrete designs of user interfaces for preference construction (chapter 4) and value-reflection (chapter 5) and thus tackled the 'what' aspect of designing human-centered decision support. This leaves us with the second focus of design, i.e. the 'how'. While we have already touched upon our design processes in the previous chapters, we will give a more detailed account of methodological considerations and reflect back on the design processes in an analytic manner. How to conduct cooperative design and engage participants into creative processes is a remaining question within HCI. *In particular, with technology becoming more complex and being used in diverse situations affecting many different stakeholders, cooperative design is of growing importance. In this chapter we first give some background in user-centered approaches to technology design and in particular cooperative design (section 2). Next, we elaborate on our compositional design method (section 3) before we go into details of the observations made in design sessions described in previous chapters (section 4) and other design work related to this thesis (section 5 and 6). We conclude with recommendations for conducting cooperative design sessions.* [1]

[1] This chapter is based on the following articles: Alina Pommeranz, Pascal Wiggers and Catholijn Jonker. Towards compositional design and evaluation of preference elicitation interfaces. HCD'11 Proceedings of the 2nd international conference on Human centered design, Springer, July 2011, pp. 586-596 and Alina Pommeranz, Ulas Ulgen and Catholijn Jonker. Exploration of facilitation, materials and group composition in participatory design sessions. European Conference on Cognitive Ergonomics (ECCE'12), Edinburgh, UK, August 2012

Imagination is everything. It is the preview
of life's coming attractions.
Albert Einstein (1879-1955)

6.1 Introduction

With technology becoming more complex and weaved into every aspect of our lives, design of computer systems and foreseeing their impact has become a challenging task. Unlike early expert systems and tools to accomplish work tasks efficiently, systems today are used by a wide variety of people and in many circumstances. Given the unpredictable use, user types and needs of end-users, it is difficult for designers to anticipate system requirements. Therefore, technology design processes are in need of end-user involvement more than ever. User-centred design (UCD) (Norman, 1988) offers diverse ways to involve users ranging from informative methods where users are consulted during requirements elicitation and usability testing to participative methods where users act as co-designers in creative processes (Abras et al, 2004). The latter methods come from cooperative or participatory design (PD) (Kensing and Blomberg, 1998) and co-design (Sanders and Westerlund, 2011). Since the early days of cooperative design many methods have been developed to engage end-users and other stakeholders in the design process. While these are useful, one of the remaining challenges is that end-users often feel that they have insufficient knowledge or are not creative (Sanders and Westerlund, 2011). Indeed, end-users may lack domain knowledge or design skills. Empowering people inexperienced in technology or design to engage in creative processes is, therefore, the focus of participatory design work and an open problem in PD (Kensing and Blomberg, 1998). In our work, we are continuously involving designers, domain experts and end-users in the design of new interfaces. Besides practically 'doing' user-centered design, we investigate how to conduct participatory design sessions including promoting engagement between researchers/designers and participants and triggering creative output. In particular, we explored set-up and facilitation of design workshops, materials to support creativity and group compositions.

In this chapter we, first, present more background on cooperative design. Subsequently, we describe observations from the design cases presented earlier in this thesis and two additional cases of projects related to this thesis. Overall, we varied on set-up, materials used and group compositions in design workshops and reflect back on how these aspects did or did not support our goal of user engagement and triggering creativity. This as a first step towards a more analytic investigation of cooperative design sessions that is needed to advance theory on supporting participants in cooperative design.

6.2 Background

6.2.1 User-centered and cooperative design approaches

UCD approaches (Abras et al, 2004) commonly involve users as informants and testers, e.g. to elicit domain knowledge and needs through interviews or in usability tests. This engagement is one-directional and emphasizes the role of the designer as the sole creator of the technology that, although informed and tested, is imposed on the user in its final form. To design human-centered systems through more bi-directional and creative engagement we have to turn to cooperative approaches, e.g. participatory design (PD). Due to its historical scope PD has led to methods for envisioning a future (Kensing and Madsen, 1992) involving changes in the technical, social and political environment in which they are situated. As our work largely omits the political aspects of PD we use the general term cooperative design in the following. However, the methods we use were partially adapted from PD. Similarly recent trends in the design of software systems have started to use PD methods based on the belief that "active user involvement in the software development process leads to more useful and usable software products" (O'Neill, 2000). Another new approach is co-design (Sanders and Westerlund, 2011) which focuses less on the organizational context and more on services and products in general. This creativity-based approach to engaging stakeholders introduces the notion of co-design spaces. Co-design spaces can refer to the physical design space a team works in, the space constituted by participant activities and the future solutions being developed. Our work focuses on the second aspect, i.e. creative participant activities.

6.2.2 Aspects of participation in cooperative design

User Empowerment

As PD arose from a movement towards emancipation of workers in Scandinavia in the 1970s, user empowerment became its central theme. Empowerment is enunciated in current PD research in diverse ways, including, among others, empowering specific user groups, enabling direct democracy on social and political matters and strengthening the users' position in design processes (Ertner et al, 2010). Other enunciations are targeted at the role of the researcher/facilitator (see next subsection). We focus on the position of the user and other stakeholders in cooperative design activities.

Roles in cooperative design

With a shift from UCD to cooperative approaches the roles of researchers, users and designers have shifted as well (Sanders and Stappers, 2008). The user has become a co-designer and the researcher has become a facilitator, providing guidance and tools to the users to make them co-designers. Professional designers are also still needed to provide their expert design knowledge. In some cases the researcher can also take the role of designer. However, according to Ertner et al (2010), "the PD researcher's practice is guided by unconscious assumptions and socially specific knowledge, which become reproduced and embedded in methods, categories and interpretations. By this the practitioner poses a risk of dominating the users, if they neglect to focus explicitly on deconstructing the tacit aspects of their own practice." Due to this risk the role of the facilitator should be given special focus when setting up cooperative design sessions.

Creativity triggers

Co-design was defined as "the creativity of designers and people not trained in design working together in the design development process." (Sanders and Stappers, 2008) These authors explicate four levels of creativity: creating, making, adapting and doing. Becoming co-designers requires a high level of passion and knowledge in a certain domain and "it can be difficult to get people to create ideas when they feel that they have insufficient knowledge and [...] people who are brought into co-designing experiences may feel that they are not creative" (Sanders and Westerlund, 2011). Sanders and Westerlund (ibid) suggest using ambiguous visual artifacts as creativity triggers.

In addition, Sanders et al (2010) have recently proposed a framework for categorizing tools and techniques used in PD, which they see as an important starting point to compare, discuss and make choices about the different tools and techniques. In their framework they distinguish between form (i.e. the kind of action between participants, e.g. making things), purpose (i.e. probing participants, priming them to immerse them in the domain of interest, understanding their current experience or generating ideas) and context (i.e. group composition, face-2-face or online, venue and stakeholder relationships). In our own work we focus mainly on making things, i.e. paper prototypes, with the purpose of generating ideas. Aspects of the context, such as group composition and relationships are also discussed in our work.

Tools and techniques for the purpose of making tangible things are among others

2-D collages or 3-D mock-ups. Diaries are used with individuals for all other purposes to tell or explain things. Furthermore, techniques related to enacting, e.g. games, improvisation and envisioning the future (e.g. (Kensing and Madsen, 1992)) are less used to probe participants, but more to immerse them and create new ideas collaboratively. Except diaries all tools and techniques can be used in group sessions.

Other examples of tools are prototypes ranging in fidelity from sketchy paper versions to fully functional systems or scenarios used to outline specific use cases or reflect on the effects of the system on different stakeholders (Nathan et al, 2010). Marois et al (2010) compared an introductory game, interactive illustrations (semi-functional prototypes) and storyboards (with static screenshots) as starting points for design activities with stakeholders. Storyboards made participants grasp the topic and supported creativity. Interactive illustrations engaged participants more, but were also more biasing. Artifacts are not only important creativity triggers for single participants, but also serve as boundary objects, i.e. supporting communication.

In the context of involving populations that are harder to reach, e.g. due to spatial restrictions or ability to engage, Gaver and colleagues developed cultural probes. "Cultural probes are designed objects, physical packets containing open-ended, provocative and oblique tasks to support early participant engagement with the design process [...] As data trickled in, the cultural probes inspired design responses used to foray into the design space." (Gaver et al, 2010). However, it is left up to the designer how to use probe materials to inspire designs. Similarly, technology probes (Hutchinson et al, 2003) have been used to create design spaces by collecting people's experiences and reflections on technical objects.

In our work we used mainly prototypes, either digital or paper versions, and single interface elements. The Reflections prototype described in the previous chapter comes close to the idea of a technology probe. In the following we focus mainly on artifacts used in cooperative group sessions with the aim to create paper prototypes. A concrete method we developed and used in some of our workshops is described next.

6.3 Towards compositional user-prototyping

In our early work on designing interfaces, specifically for preference elicitation, we found that existing interfaces were not supporting the users' cognitive characteristics well. Therefore new interfaces had to be developed to address this gap. However, clear guidelines for the design of such interfaces were lacking and useful literature and design ideas came from such diverse fields as psychology or recommender systems.

Furthermore, designing the one perfect interface to support a cognitive activity of the user, e.g. constructing a preference profile, to be used in many applications seemed impossible as such cognitive activities are highly influenced by the current use context and the user's cognitive characteristics. This means that for different user groups and applications (e.g. recommender systems, DSS, NSS) the interfaces would be varied in order to yield the best results (e.g. an accurate preference model). Starting to design new interfaces from scratch, however, costs valuable time and money. To reduce costs we propose the following approach.

In an attempt to combine knowledge from different fields (psychology, behavioral decision making, recommender systems) and integrate it with participatory design approaches we developed a compositional approach to design and evaluation of (preference elicitation) interfaces (Pommeranz et al, 2011). The underlying assumption is that similar to component-based software engineering (Heineman and Councill, 2001) elements that have been used successfully in previous designs can be reused across design contexts. To combine elements in a sensible and efficient manner one has to identify characteristics of the use context (for decision making e.g. task goal, importance of the decision, available time, the number of alternatives and people involved) and the user characteristics and be able to select or create elements fitting these aspects. This knowledge is not readily available in every design context, but can be accumulated over time in order to support designers in their future endeavors.

Overall, we suggest using existing elements (or elements created according to guidelines derived from the literature) in participatory design sessions. The designer/researcher can combine elements in a number of prototypes to show how each element acts in context of a complete interface. Participants are asked to inspect the prototypes but specifically focus on evaluating each element and create ideas of how to combine elements from several prototypes and add elements that are missing so far. In this way participants' reluctance to engage in creative activities can be mitigated, as they can use existing elements as starting points. In addition, due to the focus on elements and their composition one can test a large number of elements with a handful of interfaces.

In the following we give an outline of the proposed compositional approach.

6.3.1 Interface elements

The core of our approach form what we call interface elements. Generally, an interface can be seen holistically or as a collection of elements. Simple elements (widgets) like

buttons, list boxes, and comboboxes, fit all cognitive styles. These can be embedded in more complex elements, e.g. a spell checker in a word processor. Complex elements often represent a concept or an idea rather than just a simple way of entering data. Complex elements are meaningful in terms of a particular user goal, cognitive style and fit with a particular environment. We focus on combining complex elements. The ultimate goal is to develop and store many elements in a database with information on how each element relates to contextual factors, styles, and design principles. This database should contain only evaluated elements and be available to interface builders to speed up the design process.

6.3.2 The approach step-by-step

Step1: Defining the use context

The starting point of our compositional approach is the definition of the application's use context, i.e. in decision support systems: the general purpose or task goal, e.g. recommendation or personalized decision advice, the number of people involved, time constraints, the size of outcome space and the decision importance.

Step 2: Determining relevant cognitive styles of target users

To assure optimal adaptation to the user's way of handling information the next step is to determine the cognitive styles that are relevant in the given use context. Next, the designer needs to define whether the target user group consists of people with all styles in the chosen set of styles or whether to focus on a specific subgroup that shares one style. In the latter case it is enough to design one interface. Otherwise, it is possible to design one interface for each style or an adaptive interface covering all styles and can be adapted by the user.

Step 3: Selecting/Creating interface elements

Given a database with usable interface elements and their relationships to the use context factors and cognitive styles, this step consists of querying the database to return all elements fitting the use context defined in step 1 and the style defined in step 2. In case the query does not return enough elements to cover all styles, the designer has to create new elements. In case these elements end-up in the final design due to positive evaluation these elements will be added to the database. New elements can be created by the designer or in cooperative design sessions with users. In the latter the facilitator needs to assure that relevant design principles are considered in

the final designs. For the design sessions standard creative techniques can be applied (brainstorming, thinking hats, etc.).

Step 4: Composing interface prototypes

The set of elements is combined into complete interfaces. Ideally, all possible combinations should be developed and evaluated by the target users to find the optimal one. However, with a high number of elements this is an impossible endeavor. Therefore, the designer should pick a subset of all possible interfaces that covers the relevant design principles and styles.

Step 5: Compositional evaluation

As the core of the approach is the compositionality of elements, we evaluate the interfaces compositionally. That means the goal of the evaluation should not be to find the best interface in the designed set but to ask users to evaluate the different elements and offer ideas of how to combine them. In the compositional evaluation the participants interact with all interim interfaces. We suggest a formative evaluation (e.g. think aloud) to encourage discussion of ideas and constructive feedback. This step can be part of or preceding a cooperative design session.

Step 6: Composing new interface

After participants interacted with the interfaces the user feedback needs to be applied to the design of a new set of interfaces. In cooperative design sessions participants combine the existing elements in a new way and add new ones if needed. To understand the design rationale of new combinations, it is beneficial to split participants in several groups and have them present their final designs to each other. One consideration regarding cognitive styles is to group participants with the same characteristics in a group and thus develop an interface for each style. In the final system the interface could then be adapted to the current user.

Step 7: Usability Testing and Optimization

The last step is an iteration of standard usability testing with target users to refine the end-design(s) in the context of the real application.

6.3.3 Application of the approach

The approach was developed and applied in parts of the design case presented in chapter 4 of this thesis. A more detailed description of how the approach was instantiated in that work has been published in (Pommeranz et al, 2011). In two additional projects, we have used adapted and improved versions of the general approach outlined above. In the following we describe observations from participatory design work presented in earlier chapters (4 and 5) and the two other projects that have followed the compositional approach to interface prototyping involving a number of participatory design sessions.

6.4 Case 1: Preference elicitation revisited

In Chapter 4 we have already described in detail how we conducted cooperative design sessions with end-users to design a preference elicitation interface. We have focused on the outcomes of the sessions with the goal to create guidelines for designing such interfaces. Here, we discuss the methodological observations. First, we quickly recall the set-up of the session.

To design a preference elicitation interface we chose to set up a collaborative prototyping workshop with eight participants (5 male, 3 female) explicitly chosen to have mixed backgrounds including design and technology. The workshop consisted of two parts, a group discussion and participatory design session to create paper prototypes. We used four digital prototypes as creative triggers, which represented four extreme versions based on different interaction styles. To make sure each participant could form an opinion about these prototypes, each participant used each prototype for 10 minutes before the workshop.

6.4.1 Material

We created paper versions of all interface elements we had used in the four digital prototypes, e.g. a virtual agent, post-its, so-called value charts or a tag cloud, as well as standard interface elements such as text fields, checkboxes, sliders, etc. This was inspired by the PICTIVE technique introduced by Muller (1991). Additionally, we had a number of blank papers, pens and scissors to give participants the chance to create their own interface elements. These materials were used by the participants in the second part of the workshop to design their own preference elicitation interfaces.

6.4.2 Procedure

After a short briefing, we started a general discussion about the different prototypes and interface elements. The discussion with the whole group took about 20 minutes. After that we split all participants into two groups of four participants. Each group was provided with the same set of materials described above and instructed to use the material to create their own version of a preference elicitation interface. They were encouraged not only to combine the elements existing in the four presented prototypes but also create new ones. This part of the creative session was planned for about 30 minutes. However, since both groups were not done within that time frame, the session was extended to 1 hour. The session ended with a presentation of the two groups' results to each other to understand their design rationales.

6.4.3 Observations

We found that the discussion in the beginning of the workshop was lively, although participation was not evenly balanced. While some participants were more dominant others resided to the background. As participants had all interacted with the prototypes they had clear ideas of what elements of each prototype they liked or disliked. In the discussion new ideas for combining elements from several prototypes already emerged and were concretized. During the prototyping session we noticed that participants focused mainly on existing elements (although we encouraged them to create their own new ones) and the outcomes of the two groups were very similar to each other and to combinations that had been discussed in the beginning. Thus, we can conclude that while the digital prototypes were useful in engaging the participants easily, the results of the prototyping were strongly influenced by the prototypes and less innovative than expected.

6.5 Case 2: Value-reflection revisited

The second design case described in chapter 5 of this thesis focused on designing tools for people to reflect on their personal values. To understand how people would want a value-reflection tool to be and how they would use it, we conducted a workshop with four participants inspired by the Future Workshop method (Kensing and Madsen, 1992). As we found values to be a difficult concept to grasp for people we decided to first engage with experts (life counselors) and built a digital prototype –called Reflections– based on guidelines established together with the experts. This prototype was used by the participants for up to four weeks during which they were in active

dialog with the researchers. In the workshop the prototype was used as a starting point for critique. This design case did not follow the compositional approach outlined above, but still revealed some interesting considerations with regard to methodological choices in design workshops with end-users.

6.5.1 Material

The Reflections prototype was used as a starting point of our cooperative design work. It was a fully functional website allowing participants to enter different types of reflections (stories, images, questions etc.) that we adapted from the experts' practice. Furthermore, functionality for analyzing values and sharing reflections with friends was provided. In addition, we provided paper and colored pens for sketching design ideas.

6.5.2 Procedure

Goal of this 1.5 hour workshop was to collect ideas for tools that help people in reflecting on their values. Participants could sketch ideas but concrete prototypes were not aimed at. After a short briefing the workshop was structured into critique, fantasy and implementation phases, in which participants shared their feedback briefly (30 sec) and then placed it on a designated wall. Speaking time was limited to allow for fair participation. As a starting point for the critique phase we used the Reflections prototype. At the start of the fantasy phase we brainstormed metaphors for the prototype with the whole group and then allowed 15 minutes of individual sketching before sharing ideas. After the workshop we talked to two participants to receive feedback.

6.5.3 Observations

In the use period of the prototype some users engaged in longer discussions about conceptual problems and their possible solutions. In the workshop, the prototype was a successful trigger for critique. All participants had prepared several critical points and more were triggered through the engagement with others. While the critique went smoothly we expected the fantasy phase to be more difficult for non-designers. However, especially the metaphor generation triggered creativity and consideration of many perspectives (e.g. meditation, consultant, diary, dream, or conflict resolver). Participants used metaphors as inspiration for their sketches without being instructed to. Providing pen and paper triggered participants to create concrete designs; some as sketches, some as textual notes. This shows how open-ended triggers

can be appropriated by participants according to their own skills. All participants engaged equally and enthusiastically in the workshop and did not mention doubts of lacking ability to contribute. The enthusiasm, however, may also be ascribed to the voluntary nature of participation in the workshop. Still, one could expect that even initially motivated participants feel less encouraged during the workshop due to group composition or facilitation. This, however, could not be observed. After the workshop two participants told us that they enjoyed the method more than common brainstorming sessions. They pointed out that metaphors helped to overcome a barrier to creativity and inspired their ideas.

6.6 Case 3: The Pocket Negotiator: Linking interests to issues

In the previous two chapters we argued for the importance of awareness of values/interests and issue preferences for taking decisions or negotiating. Furthermore, we described how to design interfaces for these aspects. However, to get from underlying interests to concrete issues one can negotiate on with another party, links must be established between the two concepts. That means, a user has to specify how several issues are linked to the interests she wants to satisfy. To design an interface for this kind of input for a decision support system, we used the compositional approach in an adapted form, i.e. in the course of several design workshops.

This series of workshops was set up with the aim to investigate the role of the facilitator, the influence of material used on design activities and outcomes and the composition of groups. We set up four creative prototyping workshops with the same goal of designing an interface prototype, but with variations of participants and group compositions, different types of briefing, evaluation and materials. Workshop A was held with 16 people divided into groups of four with different backgrounds. Workshop B and C were held with four people, non-designers and designers respectively, which were divided into groups of two. Workshop D was held with 12 people divided into two groups of six, each consisting of one designer, one domain expert and 4 end-users.

6.6.1 Materials

As described above we found in the preference workshop that when participants were provided with digital prototypes they had difficulties creating innovative designs. That is why we used paper prototypes of different levels of complexity this time. Session A was provided with very simple, abstract prototypes developed beforehand

by the facilitator. In session B and C we did not provide any prototypes to compare how designers and non-designers use simple office supplies for ideation and creation. In session D, participants were provided with more elaborate prototypes created in earlier user sessions. In addition, participants were given paper interface elements from the same prototypes that they could use to build their interface designs. All groups additionally got coloring pens, post-its, papers, scissors, glue and markers to convey ideas and create their interface designs.

6.6.2 Procedure

Each workshop took about two hours and consisted of a 15-minutes briefing, a divergent process, a convergent process (both prototyping) and a presentation of the results. Short breaks were included allowing people to recover from fatigue. In B, C and D there was an evaluation, which was skipped in A due to time constraints. In D evaluation involved a participant from each group, in the others merely the facilitators. Once the participants were briefed with the problem domain they were encouraged to explore and create different solutions in the divergent stage for 30 minutes. To trigger creativity, participants were provided with materials. Then, participants were instructed to converge towards a prototype in the next 30 minutes. Last, the groups presented their prototypes to each other in about 20 minutes. We gathered observational notes, paper prototypes and evaluation interview data (workshop D).

6.6.3 Observations

When people were given prototypes (A and D) they demonstrated a better understanding of the problem. We were surprised to find that participants utilized the paper prototypes by folding, coloring and arranging them on top of each other to create prototypes instead of using the single elements. They remarked that "we were already talking about those prototypes for half an hour, so when we folded a paper prototype, or cut an element from one of the prototypes, everyone knew what it meant." This interaction was not facilitated by the digital prototypes used in case 1 above, which could be a reason for the less innovative outcomes. We also noticed that designers and non-designers interacted with the material differently. While all groups immediately stuck post-its on the 'screen', non-designers kept working with them by writing on them, rearranging them and drawing other interface elements around them until they reached a final prototype. Designers more often simply left the post-its on the sheet, but used pens to draw other ideas while discussing and sketching many possibilities and analyzing the problem in depth. When they agreed on a final idea, they cleared

the "screen" and started the prototype from scratch. Overall, designers created more complex interfaces. It seemed beneficial in terms of discussion, collaboration and outcomes when groups included a domain expert, a designer and several end-users. In some groups one or two people dominated the whole prototyping process. One participant had "a lot more to add to the design but I could not get time to speak and convince the design partners". Thus we suggest having a designated moderator in each group to steer the collaboration and manage the speaking time per participant. Furthermore, we learned that the facilitation of cooperative workshops is crucial as a lack of it can lead to major breakdowns. In workshop A and B we asked people to come up with their own negotiation domain which was difficult for people and took much more time than we had allocated. Furthermore, we found that breaks and question times are crucial to allow for clarifications and focus shifts that can help participants get back to their work with fresh ideas. Last, evaluations involving participants are helpful to establish bonding between researchers and participants allowing for feedback on the process, insights into participants' views, and more engaged participants.

6.7 Case 4: Support for communication professionals

Focus of the design was to develop a decision support system (DSS) to be used by communication professionals working with the government to design communication strategies. These strategies are employed in campaigns, e.g., for traffic safety or healthier living. It is believed that such strategies could be enhanced by using research outcomes and models developed in the social sciences. However, there is a gap between the work practice and academic research done in this area. The Dutch government is currently putting effort into closing this gap and the development of the DSS we worked on to make social science knowledge accessible to practitioners is part of these efforts. As we knew little about the practitioners' work practice, we decided to follow a cooperative design approach making them co-designers in the process.

We set up four end-user (i.e. campaign managers) workshops building upon each other to move from an investigation of their current work practice to concrete development of a DSS that fits into their practice. This time some participants took part in all workshops. The workshops' focus moved from understanding the current workflow to ideas for DSS and prototypes.

6.7.1 Materials

Different materials were used in the four workshops. In the first workshop the focus was on understanding the current workflow. To trigger discussion the facilitators used propositions formulated in extreme ways that people rated their agreement with on a continuous scale. Materials in the following workshops were office supplies like paper, pens and post-its, cut-out interface elements, several paper prototypes and one digital prototype.

6.7.2 Procedure

The first workshop focused on getting to a shared understanding of the current workflow and started with a warming-up exercise using propositions written by the facilitators to probe the participants' roles. After warming up, a longer discussion followed focused on "what makes a strategy a good strategy?" and "which role does the campaign manager have in designing the strategy?" Participants had to write their answers to the questions on post-its which were then grouped to trigger discussion. Two paper prototyping workshops followed. The first was divided into two parts, one where participants worked in groups of two to design an interface prototype from scratch, and one where the researcher presented his idea in the form of paper interface elements. Participants were asked to combine their own and the researcher's elements into a prototype. Based on the outcome, the researcher designed a new paper prototype that was used in the second paper workshop. Here, the interface was introduced element-by-element. Participants could remove, add or move elements. In the workshop four participants worked together, three who participated earlier. In the last workshop two participants together evaluated the final digital prototype built based on the previous workshops. All workshops were video recorded for later analysis.

6.7.3 Observations

In this case, we focus on the overall set-up instead of concrete interactions during the workshops. Although the process was considered to be rather labor-intensive, participants mentioned that they liked the way in which the design was established. They reflected consciously about what a computer system can or cannot do. Furthermore, according to the participants they felt that their involvement was representative for other campaigns managers. Overall, their initially rather negative opinion about the usefulness of embedding social science models in their work improved during the workshops. One participant said: "I can imagine that [such a system] has a

long term perspective." Problems found in case 1 that participants were influenced by the existing prototypes were mitigated through the different set-up. In the first part of the second workshop participants started with a completely blank "screen" and were asked to design the system from scratch. Thus innovative, unbiased ideas could emerge. However, in the second part, when the facilitator introduced his own ideas, it showed that the researcher had taken a completely different approach to the design. This lead to difficulties with combining user-created elements with those of the designer. Besides the difficulty such breakdowns in the flow of the process lead to new discussions which open up for new perspectives and enhance the shared understanding between researcher and participants.

In the following we summarize our observations and give some recommendations for cooperative design work with end-users and experts.

6.8 Conclusions

As technologies are becoming more complex and their use more unpredictable, cooperative design processes are more important than ever. However, motivating and engaging people in cooperative design is hard, due to participants feeling a lack of confidence or knowledge. We believe that as HCI researchers it is necessary not only to engage in cooperative design, but also to reflect back on the experiences gathered in workshops and systematically analyze effects of different set-ups. When sharing these analyses with other researchers we can add theoretical considerations to the many practical accounts of PD.

We have presented a compositional method that provides participants in design sessions with interface elements as a starting point for creative prototyping. By having the designer prepare prototypes using these elements the participants can already experience possible elements in practice. They are further encouraged to re-combine these elements and add new ones. We have applied this method in three design cases and shown that within short time many ideas could be discussed and new interfaces were created by the participants. However, in the first case we have also seen a lack of innovativeness. Subsequently, we have used the overall idea of combining interfaces elements in several workshops but have adapted the approach in each case.

In particular we have varied the number of participants, the background of participants, the material, the facilitation and the set-ups and reflected back on the lessons learned from each workshop. To summarize the most crucial points, we found differences in how designers and end-users are able to utilize the given material. If

possible, groups should be composed including different roles, e.g. a moderator, design and end-users. The involvement of domain experts is also crucial. In our second design case it was difficult to engage directly with end-users as values are an abstract concept. Therefore, we engaged first with experts in the domain of counseling and value-reflection and created a prototype based on their input. Thus, in cases where the domain poses difficulties to the early engagement of users it is recommended to involve experts first. In other cases experts and users can be involved in the same sessions which can lead to interesting discussions and design outcomes considering different perspectives.

If sessions are carried out with end-users only we strongly suggest using prototypes that can trigger creative thinking. While designers can easily work with abstract materials, more concrete prototypes have been proven useful to give end-users a deeper understanding of the domain, possibilities, and trigger new creations. However, the choice of material is crucial. Whereas digital prototypes may be useful to convey concrete interaction, they can strongly inhibit innovation and influence the outcome. Paper prototypes did not have the same effect, maybe due to their unfinished touch. In addition, they are more accessible in the design process and can support focusing the communication. Therefore, when combining elements participants preferred cutting the existing prototypes over same already cut elements.

Another way to avoid participants being influenced by prototypes was the set-up chosen for case 2 where participants were first asked to create a prototype from scratch with paper, pens and post-its followed by several workshops with the same participants converging more and more towards a digital prototype. However, it turned out to be difficult for the participants to merge their ideas with those of the researchers. A benefit of the continuous set-up of several workshops with the same participants was that throughout the process participants who started out with skepticism towards the introduction of a DSS in their workplace became more positive. Furthermore, we have seen that using a digital prototype prior to design workshops and critiquing it in the session can be an icebreaker and enhance people's confidence in their own contribution. The future workshop set-up in the last case worked well to structure the activities in the workshop and did not influence the outcomes. The limited speaking time avoided domination by single participants as in case 1 and metaphors were useful triggers to creativity. Last, the role of the facilitator has to be taken seriously. Small violations, such as skipping question times, can lead to breakdowns in the process. Explicit breaks can lead to focus shifts which are beneficial to creativity. We hope that this work encourages others to work on a research agenda focusing on similar analysis of cooperative design work in order to build a theoretical account of PD.

6.9 Bibliography

Abras C, Maloney-Krichmar D, Preece J (2004). User-Centered Design. In Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications.

Ertner M, Kragelund A M, Malmborg L (2010). Five Enunciations of Empowerment in Participatory Design. In Proceedings of the Participatory design conference, PDC'10, Sydney, Australia, 191–194.

Gaver B, Dunne T, Pacenti E (1999) Design: Cultural probes. Interactions 6:1, 21–29.

Heineman, GT and Councill, WT (2001) Component based software engineering: putting the pieces together. ACM Press.

Hutchinson H, Mackay W, Westerlund B, Bederson B B, Druin A, Plaisant C, Beaudouin-Lafon M, Conversy S, Evans H, Hansen H, Roussel N, Eiderbäck B (2003) Technology probes: inspiring design for and with families. In Proceedings of the ACM SIGCHI conference on human factors in computing systems, CHI '03, Fort Lauderdale, FL, 17–24.

Kensing F, Madsen K H (1992). Generating visions: future workshops and metaphorical design. In Design at Work: Cooperative Design of Computer Systems, 155–168.

Kensing F, Blomberg J (1998) Participatory Design: Issues and Concerns. Computer Supported Cooperative Work 7, 167–185.

Marois L, Viallet J-E, Poirier F, Chauvin C (2010). Experimenting Introductory Tools for Innovation and Participatory Design. In Proceedings of the Participatory Design Conference, PDC'10, Sydney, Australia, 259–262.

Muller MJ (1991) Pictive-an exploration in participatory design. In Proceedings In Proceedings of the ACM SIGCHI conference on human factors in computing systems, CHI '91, New Orleans, LA, 225–231.

Nathan L, Klasnja P V, Friedman B (2007) Value Scenarios: A Technique for Envisioning Systemic Effects of New Technologies. Ext. Abstracts CHI '07, San Jose, CA, 2585–2590.

Norman D A (1988) The Design of Everyday Things. New York, Doubleday.

O'Neill E (2000) User-developer cooperation in software development: building common ground and usable systems. London, Springer Verlag.

Pommeranz A, Wiggers P, Jonker C M (2011) Towards Compositional Design and Evaluation of Preference Elicitation Interfaces. HCD'11 Proceedings of the 2nd international conference on Human centered design, Orlando, FL, 586–596.

Sanders EB-N, Stappers P J (2008). Co-creation and the new landscapes of design. CoDesign: International Journal of CoCreation in Design and the Arts, 4(1):5–18.

Sanders EB-N, Brandt E, Binder T (2010) A Framework for Organizing the Tools and Techniques of Participatory Design. In Proceedings of the Participatory design conference, PDC'10, Sydney, Australia, 195–198.

Sanders EB-N, Westerlund B (2011) Experience, exploring and experimenting in and with co-design spaces. In Proceedings of the Nordic Design Research Conference 2011, Helsinki, Finland, 1–5.

CHAPTER 7

DISCUSSION AND CONCLUSION

This chapter revisits and answers the research questions posed in Chapter 1. Furthermore, limitations are discussed, contributions of the thesis are outlined and recommendations for future work are given.

I think and think for months and years,
ninety-nine times, the conclusion is false.
The hundredth time I am right.
Albert Einstein (1879-1955)

7.1 Discussion

In this thesis we researched the most important aspects for designing decision support systems accounting for human capabilities and needs and was guided by the main research question 'How can we design user-system interaction for human-centered decision support?'. In the following we revisit and answer the detailed research questions posed in Chapter 1 (section 1.4).

The first set of research questions aimed at defining the design space – the 'what' to design – and are answered in Chapter 3.

1. What functionality is crucial for a NSS from an expert perspective to overcome typical problems in negotiations?
2. What are the needs of end-users with respect to a NSS?
3. In which social situations would people accept the use of a NSS?

To answer questions 1 and 2 we conducted focus groups with expert negotiators and possible target users. To help participants envision a NSS we created five scenarios (in form of videos) that showed when and how people could use such a system. We did not show much concrete functionality and used sketched interfaces to make sure participants understood that there was no existing prototype yet.

From the expert discussions we could extract several dominant themes. First of all, the experts agreed that a NSS adds higher value in the preparation and training phase than during a negotiation. Especially in face-to-face situations they considered it not feasible, but awkward and producing too much cognitive load. Main focus of the preparation is to create an accurate preference profile. Training needs to be interactive, e.g. provided through a virtual training environment and the NSS needs to react intelligently.

Context-sensitivity was another theme. According to the experts, the reasoning of the system should take into account the current atmosphere, non-verbal communication and emotions that may influence the decision making process. A crucial aspect is that the system should be able to adapt to the users' capabilities, e.g. the skill level or bidding strategies of the user.

From the user's perspective we found that people had a critical attitude towards the value of a NSS. Many emphasized that it is important that the user stays independent

from the device instead of following its advice blindly. Furthermore, it is of importance that the advice is presented in a way that is comprehensible to the user. A strength they saw in such systems was that it is helpful in the training and to organize things during the negotiation process. Some participants felt that insecure people would feel more supported and confident.

Based on the analysis of the observations and notes on the group discussions we developed a set of 12 guidelines for the design of NSS (e.g. 'Advice from a NSS should consider information about the context of the negotiation' or 'A NSS should be able to adapt to the user's skill level and experience and more in specific to the user's bidding behavior.').

With regard to question 3 we found in the analysis of our social acceptance survey that not only functionality and usefulness play a role, but also social aspects like the subjective norm and social acceptance. A NSS is not only a tool people use to fulfill a certain task but it is a social device depending on the use context. Therefore, the designer has to determine in which context the device should be used and fit the design to the context and its social norms. Furthermore, our survey has shown that the respondents value the opinions of close friends or family highly, both for deciding whether to use a NSS and when taking decisions during the negotiation. Based on that we recommend designers to design such a system as a companion to the users, and consider integration with social networks.

Once this first exploration was completed our work focused on the preparation phase and in particular the elicitation of preferences. The first part of the research showed that this is a crucial step to allow for intelligent and accurate advice of the system. Other aspects discussed above, e.g. a virtual training, were handled by other colleagues, but are outside the scope of this thesis. We aimed at focusing the system's support on reflecting on what is most important to a user in the given decision context and creating a preference profile. We did so because of the findings discussed above that people don't want to depend on the tool, experts think that being well prepared and having a good understanding of one's own preferences and goals is the most important step, and finally the technology relies on a complete and correct preference profile and is not yet at the point that it can make decision for you, it can only give advice, that the user has to understand. In particular, we investigated the following research questions in Chapter 4:

4. How do people create preferences?
5. What preference elicitation methods exist?
6. What ways do people prefer to express their preferences in interfaces?
7. How can interfaces be designed to fit the user's expression of preferences?

To answer question 4 we reviewed literature in diverse areas, mainly psychology and behavioral decision making. Preferences are affective judgments of alternatives/ items or their attributes. We found that the human process of developing preferences is highly constructive. This view implies that people construct their preferences at the time a valuation question is asked. Furthermore, the decision process itself and the context play a major role in the construction process. There are different views on how people construct their preferences, including among others, that the process depends on memories, the available cognitive resources or is influenced by psychological effects, e.g. based on how information is presented or a valuation question is phrased. People tend to minimize cognitive efforts, which could lead to pitfalls, such as prominence effects, where people focus too quickly on a single dominant attribute. One way to overcome problems with this constructive nature, is to focus on underlying values, which are stable over longer periods of time. Preferences for concrete attributes or alternatives can be created in a decision-context with a focus on achieving these values.

Preference elicitation methods (see question 5) used in DSS and recommender systems range from implicit, e.g. machine learning approaches, to explicit ones, e.g. answering a number of valuation questions. In product recommenders with a large number of items common implicit methods are filtering products based on their content (i.e. comparing attributes to other products a user liked) or user similarity (i.e. suggesting products that similar users liked).

We focused on explicit methods, because we considered it important that the user understands the relation between her preferences and the possible outcomes of the decision making process. Common explicit methods in DSS are absolute measurement, which requires users to enter specific values for attributes and importance weights, and pairwise comparison, which forces users to make direct trade-offs between attributes. These methods are not suited to the cognitive abilities of humans as people do not think in concrete numbers, and trade-offs require a lot of cognitive effort. User-centered methods are few, but can be found in recommender system research, e.g.

conversational methods that ask users to critique and tweak examples and allow users to see the immediate effect of their input. In addition, the majority of existing methods does not consider underlying values or affective input.

To answer question 6 we conducted two user experiments. The first experiment focused on people's effort and liking of different ways of entering preferences, i.e. ranking and rating of either attributes or alternatives on a numbered scale or affective input device and a navigation through the alternative space. The results confirmed that cognitively less demanding ordering or rating tasks were perceived as less effortful and liked most by users. Effort, however, was not always an indicator of how much a method was liked. Affective feedback and navigation were rated significantly higher in effort than other methods, but still high in liking.

To further investigate user's motivation of giving detailed preference feedback we conducted a second experiment in which people were asked to enter preferences on images and music samples with as much detail (using input devices from thumbs- and stars-ratings, affective feedback and free-form feedback) as preferred for each item. We found that familiarity, ownership and having an opinion about that item are the main factors in influencing the preference detail people are willing to give, and thus the amount of effort they are willing to put into giving feedback. As we found only a small difference in detail for pictures versus music, we tentatively concluded that the willingness to give feedback is not triggered by content types. Our results also show that multidimensional affective feedback is used when people have the choice to do so. Moreover, people in general prefer to give more feedback in the form of multidimensional affective feedback than to give more feedback using a finer grained one dimensional method.

Given these first results we subsequently engaged in design activities involving end-users to answer question 7. We created four digital prototypes that used insights from the above and asked users to interact with them and evaluate them in a compositional way, i.e. focusing on the different interface elements rather than the complete interfaces. In a follow-up prototyping workshop with two teams of end-users, participants were asked to create new prototypes according to their own needs. We provided the interface elements of the first prototypes as paper versions as well as other general elements like buttons, textboxes etc. and additional crafting material. This process allowed us to evaluate many ideas/elements in an efficient way and resulted in two new prototypes that were later combined and implemented in a prototype NSS for job contract negotiations. Based on the results from the experiments, group discussions in the prototyping sessions and the resulting designs we developed four guidelines for

designing preference elicitation interfaces presented in Chapter 4 (section 4.6).

One aspect of the resulting design was that underlying values, preferences and example outcomes were visible in the same space to allow users to explore their interaction. We used preset value profiles to elicit values in this design. However, participants felt less comfortable with choosing a preset profile, as their value systems were more complex than reflected by the profiles. This result led to the next set of research questions investigated in Chapter 5:

8. Why are values difficult to assess?
9. How do experts support people in assessing their values?
10. How can we design tools that help people reflect on their values?

With regard to question 8 research in psychology and behavioral science shows that although values are seen as guiding principles and particularly relevant for decision making, they operate on an unconscious level. Because of their abstract nature people find it difficult to assess and articulate values. An exploratory photo elicitation study we conducted with participants of different ages and occupational backgrounds showed strong personal differences in ways people prefer (or are able to) express their values. Whereas some were able to take pictures that reflected their values, others could easily tag images with single value terms and others could more easily talk about what is important to them. Overall, we found a lack of reflection on values in everyday life and a lack of the ability to abstract from concrete situations to high-level value systems.

In order to design digital tools to support people in reflecting on and assessing their values, we found it was difficult to engage directly with possible end-users – as we did in the design of preference elicitation tools – due to the difficulties named above. Therefore, we first engaged with experts, i.e. life and job counselors, to answer question 9 (see section 5.4). From expert sessions we could extract the following themes that are important in supporting people's value-reflection: uniqueness, individual truth, trust, guidance, emotional triggers, and social aspects.

To assess people's values counselors have to take into account the following. Every client and, thus, every session is unique and has to be provided with different methods to reflect. A related aspect is that often reflections of people seem contradictory to an outstander, but make sense to the person herself. A counselor aims at not being

judgmental in this case, but promoting individual truth. Another overarching theme was the role of trust. It is often difficult for people to open up and discuss their intimate experiences. This difficulty can be reduced through the built-up of a trusting relationship between the counselor and client. Once trust is established it is important for the counselor to provide guidance to the client to get from concrete descriptions of experiences and feelings to abstract value concepts. The majority of methods use emotional triggers as a starting point for reflection, e.g. images, music, writing accounts of experiences, etc. Concrete methods are, among others, association cards, storytelling, metaphors or reflection questions. In some cases group sessions are conducted as the social setting helps people to see things from a new perspective.

To investigate question 10 we created a prototype for value-reflection in the context of job choices based on the above expert input, which we employed as a probe to trigger user feedback in three subsequent studies (section 5.5). First, we conducted an interactive user study, in which participants used the prototype for up to four weeks in their daily life and were in continuous contact with the researcher through a message function within the prototype. After use participants filled in a questionnaire focusing on the usability and usefulness to reflect as well as people's attitudes towards awareness of values and their feeling after using the tool. We obtained positive results with regard to people believing that more awareness of values is beneficial for a job negotiation and on their level of preparedness for the job negotiation. However, views on the tool itself and its ability to support reflection were diverse. From this study we received detailed feedback on bugs, aesthetic problems and conceptual problems (see section 5.5.1).

To explore these problems and develop ideas for improvements, we conducted a user workshop with a subset of the same participants. This workshop was divided into critique, fantasy and implementation phase and used metaphor generation and sketching as creativity tools. The workshop was useful in pointing out the main challenges of creating a digital tool for value-reflection. Some themes that were discussed overlapped with the expert themes. A major discussion arose around the aspects of system guidance and a more free-form style. Participants generally wished for a personal and private feel which could be achieved through a free-form diary-style interface. However, at the same time too much freedom may demotivate the users. Guidance is useful, especially when getting from concrete reflections and own values to work-related values. For this a conversational approach was suggested. Furthermore, participants wished for the system to give them new, and maybe unexpected insights about their own values. They also thought transparency about the goals of each reflection tool was necessary. Integration with social networks was also discussed but

stood in conflict with a private feel and was, therefore, not preferred by all users.

Besides interesting discussions, the workshop also provided concrete design ideas sketched on paper, e.g. the use of a value pie chart that adapts to user input. In order to develop concrete design guidelines for value-reflection tools we tested the ideas of the small sample of users with a bigger number of people. Therefore, we conducted an online survey using design sketch pairs for each idea that was discussed in the user workshop and asked a large sample of people about their preferences. To get deeper insights we also elicited people's level of reflection and personal characteristics.

We found clear preferences for four out of ten design ideas, i.e. for users to add their own values to the systems, personal questions for reflection, receive immediate, visible feedback about value frequencies while filling in new a reflection and a dialog with the system to understand and assess values. Furthermore, the majority of respondents was against an integration with social networks. From the insights gained through this survey and the detailed discussions with experts and users we created five design guidelines for digital value-reflection tools (e.g. 'Aim for a trustworthy design of the tool through careful implementation of privacy, dialog with the user and transparency', see section 5.7.1).

Besides the concrete design guidelines for DSS, we also aimed at answering the following research questions on the methodological aspects, i.e. the 'how' to design:

11. Which design and research methods are useful in the design of human-centered DSS?

12. How can we involve end-users and domain experts actively in the design process?

Throughout our research, we chose a mix of design and research methods and selected carefully at each stage of the research which method would lead to the appropriate results. In order to answer question 11 we reflect back on the chosen methods and point out some of the strength and weaknesses we found.

In the exploration phase we conducted focus groups with experts and users. These were useful in getting in-depth insights into people's attitudes with regard to the envisioned system. To support people getting an idea of the type of system we wanted to develop we constructed several filmed scenarios that showed how the system could be used. While the videos were useful as an icebreaker in the groups, they also biased

the participants towards the designer's first ideas. This limits the amount of insights into aspects not thought of by the designer beforehand. One way to mitigate this risk could be to conduct a number of focus groups with the same participants, first without and then with concrete scenarios. Generally, the scenarios were helpful for the project team to achieve a shared understanding of the goals of the project.

After the exploration concrete hypotheses need to be tested and designs created. For the first, experiments were useful. Especially conducting them online, when possible, can be efficient with regard to time and the number of participants. However, it needs to be considered in how far the sometimes artificial setting may influence the results. Instead of testing different preference input methods in a lab setting, another way would have been to explore how people use real systems that use the same methods.

To design human-centered DSS we chose for active involvement of the users in design workshops. Besides the actual design outcomes of these sessions, the process itself and the participant's discussions provide interesting insights into people's way of thinking and their needs. These insights are difficult to obtain from simple prototype evaluations where the focus lies more on the system and its flaws instead of the user.

However, creating opportunities for the experts and users to get involved is not easy. Thus, we focused our analysis also on answering question 12. To receive good feedback, participants need to be motivated and feel that they are able to contribute. With regard to experts and users, we found that it is important that their involvement matches their roles or expectations. In our experience, experts felt more comfortable in being interviewed, as it allows them to provide their expert knowledge, than actually design an interface. With regard to users we found that giving them the opportunity to provide feedback during the use, as in the value-reflection study, increased their engagement and motivation for the subsequent workshop. Chapter 6 discusses aspects of setting up cooperative design workshops in-depth and presents several case studies. In particular, we presented a compositional design method that focuses on creating and evaluating design ideas in an efficient way. Furthermore, we investigated facilitation, group composition and creativity triggers in the workshops.

Overall, we believe, our choice of methods was useful and especially the mix of different methods resulted in insights from different viewpoints. In addition, it was helpful to involve domain experts and end-users, as the experts can provide in-depth insights into existing challenges, e.g. pitfalls of decision making that users are not aware of, and users know better how they would like to use a system. Active involvement of end-user in participatory design processes is beneficial to understand the user and design interfaces that users will more easily accept. However, it requires a

lot of time. Furthermore, the process could have been enhanced by a better integration with actual implementation of other parts of the system. By that the underlying reasoning and technical implementation could have been adapted to the user's mental model and the developed interfaces could have been tested in their actual use context.

7.2 Limitations of results

Some of the results, e.g. the design guidelines presented in Chapter 3, are specific to the design of NSS, which was the goal of the project (Pocket Negotiator) in which this thesis was carried out. In particular, aspects of using the system face-to-face with the opponent or the bidding functionality are not applicable for DSS in general. However, aspects like the importance of preparation and transparency are applicable to DSS and NSS alike. The work carried out in Chapters 4,5 and 6, however apply to DSS as well.

Furthermore, evaluations of the interfaces presented in this thesis were mainly done in a stand-alone manner, due to a lack of progress on the overall Pocket Negotiator system. In such cases we cannot draw conclusions on the effect of an interface on the overall decision or negotiation outcome. The preference elicitation interface was evaluated in a usability study of an early Pocket Negotiator prototype. However, the implementation of the interface had to be adapted from the original outcomes of the participatory design work presented in chapter 4. This and the overall usability of the Pocket Negotiator had major influences in the evaluation of the interface. While users understood the concepts and were generally able to use the interface, many flaws were pointed out. One example is the fact that in the implemented version several post-its for the same issue can be created but move simultaneously along the y-axis when one is dragged. This is due to the system's preference model requiring one importance weight per issue (and not per issue-value pair).

This example also points to the limits of the methodological approach taken. Participatory design approaches can lead to challenges when the core of the system is implemented in a way that poses restrictions on the interface. Generally, this can be mitigated by an integrated approach to frontend and backend design. However, this was not given in the current project. Within the project efforts are going on to integrate the work of several researchers into the current prototype and an overall user evaluation is planned.

In addition, due to time restrictions the value-reflection tool is currently in a prototypical state that needs further development and testing as part of real-life decision making processes. Last, all interfaces were developed for the job domain,

i.e. to make career decisions or negotiate a job contract. While it can be assumed that the general process of constructing preferences and reflecting on values would be supported by the interfaces in other domains, this still has to be shown.

7.3 Conclusion

This thesis raises awareness of human-centered issues in the design of decision support. Overall, our contributions to DSS research are three-fold: (1) guidelines for DSS design from an expert and user perspective, (2) results from concrete design studies with regard to preference and value elicitation, and (3) an analysis of appropriate methods with regard to involvement of stakeholders in the design process.

In order to create human-centered DSS we worked in an interdisciplinary manner using knowledge from psychology and behavioral science to create interfaces supporting human characteristics of thinking, self-reflection and decision making. In addition, our focus was on finding appropriate ways to give people a voice and enhance communication between designers and stakeholders, i.e. in our case experts and users. The majority of DSS and NSS are still seen as knowledge-based systems and are engineered to provide expert knowledge to their users. This knowledge elicitation paradigm is one-directional and does not allow for creating new knowledge and new interactions. Focusing instead on sharing knowledge through engagement of designers and users as well as developing new designs together will enhance design outcomes and acceptance of these outcomes by users.

By taking this cooperative approach, we worked towards a system design that supports individuals to gain an understanding of their values, preferences and behaviors in order to take *their own informed and satisfying decisions* and communicate them to others. This is a novel approach to DSS/NSS design which has so far focused on automation and analytical tools taking away choices from the user to prevent human error.

In conclusion, our results help designers of DSS to create human-centered systems that support users in gaining self-awareness of their values, understanding the link between their values and preferences and take their own decisions. We have taken an important step towards reflective DSS that empower people in their decision making, instead of automating their decisions.

7.4 Future work

In the future, our proposed interfaces and design guidelines need to be implemented in different types of DSS for various domains and evaluated in diverse real-life settings with many users. The research was conducted as part of the Pocket Negotiator project aimed at building NSS that complement the users' and system's capabilities in order to enhance negotiation outcomes. At this moment the project focuses on job and real estate negotiations, however, current implementation efforts are aimed at offering a wider range of negotiation domains.

Furthermore, the value-reflection system (Chapter 5) needs to be improved, integrated into the current system and evaluated in studies lasting at least several weeks, as value awareness takes time.

Besides the specific implementations and evaluations, important aspects for future investigation are the role of trust, user-system collaboration, explanation and shared mental models. The importance of these aspects was claimed by the experts in the focus groups we conducted which pointed to the role of training and intelligent advice of the system, and further support by participants of the user focus groups and social acceptance survey. In addition, user-system collaboration and trust were considered by experts and user with regard to the value-elicitation tool.

In the Pocket Negotiator project we have touched these topics partially but more in-depth investigation is needed. In a side-project we investigated how to enhance trust between the user and the system. Trust is an important concept when it comes to the adoption and reliance on technology, and even one breach of trust can highly influence user perception of that technology. In empirical studies of several systems, among them the Pocket Negotiator, we identified ambiguity, transparency and open interpretation as important interface qualities that can lead to increased trust. The rationale behind this is that, in parallel to human-human trust, systems that show their vulnerabilities through transparency and leave room for the user to interpret the results (e.g. presenting information in a way that leaves room for interpretation) based on her own knowledge seem to be more trustworthy. This research needs to be followed up on and an in-depth investigation of how to design for these qualities is needed.

Furthermore, efforts with other colleagues aim at investigating how to achieve a shared mental model of a decision (or negotiation) domain between the system and the user. Transparency, explanation and collaboration between the system and the user need to be considered as part of these investigations.

While these future work strands are focused on the design of DSS, we believe that more work is also needed on the methodological level. In Chapter 6 we called for more systematic analysis of participatory design workshops. We have started to look at the influence of facilitation, group compositions and materials used in workshops on stakeholder engagement and design outcomes. This work needs to be expanded to more cases and also more aspects of workshops such as location, set-up or dealing with diverse stakeholders. With respect to the latter issues such as reaching a common ground in communication, sharing knowledge and negotiating design trade-offs are of interest within HCI and in particular needed for value-sensitive approaches. We have touched the communication of values between designers and stakeholders in previous work (see [3] in the list of publications of this thesis). However, future work is needed to create tools for stakeholders in design processes to reflect on, share and trade-off their values.

APPENDIX A

A.1 Questionnaire - English version

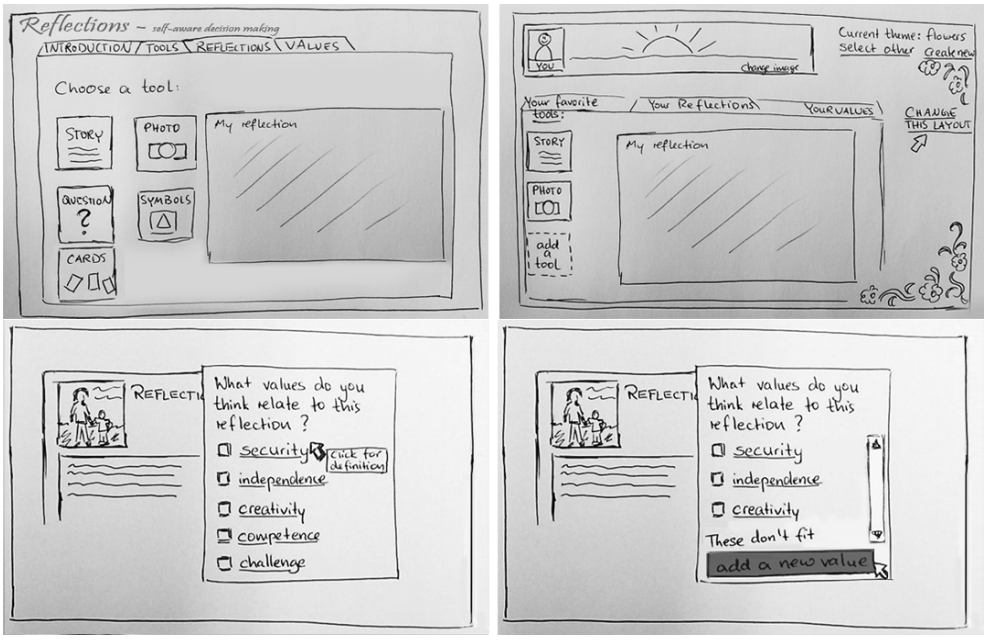
(Unless otherwise specified in the footnotes the answers were measured by a 7-point Likert scale)

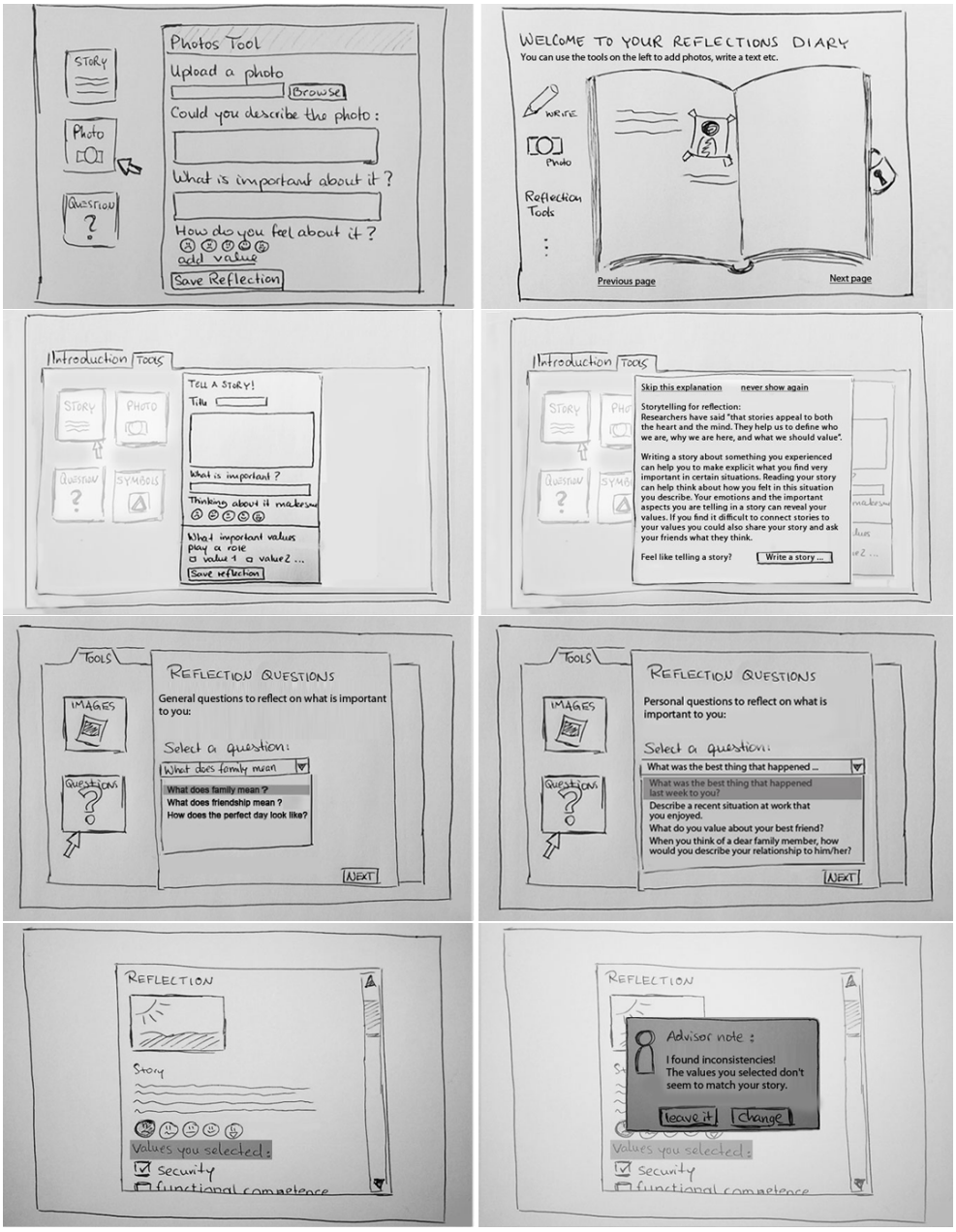
A. Appendix A

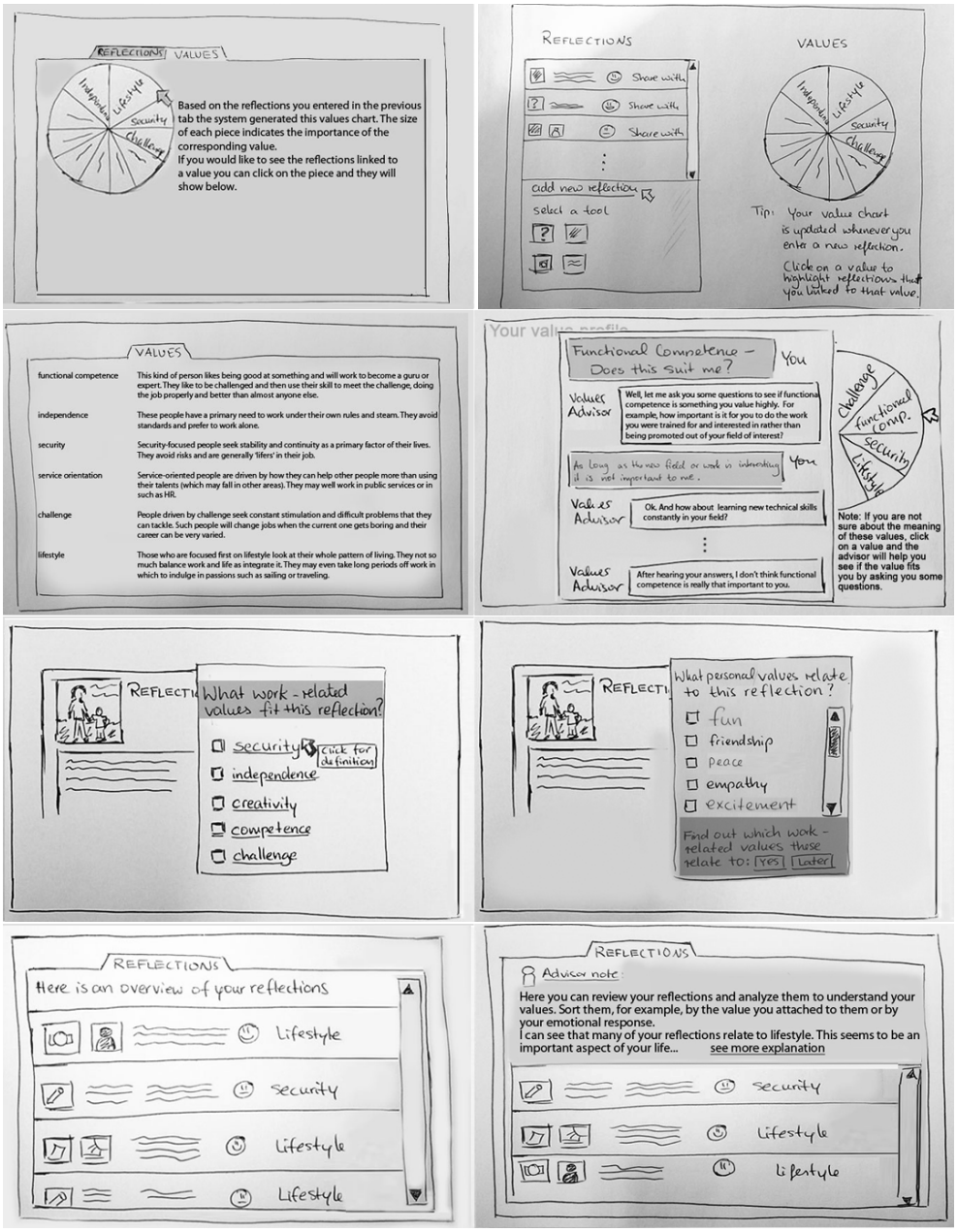
Item/Construct	Question
	Before all scenarios
GEN	What is your gender? (male/female)
COU	What is your nationality? (open)
EDU	What is your level of education? (No degree, vocational training, university degree)
AGE	How old are you? (open)
CSK	How many hours do you spend using computers per week? (open)
(NEX)	
NEX 1	How many houses have you sold? (open)
NEX 2	How many houses have you bought? (open)
NEX 3	How many job interviews have you had? (open)
NEX 4	Is negotiation an important part of your job? (yes/no))
(NAT)	
NAT 1	Negotiation is a game.
NAT 2	I try to avoid negotiations.
NAT 3	I enjoy negotiations.
NAT 4	Negotiations are a necessary must.
(NSK)	
NSK 1	I am a good negotiator.
NSK 2	I would rather negotiate myself if the negotiation task is simple.
NSK 3	I would let someone else negotiate for me if the negotiation task is simple.
NSK 4	I would rather negotiate myself if the object of the negotiation is important for me.
NSK 5	I would let someone else negotiate for me if the object of the negotiation is important for me.
	After each scenario
IU	I would use the Pocket Negotiator (PN) in the situation shown in the video/picture.
SN	Most people who are important to me would think a Pocket Negotiator is useful in this situation.
(SA)	
SA 1	I think it is socially acceptable to use a PN in this situation.
SA 2	I think the opponent would think it is socially acceptable to use a PN in this situation.
Specific	
train	I expect a PN to prepare me in a short (1-2 hours) time before a negotiation.
f-2-f	A PN would be useful to propose new options for the negotiation.
coll.prep.	I expect a PN to help me organizing data (e.g. information from the internet).
phone	I expect from a PN to give me a clear overview of the negotiation process.
car dealer	I believe the advice that the PN gives is useful for the negotiation.
Comment	
COM	Could you please explain what you based your ratings on? (open)
	After all scenarios
PNA	My attitude towards using a PN is positive.
BC	
BC 1	I would probably feel comfortable using a PN on my own.
BC 2	Learning to operate a PN would probably be easy for me.
BC 3	I would probably understand how to use a PN.
USE	
USE 1	A PN would help me to reach a better outcome in a negotiation.
USE 2	I would feel more confident in the negotiation while using a PN.
USE 3	I will learn how to negotiate better through using the PN.
USE 4	Using a PN would increase my productivity.
USE 5	Using a PN would increase my negotiation performance.
USE 6	Using a PN would enhance my effectiveness in negotiations.
USE 7	Using a PN would make negotiations easier for me.
USE 8	Overall, I find the PN useful for house/job negotiations.
OCM	Please feel free to enter comments here: (open)

APPENDIX B

B.1 Sketches




Photos Tool
Upload a photo
Browse
Could you describe the photo:
What is important about it?
How do you feel about it?
add value
Save Reflection
STORY
Photo
QUESTION
WELCOME TO YOUR REFLECTIONS DIARY
You can use the tools on the left to add photos, write a text etc.
WRITE
Photo
Reflection Tools
Previous page
Next page
Introduction
Tools
TELL A STORY!
Title
What is important?
Thinking about it makes...
What important values play a role
value 1
value2 ...
Save reflection
STORY
PHOTO
QUESTION
SYMBOLS
Skip this explanation
never show again
Storytelling for reflection:
Researchers have said "that stories appeal to both the heart and the mind. They help us to define who we are, why we are here, and what we should value".
Writing a story about something you experienced can help you to make explicit what you find very important in certain situations. Reading your story can help think about how you felt in this situation you describe. Your emotions and the important aspects you are telling in a story can reveal your values. If you find it difficult to connect stories to your values you could also share your story and ask your friends what they think.
Feel like telling a story?
Write a story ...
Tools
IMAGES
QUESTION
REFLECTION QUESTIONS
General questions to reflect on what is important to you:
Select a question:
What does family mean
What does family mean ?
What does friendship mean ?
How does the perfect day look like?
NEXT
Personal questions to reflect on what is important to you:
What was the best thing that happened ...
What was the best thing that happened last week to you?
Describe a recent situation at work that you enjoyed.
What do you value about your best friend?
When you think of a dear family member, how would you describe your relationship to him/her?
REFLECTION
Story
Values you selected:
security
functional competence
Advisor note:
I found inconsistencies!
The values you selected don't seem to match your story.
leave it
change


REFLECTIONS
VALUES
Based on the reflections you entered in the previous tab the system generated this values chart. The size of each piece indicates the importance of the corresponding value.
If you would like to see the reflections linked to a value you can click on the piece and they will show below.
REFLECTIONS
VALUES
Share with
Share with
Share with
add new reflection
select a tool
Tip: Your value chart is updated whenever you enter a new reflection.
Click on a value to highlight reflections that you linked to that value.
VALUES
functional competence
This kind of person likes being good at something and will work to become a guru or expert. They like to be challenged and then use their skill to meet the challenge, doing the job properly and better than almost anyone else.
independence
These people have a primary need to work under their own rules and steam. They avoid standards and prefer to work alone.
security
Security-focused people seek stability and continuity as a primary factor of their lives. They avoid risks and are generally 'lifers' in their job.
service orientation
Service-oriented people are driven by how they can help other people more than using their talents (which may fall in other areas). They may well work in public services or in such as HR.
challenge
People driven by challenge seek constant stimulation and difficult problems that they can tackle. Such people will change jobs when the current one gets boring and their career can be very varied.
lifestyle
Those who are focused first on lifestyle look at their whole pattern of living. They not so much balance work and life as integrate it. They may even take long periods off work in which to indulge in passions such as sailing or traveling.
Functional Competence – Does this suit me?
You
Values Advisor
Well, let me ask you some questions to see if functional competence is something you value highly. For example, how important is it for you to do the work you were trained for and interested in rather than being promoted out of your field of interest?
As long as the new field or work is interesting it is not important to me.
You
Values Advisor
Ok. And how about learning new technical skills constantly in your field?
Values Advisor
After hearing your answers, I don't think functional competence is really that important to you.
Note: If you are not sure about the meaning of these values, click on a value and the advisor will help you see if the value fits you by asking you some questions.
What work - related values fit this reflection?
security
click for definition
independence
creativity
competence
challenge
What personal values relate to this reflection?
fun
friendship
peace
empathy
excitement
Find out which work - related values these relate to: Yes Later
REFLECTIONS
Here is an overview of your reflections
Lifestyle
Security
Lifestyle
Lifestyle
REFLECTIONS
Advisor note:
Here you can review your reflections and analyze them to understand your values. Sort them, for example, by the value you attached to them or by your emotional response.
I can see that many of your reflections relate to lifestyle. This seems to be an important aspect of your life...
see more explanation
Security
Lifestyle
Lifestyle

B.2 Questionnaire items

Item	Statement/Question part 3b
item3.0	My Facebook (or similar network) account should ... a. be used to create a profile of me on the Reflections website. b. be used to create reflections based on status updates in Facebook. c. be used to connect me to my Facebook friends on the Reflections website. d. never be used on the Reflections website. e. The above does not apply, as I don't have a Facebook (or similar network) account.
item3.1.	I expect the system to help me create links between personal reflections and values.
item3.2.	I know better than the system how my personal experiences relate to abstract values.
item3.3.	I want the system to teach me something I did not know about myself.
item3.4.	Looking at (old) photographs often makes me reflect on my past experiences.
item3.5.	I often play songs that remind me of a certain situation or experience.
item3.6.	Artworks inspire me to think beyond what I see.
item3.7.	My old diary entries make me think of who I was or who I am.
item3.8.	I only share reflections on myself with people I trust. The same is true for a digital system.
item3.9.	I need to know that my data will be held private; otherwise I would not enter reflections on myself.
item3.10.	A breech of privacy would make me stop using the system immediately.

Item	Statement/Question part4
item4.1.	I know exactly what a value is.
item4.2.	The concept of values is still hard to grasp.
item4.3.	I know what my values are.
item4.4.	I know how my values relate to my decision making.
item4.5.	I have a clear idea of my life goals.
item4.6.	I know exactly how my life goals relate to my values.
item4.7.	More awareness about my values will be beneficial to choose a job/career.
item4.8.	More awareness about my values will be beneficial in a job negotiation.
item4.9.	I would use a digital tool (designed similar to the interfaces presented before) to self-reflect.
item4.10.	I think, using a digital tool to self-reflect would help many people in making better decisions.
item4.11.	Imagine an upcoming job negotiation. After using a digital tool for self-reflection I would probably feel ... ...self-confident. ...well-prepared. ...know exactly what I want. ...know what I already know. ...nothing of the above.

Summary

Taking major life decisions, e.g. what career to follow, is difficult and sometimes emotional. Besides finding out what exactly one wants for oneself, part of the decision making process consists of considering the long-term consequences of the decisions and being empathetic for loved ones affected by the decisions. Furthermore, such life decisions can involve negotiations with other parties, e.g. the seller of one's desired house or representatives of the company where one wants to work. Besides these soft issues, decision making deals with establishing and browsing alternatives as well as weighing options according to one's preferences. This can be difficult due to the vast number of alternatives to consider and information overload, especially for untrained decision makers. Digital decision support systems (DSS) promise help in this process. However, the currently prevailing focus in DSS on economic models and technical solutions for expert decision makers limits their applicability and success in life decision support for untrained decision makers. We believe, DSS research is in need of a human-centered approach focusing on the users' cognitive capabilities and needs to ultimately support people in taking their own, informed decisions. The dissertation was guided by the main research question "How can we design user-system interaction for human-centered decision support?" In particular, the investigations were two-fold, i.e. focusing on requirements and concrete design guidelines and on the methodology involving the stakeholders in the design process (*Chapter 1*).

The central concept in our work is the interaction between the user and the system. As outlined in *Chapter 2* several perspectives on interaction exist and have evolved within the field of Human Computer Interaction (HCI). These range from approaches to formalize the interaction and seeing the human user as an information processor to phenomenological approaches that view interaction as tied to the physical and social environment. In the later view, the user's background, characteristics and the use situation influence each interaction with a digital artefact's and thus need to be studied together. So-called user-centered design approaches are nowadays prevailing in the study of how systems should be designed with the human user in mind. In addition, co-design approaches even involve the future users and other stakeholders by making them active design partners, who take part in creative activities and decision making. In our work we carefully selected at each stage of the research which methods to employ in order to answer the questions at hand. These included quantitative methods, e.g. surveys, allowing us to investigate attitudes of a larger population and qualitative methods, e.g. prototyping or design workshops, to explore design ideas and user needs with experts and users.

In the early phases of the research, presented in *Chapter 3*, we focused on exploring the design space for DSS and in this case a particular type of DSS, i.e. a negotiation support system (NSS). The focus of our project was a NSS to be used by one negotiation party to prepare and carry out the negotiation with the opponent party. In particular, we were curious in which situations it would be socially accepted to use a NSS and what functionality such as system should have according to negotiation experts and end-users. We used focus groups with experts and users and an online user survey to obtain results. To trigger participants' thoughts on different possible use contexts we provided filmed scenarios portraying five different contexts. Focus of the experts groups was the functionality of the system and from their discussions we extracted 12 overall design guidelines for NSS. They highlighted the importance of a good preparation for successful negotiations. The user groups and survey pointed to social issues. In particular, the survey results show that the use context is an important factor influencing the social acceptance of NSS. The survey's respondents would not accept the use in face-to-face situations when the relationship to the opponent was important. However, when the relationship is less important, i.e. with a car dealer, it is more accepted. Surprisingly, the subjective norm is the most dominant factor influencing the intention to use a NSS. People value opinions of their close ones highly when deciding whether to use an NSS and also ask them for advice when negotiating. In summary, we learned that the preparation phase is crucial for a good negotiation outcome. In addition, the system should be a companion throughout the negotiation process and social functionality would be beneficial.

The concrete design of tools for preparation was the focus of the next steps in our research. *Chapter 4* describes several studies investigating how to design interfaces for preference elicitation – a major part of the preparation. A problem with regard to preference elicitation in current systems is that interfaces are not suited to the constructive nature of human preferences and are often based on rational, quantitative models that do not match the mental models of people. To close this gap between people's cognitive capabilities and user-system interaction we conducted three studies dealing with (1) different ways of entering preferences, (2) factors influencing a user's motivation to enter preference details and (3) participatory design of interfaces that support the human preference construction process. The first study showed that for preference elicitation methods with similar input, e.g. rating or ordering tasks, generally effort and liking were negatively correlated. However, in cases where the process (navigation) and the type of feedback (affective) was more sophisticated in terms of expressive power and understanding of one's own preferences, participants rated the methods high in liking even though the results show a substantial increase in

perceived effort or are less easy to use. The follow-up experiment further revealed that familiarity with and a formed opinion about an item led people to give more detailed preference feedback. Interestingly, people, who entered more detail did so with giving (2-dimensional) affective input in the majority of cases. The third study was focused on creating concrete designs in co-design with end-users. From the study we learned that an important aspect of the process is that it allows people to understand their own preferences and that people feel in charge of creating their profile as opposed to just answering questions used by the system to build the profile. In particular, being able to explore their preferences from different angles including underlying values and consequences (in form of rankings of decision outcomes) within the same interface supported people's process of constructing their preferences. Participants liked design elements that supported this exploration in a natural way that allowed immediate visual feedback. The resulting prototype of the third study is one contribution to the field and was further implemented in a NSS and tested in usability studies. The main contribution are four design guidelines for user-centered preference elicitation (e.g. 'affective feedback should be considered as a way for specifying detailed preference feedback with multiple dimensions.')

In our work on preference elicitation we have seen that one of the main difficulties lies in the constructive nature of people's preferences. This means that preferences can be influenced by the current decision context, they can change over time, and people often do not know their preferences. To support people to construct an accurate preference model in a given context and mitigate risks of unforeseen consequences, we have seen that underlying values are a useful concept to focus on in the preference construction. Personal values are stable over longer periods in people's lives and can be seen as guiding principles. When focusing on values in decision making and sharing thoughts on values with negotiation partners, more satisfying outcomes can be reached. However, people's value systems are complex and difficult to assess or articulate. Currently there are no digital tools available to support people in defining their value profiles with regard to decisions. *Chapter 5* therefore deals with the design of such a tool. Based on expert advice we designed a prototype for value-reflection, which was used in subsequent user studies and an online survey. In these studies we identified a number of important design considerations. One prevailing aspect was the uniqueness of people with regard to self-reflection. It was considered crucial that value-reflection tools offer many different ways to reflect to enable users to choose the ones they suit them best and allow the users to add their own values. Generally, tools that trigger emotional reactions work best to begin a reflection process. Furthermore, a personal feel, e.g. through a diary-style interface, is important to the users. Related

to the personal feel was the importance of establishing trust between the user and the system. This could be achieved through carefully considering privacy options, a dialog-style interaction with the system and transparency allowing the user to see immediately how their value profile is adapted while they enter reflections. An interesting discussion arose considering the system's guidance, which was deemed necessary at some points, e.g. in getting from a reflection to abstract values, and hindering the intimate, unique way of reflection at other points. From the results we compiled a list of design guidelines for value elicitation tools (e.g. 'Aim for a trustworthy design of the tool through careful implementation of privacy, dialog with the user and transparency'.).

Throughout our design work briefly outlined above we have focused on cooperative methods involving end-users and other stakeholders in creative activities, i.e. prototyping or developing design ideas. In particular, we were interested in how to support people to envision a future system and take part in creative activities without a formal education in design. We developed a design method involving participants in compositional prototyping activities. Furthermore, we varied facilitation, group composition and creativity triggers in several design workshops carried out in activities part of or related to this thesis. An important aspect of design-based research is to critically reflect on the process and outcomes. *Chapter 6* offers a reflection on our methodological choices. To summarize, we found it beneficial to involve people with different roles, i.e. end-users, domain experts and designers, in design workshops. As these groups of people use provided materials differently it is useful to provide prototypes as boundary objects to trigger creative exploration and shared understanding. Furthermore, interface elements or paper prototypes as well as a structured brainstorming process can support people to generate new design ideas.

As elaborated in *Chapter 7*, the dissertation's contributions to HCI and DSS research are three-fold. We provide (1) crucial design considerations for DSS, (2) a list of concrete design guidelines to be used by others in the field, and (3) an analysis of appropriate methods with regard to engagement and empowerment of stakeholders in the design process. Design considerations were derived from investigating literature from several research disciplines (psychology, behavioral decision making, decision theories, artificial intelligence and human computer interaction) and exploratory studies with domain experts and users. These highlighted the crucial preparation phase of decision making and negotiations and the social factors of the process. Design guidelines were derived through design-based research involving experts and users at different stages for the overall design of NSS and in particular preference construction and value-reflection in DSS. A dominant theme was the delicate balance

between supporting human ways of thinking and reflecting and giving intelligent guidance which needs to be created by system designers. This balance can only be achieved through close, iterative interactions with end-users, domain experts and designers throughout the design process supported by skilled facilitators.

With this thesis we call for a shift in DSS research from engineering expert systems taking over decision making to designing human-centered support for people to make their own, informed decisions. We hope that others take this work as a starting point for their own research to empower decision makers in the future.

Samenvatting

Het nemen van belangrijke beslissingen, bijvoorbeeld over wat voor soort carrière te volgen, is moeilijk en soms emotioneel. Niet alleen moet iemand op zo'n moment bepalen wat hij voor zichzelf wil, maar daarnaast moet hij rekening houden met de gevolgen op lange termijn en met de consequenties voor anderen die door de beslissing worden beïnvloed, zoals een partner of kinderen. Bovendien kunnen er ander partijen in het spel zijn waarmee rekening gehouden moet worden of waarmee onderhandeld moet worden, zoals de huidige eigenaar van je droomhuis of de baas van het bedrijf waar je wilt werken. Belangrijke aspecten van besluitvorming zijn het exploreren en bepalen van verschillende alternatieven en het afwegen van opties op basis van eigen voorkeuren. Dit kan moeilijk zijn, met name als de beslissingsnemer zich, door een groot aantal alternatieven, geconfronteerd ziet met een overvloed aan informatie.

Digitale beslissingsondersteunende systemen (Decision Support Systems (DSS) in het Engels) kunnen uitkomst bieden. De huidige nadruk bij het onderzoek naar DSS op economische modellen en technische oplossingen voor professionele beslissers, beperkt echter de toepasbaarheid en het succes van deze technologie voor ongetrainde beslissers. Wij denken, dat er in DSS onderzoek behoefte is aan een mensgerichte aanpak. Beslissingsondersteunende systemen moeten zijn afgestemd op het cognitieve vermogen en de behoeften van een gebruiker om uiteindelijk mensen te ondersteunen bij het nemen van hun eigen, geïnformeerde beslissingen.

De centrale onderzoeksvraag in dit proefschrift is: " Hoe kunnen we de interactie tussen gebruiker en een DSS zo ontwerpen dat de gebruiker centraal staat?" Het onderzoek richt zich op twee aspecten: (1) de eisen en concrete ontwerprichtlijnen voor DSS en (2) de methodologie met betrekking tot het integreren van gebruikers in het ontwerpproces. (hoofdstuk 1).

Het centrale concept in dit onderzoek is de interactie tussen de gebruiker en het systeem. Hoofdstuk 2 beschrijft verschillende perspectieven op interactie zoals deze worden gehanteerd binnen het vakgebied van de mens-machine interactie (engels: Human-Computer Interaction – HCI). Deze variëren van formele modellen en het beschouwen van de menselijke gebruiker als een informatie-processor, tot aan fenomenologische benaderingen die interactie zien als onlosmakelijk gebonden aan de fysieke en sociale omgeving. In deze laatste zienswijze, wordt elke interactie met een digitaal artefact beïnvloed door de achtergrond en kenmerken van de gebruiker en de gebruikssituatie; deze moeten dus ook als geheel worden onderzocht.

Het zogenaamde user-centered design (UCD) is tegenwoordig de meest gebruikelijke benadering tot de studie van het ontwerpen van systemen met de menselijke gebruiker in het achterhoofd. Co-design benaderingen gaan een stap verder en betrekken de toekomstige gebruikers als actieve ontwerppartners die deelnemen aan creatieve activiteiten en de besluitvorming in het ontwerpproces. In ons werk hebben we in elke onderzoeksfase zorgvuldig bepaald welke methoden te gebruiken. Deze omvatten kwantitatieve methoden, b.v. enquêtes om de houding van gebruikers te onderzoeken, en kwalitatieve methoden, zoals prototyping of gebruikersworkshops met experts en gebruikers voor het ontwerpen van ideeën en het onderzoeken van de behoeften van de gebruikers.

In de beginfase van het onderzoek, gepresenteerd in hoofdstuk 3, hebben we ons gericht op het verkennen van de ontwerpruimte voor DSS. We waren benieuwd in welke gevallen een onderhandeling-ondersteuningssysteem (negotiation support system (NSS), een speciaal geval van een DSS) sociaal zou worden geaccepteerd en welke functies deze systemen moeten hebben volgens experts en eindgebruikers. We hebben focusgroepen met deskundigen en gebruikers opgezet en een online gebruikersonderzoek gehouden. Om bij de deelnemers gedachten over verschillende gebruikscontexten in gang te zetten, hebben wij gefilmde scenario's met vijf verschillende contexten laten zien. De discussies hebben geleid tot 12 algemene ontwerprichtlijnen voor NSS.

Focus in de groepen met experts was de functionaliteit van het systeem. De experts benadrukten het belang van een goede voorbereiding voor succesvolle onderhandelingen. Het onderzoek met gebruikers wees op sociale aspecten van NSS. Met name de resultaten van de enquête tonen aan dat de context voor het gebruik een belangrijke factor is die de sociale acceptatie van NSS beïnvloed. De respondenten zouden geen gebruik maken van NSS in face-to-face situaties waarin de relatie met de tegenstander belangrijk is. Wanneer de relatie minder belangrijk is, is het acceptabeler.

Verrassend is dat de subjectieve norm de belangrijkste factor was die de intentie om een NSS te gebruiken, beïnvloed. Mensen gaan vooral af op meningen van hun familie en vrienden als zij beslissen of ze een NSS gebruiken. Samengevat leerden we dat de voorbereidingsfase cruciaal is voor een goed onderhandelresultaat. Daarnaast zou het systeem een begeleider bij het onderhandelingsproces moeten zijn en functionaliteit die overleg met een sociaal netwerk mogelijk maakt, zou voordelig zijn.

Concrete ontwerpen van digitale tools voor de voorbereidingsfase van een beslissingsproces was de focus bij de volgende stappen in ons onderzoek. Hoofdstuk 4 beschrijft een aantal studies die onderzoeken hoe interfaces voor het bepalen van voorkeuren moeten worden ontworpen. Een probleem met huidige systemen voor

de elicitatie van voorkeuren is dat deze geen rekening houden met de constructieve aard van menselijke voorkeuren. Zulke interfaces zijn vaak gebaseerd op rationele, kwantitatieve modellen die niet overeenkomen met de mentale modellen van mensen. Om dit gat tussen cognitieve vermogens van mensen en de gebruiker-systeeminteractie te dichten, voerden wij drie studies uit met betrekking tot (1) verschillende manieren van invoeren van voorkeuren, (2) factoren die van invloed zijn voor de motivatie van gebruikers om gedetailleerde voorkeuren in te voeren en (3) het participatieve ontwerpen van interfaces die ondersteuning bieden voor het menselijke proces van bepalen van voorkeuren. Gebaseerd op de resultaten hebben wij vier ontwerprichtlijnen voor user-centered elicitatie van voorkeuren ontwikkeld. Een andere bijdrage is het prototype dat wij hebben ontworpen met eindgebruikers en hebben geïmplementeerd in een NSS en getest in bruikbaarheidsstudies.

Bij het eliciteren van voorkeuren hebben wij gezien dat een van de grootste moeilijkheden in het constructieve karakter van menselijke voorkeuren ligt. Dit betekent dat de voorkeuren kunnen worden beïnvloed door de context waarin de beslissing genomen wordt, dat voorkeuren over de tijd kunnen veranderen, en dat mensen vaak hun voorkeuren niet precies weten.

Om mensen te ondersteunen om een nauwkeurig voorkeurmodel te construeren in een gegeven context en de risico's van onvoorziene gevolgen te beperken, hebben wij gezien dat onderliggende waarden een nuttig concept zijn om aandacht aan te besteden aan het begin van de elicitatie van voorkeuren. Persoonlijke waarden zijn stabiel over langere periodes in het leven van mensen en kunnen gezien worden als leidende principes. Door het benadrukken van waarden in de besluitvorming en het delen van gedachten over waarden met onderhandelingspartners kunnen meer bevredigende resultaten worden bereikt. Mensen hebben echter complexe systemen van waarden die moeilijk zijn om te articuleren of te beoordelen. Momenteel zijn er geen digitale hulpmiddelen beschikbaar die mensen ondersteunen bij het bepalen van hun waarden. Hoofdstuk 5 behandelt daarom het ontwerp van een dergelijke tool. Op basis van deskundigenadvies hebben wij een prototype voor reflectie op waarden ontworpen, dat wij hebben gebruikt in de volgende gebruikerstudies en een online enquête. Gebaseerd op de resultaten hebben we een lijst van ontwerprichtlijnen samengesteld voor digitale tools voor de elicitatie van waarden.

Tijdens het ontwerpen hebben wij gewerkt met coöperatieve methoden waarbij eindgebruikers en andere belanghebbenden meededen aan creatieve activiteiten, dat wil zeggen het ontwikkelen van prototypes of ontwerpideeën. We waren geïnteresseerd in hoe we mensen kunnen helpen om zich een toekomstig systeem voor te stellen en

deel te nemen aan creatieve activiteiten zonder een formele ontwerpopleiding. We hebben een ontwerpmethode ontwikkeld waarbij gebruikers en experts door middel van herhaaldelijke compositie van bestaande en nieuwe elementen kunnen deelnemen aan de ontwikkeling van een prototype. Verder hebben we, verspreid over meerdere design workshops – uitgevoerd in onderzoek voor, of verbonden met dit proefschrift – gevarieerd in de geboden ondersteuning, de groepssamenstelling en de triggers voor creativiteit, om zo de invloed van deze factoren te kunnen bepalen.

Een belangrijk aspect van design-based onderzoek is een kritische reflectie op het proces en de resultaten. Hoofdstuk 6 biedt een reflectie op onze methodologische keuzes. Samengevat hebben we bevonden dat het gunstig is om in design workshops mensen te betrekken met verschillende rollen (de eindgebruikers, domeindeskundigen en ontwerpers). Aangezien deze groepen mensen verschillende materialen verschillend gebruiken, zou het nuttig zijn om prototypes als zogenaamde boundary-objecten te verschaffen om creativiteit en begrip te versterken. Bovendien kunnen interface-elementen of papieren prototypes en een gestructureerd brainstormproces helpen om nieuwe ontwerpideeën te genereren.

Zoals uitgewerkt in hoofdstuk 7 zijn de bijdragen van het proefschrift aan HCI en DSS onderzoek drieledig: (1) belangrijke overwegingen voor DSS ontwerp, (2) een lijst van concrete ontwerprichtlijnen, en (3) een analyse van geschikte methoden met betrekking tot de betrokkenheid en empowerment van de belanghebbenden in het ontwerpproces. Overwegingen voor het ontwerp kwamen uit literatuuronderzoek in verschillende onderzoeksdisciplines (psychologie, gedragswetenschappen, beslissingstheorieën, kunstmatige intelligentie en mens-computer interactie) en exploratieve studies met domein experts en gebruikers. Deze wezen op het belang van de voorbereidingsfase van besluitvorming en onderhandelingen en de sociale factoren van het proces. Ontwerprichtlijnen werden afgeleid van het ontwerpgebaseerde onderzoek met deskundigen en gebruikers voor het volledige ontwerp van NSS en in het bijzonder het bepalen van voorkeuren en reflectie op onderliggende waarden. Een overheersend onderwerp was de balans tussen enerzijds het ondersteunen van menselijke manieren van denken en reflecteren en anderzijds het geven van intelligente begeleiding door het systeem die door systeemontwerpers wordt gecreëerd. Dit evenwicht kan alleen worden bereikt door nauw samen te werken met eindgebruikers, domeindeskundigen en ontwerpers tijdens het ontwerpproces, ondersteund door deskundige procesbegeleiders.

Met dit proefschrift roepen wij op tot een verschuiving in het DSS onderzoek van de ontwikkeling van expertsystemen die beslissingen overnemen naar het ontwerpen

van systemen die mensen helpen om hun eigen, geïnformeerde beslissingen te nemen. We hopen dat andere onderzoekers dit werk als uitgangspunt zien voor hun eigen onderzoek om beslissers een groter vermogen om te beslissen, te geven.

Zusammenfassung

Wichtige Lebensentscheidungen zu treffen, wie etwa das Anstreben einer bestimmten Karriere, ist schwierig und manchmal emotional. Während man herausfinden muss, was man für sich selbst möchte, beinhaltet der Entscheidungsprozess auch, die Konsequenzen auf längere Sicht als auch diejenigen für andere Beteiligte (z.B. Familienmitglieder) in Betracht zu ziehen. Weiterhin können diese Lebensentscheidungen auch Verhandlungen mit anderen Parteien mit sich bringen, z.B. mit dem Verkäufer des gewünschten Hauses oder den Vertretern einer Firma, in der man gerne arbeiten möchte. Weitere wichtige Aspekte des Entscheidungsprozesses sind das Erkunden und Festlegen verschiedener Alternativen und das Abwägen von Optionen auf der Basis seiner eigenen Präferenzen. Dies kann sich als schwierig herausstellen, insbesondere wenn der Entscheidungsträger ungeschult ist und mit einer großen Anzahl an Alternativen oder mit erheblichem Informationsüberfluß konfrontiert wird.

Hilfe in derartigen Entscheidungsprozessen versprechen hier Digitale Entscheidungsunterstützende Systeme (auf Englisch: decision support systems = DSS). Allerdings beschränkt sich die Forschung momentan lediglich auf wirtschaftliche Modelle und technische Lösungen für professionelle Entscheidungsträger, was die Anwendbarkeit und den Erfolg dieser Systeme im Bereich wichtiger Lebensentscheidungen einschränkt. Wir glauben, dass DSS Forschung einen menschbezogenen Ansatz nötig hat, welcher die kognitiven Fähigkeiten und Bedürfnisse der Benutzer in den Mittelpunkt stellt, um so Menschen zu unterstützen, ihre eigenen, kompetenten Entscheidungen zu treffen. Diese Dissertation befasst sich mit der Hauptforschungsfrage "Wie können wir die Interaktion zwischen dem Benutzer und dem System gestalten, um eine menschbezogene digitale Entscheidungsunterstützung zu erreichen?" Die Untersuchungen waren zweigeteilt, auf der einen Seite fokussiert auf Anforderungen und Richtlinien und auf der anderen auf einer Methodik, die die *Stakeholder* mit in den Gestaltungsprozess einbezieht. (*Kapitel 1*).

Das zentrale Konzept der vorliegenden Arbeit ist die Interaktion zwischen dem Benutzer und dem System. Wie in *Kapitel 2* beschrieben, existieren verschiedene Interaktionsperspektiven, die im Gebiet der Mensch Computer Interaktion (auf Englisch: Human Computer Interaction = HCI) entwickelt wurden. Diese reichen von Ansätzen zur Formalisierung von Interaktion, welche den Benutzer als Informationsprozessor betrachten, hin zu phänomenologischen Ansätzen, welche Interaktion immer im Verband mit dem physikalischen und sozialen Umfeld betrachten. Im letzteren Ansatz beeinflußt der Hintergrund des Benutzers sowie seine Eigenschaften und die gegebene

Nutzungssituation jede Interaktion mit einem digitalen Gegenstand und muss daher zusammen erforscht werden. Benutzerorientierte Designansätze sind heutzutage Gang und Gäbe, um zu erforschen, wie Systeme mit Blick auf den Benutzer gestaltet werden müssen. Zusätzlich gibt es das Co-design, welches die zukünftigen Benutzer und auch andere Stakeholder involviert. Diese werden zu aktiven Designpartnern, die an kreativen Aktivitäten teilnehmen und mitentscheiden. Während der vorliegenden Arbeit haben wir in jeder Forschungsphase sorgsam Methoden selektiert, um unsere jeweiligen Fragen zu beantworten. Diese beinhalteten quantitative Methoden, wie z.B. Umfragen, die es uns erlaubten, die Meinungen größerer Personenkreise zu erfragen, sowie qualitative Methoden, wie z.B. das Erstellen von Prototypen und das Leiten von Designworkshops, um Designideen und Benutzerbedürfnisse zusammen mit Experten und Benutzern zu erkunden.

In der ersten Forschungsphase, präsentiert in *Kapitel 3*, haben wir uns darauf beschränkt, die Gestaltungsmöglichkeiten für DSS, in diesem Fall eine spezielle Art von DSS, nämlich ein verhandlungsunterstützendes System (auf Englisch: negotiation support system = NSS) zu erkunden. Der Fokus lag auf einem NSS, welches einer Verhandlungspartei hilft, sich auf die Verhandlung vorzubereiten und mit der Gegenpartei zu verhandeln. Wir waren insbesondere neugierig, in welchen Situationen es sozial akzeptabel wäre, ein NSS zu verwenden und welche Funktionen solch ein System gemäß der Experten und Benutzer haben sollte. Wir haben Fokusgruppen mit Experten und Benutzern sowie einen online Fragebogen in unseren Studien benutzt. Als Denkanstoß dienten verfilmte Szenarios, die insgesamt fünf verschiedene Nutzungskontexte porträtierten. Im Mittelpunkt der Expertendiskussion stand die Funktionalität des Systems. Basiert auf dieser Diskussion haben wir 12 allgemeine Gestaltungsrichtlinien für NSS entwickelt. Diese heben den besonderen Stellenwert einer guten Vorbereitung für erfolgreiche Verhandlungen hervor.

Die Fokusgruppen mit Benutzern und der Fragebogen wiesen soziale Aspekte auf. Insbesondere die Ergebnisse des Fragebogens zeigen, dass der Nutzungskontext ein wichtiger Indikator für die Akzeptanz von NSS ist. Die Befragten würden z.B. ein NSS nicht in persönlicher Kommunikation mit dem Verhandlungspartner einsetzen, falls die Beziehung zu diesem wichtig ist. Ist die Beziehung weniger wichtig, z.B. bei einem Autokauf, ist die Akzeptanz größer. Erstaunlicherweise ist die sogenannte subjektive Norm der dominanteste Indikator für die Intention des Benutzers, ein NSS zu benutzen. Menschen schätzen die Meinungen ihrer Anvertrauten und fragen diese auch um Rat, wenn es darum geht, ein NSS zu benutzen oder grundsätzlich eine Entscheidung zu treffen. Zusammenfassend läßt sich sagen, dass die Vorbereitungsphase ausschlaggebend ist für das Verhandlungsergebnis. Des Weiteren, sollte ein

System als Partner im Verhandlungsprozess fungieren und die entsprechende soziale Funktionalität (z.B. in Verbindung mit sozialen Netzwerken) in Betracht gezogen werden.

Das konkrete Gestalten von digitalen *Tools* zur Vorbereitung stand im Mittelpunkt der folgenden Schritte unserer Forschung. *Kapitel 4* beschreibt mehrere Studien zur Erforschung der Gestaltung von Benutzerschnittstellen zur Feststellung und Eingabe von Präferenzen – ein wichtiger Aspekt der Vorbereitung. Ein Problem aktueller Systeme in dieser Hinsicht ist, dass sie nicht an die konstruktive Natur menschlicher Präferenzen angepaßt sind und stattdessen auf rationalen, quantitativen Modellen beruhen, die nicht zu den mentalen Modellen der Menschen passen. Um diese Lücke zwischen menschlichen kognitiven Fähigkeiten und der Interaktion mit dem System zu schließen, haben wir drei Studien ausgeführt, die sich mit den folgenden drei Aspekten befassten: (1) Verschiedene Arten, Präferenzen einzugeben, (2) Faktoren die den Benutzer motivieren, mehr Details preiszugeben und (3) *Participatory Design* von Benutzerschnittstellen, die den menschlichen Prozess zur Feststellung von Präferenzen unterstützen.

Die erste Studie zeigte, dass bei Methoden mit ähnlichen Eingaben, z.B. Bewertungen oder Anordnungen, der Aufwand und das Gefallen der Methode grundsätzlich negativ korreliert waren. Dennoch gefielen den Teilnehmern der Studie diejenigen Methoden besser, bei denen die Navigation und die Art des Feedbacks raffinierter waren mit Blick auf die Aussagekraft und das Verständnis der eigenen Präferenzen. Die Ergebnisse zeigten aber auch, dass diese Methoden gleichzeitig einen erhöhten Aufwand erfordern und schwieriger zu gebrauchen waren. Ein Folgeexperiment zeigte, dass Vertrautheit und eine vorgefertigte Meinung über ein Thema, Menschen dazu bewegte, detaillierteres Feedback abzugeben. Interessanterweise bevorzugten die Teilnehmer, die mehr Details eingaben, hauptsächlich (zweidimensionales) emotionales Feedback. Im Mittelpunkt der dritten Studie stand die Gestaltung konkreter Benutzungsschnittstellen mit Hilfe von Co-Design mit Endbenutzern. Aus der Studie ergab sich, dass bei der Eingabe von Präferenzen anstelle des Beantwortens von Fragen des Systems das Verständnis der eigenen Präferenzen und die Kontrolle über das Erstellen des eigenen Profils im Vordergrund steht. Insbesondere das Betrachten der Präferenzen aus verschiedenen Sichtweisen einschließlich zugrunde liegender Werte und Konsequenzen (als Anordnung von Entscheidungsergebnissen) innerhalb derselben Benutzerschnittstelle ermöglicht es Menschen, ihre Präferenzen zu verdeutlichen. Den Teilnehmern gefielen Gestaltungselemente, die diese Betrachtung auf natürliche Art und Weise unterstützen und ein direktes visuelles Feedback geben. Die aus dieser Studie resultierenden Prototypen stellen einen Forschungsbeitrag dar

und wurden und in einem NSS implementiert und in auf ihre Nutzbarkeit hin getestet. Der Hauptbeitrag besteht aus vier Gestaltungsrichtlinien zur benutzerorientierten Erstellung von Präferenzen.

In der Arbeit mit Präferenzen haben wir gesehen, dass die konstruktive Natur von menschlichen Präferenzen eine Schwierigkeiten für die Eingabe in Systemen darstellt. Der Ausdruck *konstruktiv* bezieht sich in diesem Zusammenhang darauf, dass Präferenzen leicht vom gegebenen Entscheidungskontext beeinflusst werden, dass sie sich über die Zeit hinweg ändern können und dass Menschen ihre Präferenzen nicht immer wissen. Wir haben festgestellt, dass die zugrunde liegende Wertvorstellung ein sinnvolles Konzept zur Erstellung von Präferenzen ist, das Menschen dabei hilft, ein akkurates Präferenzenmodell in einem gegebenen Kontext zu erstellen und das Risiko unvorhergesehener Konsequenzen zu verringern. Persönliche Werte erweisen sich über längere Zeiträume als stabiler und können als leitende Prinzipien angesehen werden. Das Einbeziehen von Werten in Entscheidungen und das Teilen von Wertvorstellungen mit dem Verhandlungspartner führt zu zufriedenstellenderen Ergebnissen. Allerdings sind Wertvorstellungen oft komplex und Menschen haben Schwierigkeiten, diese Komplexität zu äußern. Zurzeit existieren keine digitalen Hilfen zur Bestimmung von Werteprofilen, die Entscheidungsfindung unterstützen.

Kapitel 5 befasst sich daher mit der Gestaltung einer solchen digitalen Hilfe. Wir haben einen auf Expertenwissen basierten Prototypen zum Reflektieren von Werten entworfen. Dieser Prototyp wurde in folgenden Nutzerstudien und einer Online-Umfrage benutzt, um eine Reihe wichtiger Gestaltungsanregungen zu identifizieren. Ein herausstehender Aspekt ist die Einzigartigkeit der Menschen im Hinblick auf Selbstreflektion. Ausschlaggebend ist, dass digitale Tools zur Wertereflektion viele verschiedene Arten der Reflektion anbieten, um es den Benutzern zu erlauben, diejenigen zu selektieren, die ihnen zusagen. Außerdem muss das System den Benutzern erlauben, ihre eigenen Werte hinzuzufügen. Zum Starten des Reflektionsprozesses eignen sich besonders Methoden die emotionale Reaktionen hervorrufen. Weiterhin ist es wichtig, dass das System ein persönliches Gefühl vermittelt, was auch zum Aufbau von Vertrautheit dient, die zur Reflektion über persönliche Dinge wichtig ist. Vertrauen kann aufgebaut werden durch Rücksicht auf die Privatsphäre des Nutzers, durch einen auf Dialog basierten Interaktionsstil mit dem System und durch Transparenz, die dem Nutzer erlaubt, direkt zu erkennen, wie sein Werteprofil bei der Eingabe einer Reflektion angepasst wird. Eine interessante Diskussion entstand rund um das Thema inwiefern das System den Nutzer lenkt. In manchen Fällen ist dies wünschenswert, z.B. um von einer Reflektion zu einem abstrakten Wert zu gelangen, in anderen hindert es die persönliche Art und Weise zu reflektieren. Anhand der

Resultate unserer Studien haben wir eine Liste mit Gestaltungsrichtlinien für digitale Tools zur Werteerhebung erstellt.

Während unserer Designarbeit haben wir uns auf kooperative Methoden beschränkt, die Endbenutzer und andere Stakeholder in kreative Aktivitäten einbinden, d.h. Prototypen erstellen oder Gestaltungsideen entwickeln. Wir waren insbesondere daran interessiert, wie man Menschen darin unterstützen kann, sich ein zukünftiges System vorzustellen und sich auch ohne eine formelle Designausbildung an kreativen Aktivitäten zu beteiligen. Wir haben eine Designmethode entwickelt, die Menschen beim kompositionalen Entwickeln von Prototypen einbezieht. Weiterhin haben wir die Unterstützung, Gruppenkomposition und kreative Auslöser in Designworkshops in verschiedenenen Aktivitäten im Rahmen dieser Dissertation variiert. Ein wichtiger Aspekt designbasierter Forschung ist es, den Prozess und die Ergebnisse im Nachhinein kritisch zu begutachten. In *Kapitel 6* betrachten wir unsere methodischen Entscheidungen rückblickend. Zusammengefasst lässt sich sagen, dass es günstig ist, Teilnehmer mit verschiedenen Rollen, d.h. Endbenutzer, Experten und Designer, in Designworkshops zu involvieren. Weil diese Gruppen unterschiedlich mit Materialien umgehen, ist es sinnvoll, Prototypen als sogenannte *boundary objects* zur Auslösung kreativer Erkundungen und zur Förderung eines geteilten Verständnisses einzusetzen. Außerdem können sowohl einzelne Gestaltungselemente und Papierprototypen als auch ein strukturierter Brainstormingprozess Menschen unterstützen, neue Ideen zu generieren.

Wie in *Kapitel 7* beschrieben liefert diese Dissertation drei Forschungsbeiträge auf den Gebieten der Mensch Computer Interaktion und der entscheidungsunterstützenden Systeme: (1) wichtige Gestaltungseinsichten für DSS, (2) konkrete Gestaltungsrichtlinien und (3) eine Analyse von angemessen Methoden mit Blick auf das Engagement und Unterstützung von Stakeholdern im Designprozess. Die Gestaltungseinsichten basieren auf Literatur aus verschiedenen Fachgebieten (Psychologie, Verhaltensforschung, Entscheidungstheorien, Künstliche Intelligenz und Mensch Computer Interaktion) sowie auf Studien mit Experten und Endbenutzern. Diese haben die Vorbereitungsphase von Verhandlungen und Entscheidungsfindung sowie die sozialen Faktoren des Prozesses als ausschlaggebend hervorgehoben. Die Gestaltungsrichtlinien für die allgemeine Gestaltung von NSS und insbesondere für die Erstellung von Präferenzprofilen und Wertereflektion wurden durch designbasierte Forschung, die sowohl Experten als auch Endbenutzer in verschiedenen Phasen mit einbezogen, erhoben. Ein dominantes Thema war die Balance zwischen der Unterstützung menschlicher Denkweisen sowie die intelligente Führung des Systems. Diese Balance kann nur durch enge und iterative Interaktion mit Endbenutzern, Experten und

Designern im Designprozess zustande kommen.

Diese Dissertation ruft dazu auf, den Fokus der Forschung entscheidungsunterstützender Systeme zu verlagern. Statt Expertensysteme zu entwickeln, die dem Nutzer Entscheidungen abnehmen, sollten wir menschenbezogene Systeme gestalten, die Menschen helfen, ihre eigenen, informierten Entscheidungen zu treffen.Wir hoffen, dass andere diese Arbeit als einen Startpunkt ihrer eigenen Forschung ansehen, um zukünftig Entscheidungsträgern unter die Arme zu greifen.

Curriculum Vitae

Alina Pommeranz was born on December 07, 1981 in Bochum, Germany. In 2001 she received her High School degree from Hellweg-Schule (Grammar School), Bochum, Germany. She continued her education in computer science with a major in media technology at the University of Applied Science Gelsenkirchen, Germany, where she received a Graduate Degree in Computer Science (in German: Diplom-Informatikerin (FH)) in 2005. During her studies she performed an internship of seven months at the Digital Media Center at the University of Wollongong, Australia in 2004.

After graduation Alina worked as an IT consultant at Triestram & Partner in Bochum before she went to Stockholm, Sweden to pursue Master of Science studies in Interactive Systems Engineering at the Royal Institute of Technology. Her thesis work on ”Exploring and designing for emotional closeness between friends” was carried out at the Mobile Life Centre, Stockholm, under the supervision of Prof. Dr. Kristina Höök.

After receiving her Master of Science degree in 2008, Alina started her Ph.D. studies in Human-Computer Interaction at Delft University of Technology, the Netherlands. As part of her Ph.D. she performed an internship of three months at the Value Sensitive Design Lab at University of Washington, Seattle, USA, under the supervision of Prof. Dr. Batya Friedman.

List of SIKS Dissertations – SIKS Dissertatiereeks

====
2009
====

2009-01 Rasa Jurgelenaite (RUN)
Symmetric Causal Independence Models

2009-02 Willem Robert van Hage (VU)
Evaluating Ontology-Alignment Techniques

2009-03 Hans Stol (UvT)
A Framework for Evidence-based Policy Making Using IT

2009-04 Josephine Nabukenya (RUN)
Improving the Quality of Organisational Policy Making using Collaboration Engineering

2009-05 Sietse Overbeek (RUN)
Bridging Supply and Demand for Knowledge Intensive Tasks - Based on Knowledge, Cognition, and Quality

2009-06 Muhammad Subianto (UU)
Understanding Classification

2009-07 Ronald Poppe (UT)
Discriminative Vision-Based Recovery and Recognition of Human Motion

2009-08 Volker Nannen (VU)
Evolutionary Agent-Based Policy Analysis in Dynamic Environments

2009-09 Benjamin Kanagwa (RUN)
Design, Discovery and Construction of Service-oriented Systems

2009-10 Jan Wielemaker (UVA)
Logic programming for knowledge-intensive interactive applications

2009-11 Alexander Boer (UVA)
Legal Theory, Sources of Law & the Semantic Web

2009-12 Peter Massuthe (TUE, Humboldt-Universitaet zu Berlin)
perating Guidelines for Services

2009-13 Steven de Jong (UM)
Fairness in Multi-Agent Systems

2009-14 Maksym Korotkiy (VU)
From ontology-enabled services to service-enabled ontologies (making ontologies work in e-science with ONTO-SOA)

2009-15 Rinke Hoekstra (UVA)
Ontology Representation - Design Patterns and Ontologies that Make Sense

2009-16 Fritz Reul (UvT)
New Architectures in Computer Chess

2009-17 Laurens van der Maaten (UvT)
Feature Extraction from Visual Data

2009-18 Fabian Groffen (CWI)
Armada, An Evolving Database System

2009-19 Valentin Robu (CWI)
Modeling Preferences, Strategic Reasoning and Collaboration in Agent-Mediated Electronic Markets

2009-20 Bob van der Vecht (UU)
Adjustable Autonomy: Controling Influences on Decision Making

2009-21 Stijn Vanderlooy (UM)
Ranking and Reliable Classification

2009-22 Pavel Serdyukov (UT)
Search For Expertise: Going beyond direct evidence

2009-23 Peter Hofgesang (VU)
Modelling Web Usage in a Changing Environment

2009-24 Annerieke Heuvelink (VUA)
Cognitive Models for Training Simulations

2009-25 Alex van Ballegooij (CWI)
"RAM: Array Database Management through Relational Mapping"

2009-26 Fernando Koch (UU)
An Agent-Based Model for the Development of Intelligent Mobile Services

2009-27 Christian Glahn (OU)
Contextual Support of social Engagement and Reflection on the Web

2009-28 Sander Evers (UT)
Sensor Data Management with Probabilistic Models

2009-29 Stanislav Pokraev (UT)
Model-Driven Semantic Integration of Service-Oriented Applications

2009-30 Marcin Zukowski (CWI)
Balancing vectorized query execution with bandwidth-optimized storage

2009-31 Sofiya Katrenko (UVA)
A Closer Look at Learning Relations from Text

2009-32 Rik Farenhorst (VU) and Remco de Boer (VU)
Architectural Knowledge Management: Supporting Architects and Auditors

2009-33 Khiet Truong (UT)
How Does Real Affect Affect Affect Recognition In Speech?

2009-34 Inge van de Weerd (UU)

Advancing in Software Product Management: An Incremental Method Engineering Approach

2009-35 Wouter Koelewijn (UL)
Privacy en Politiegegevens; Over geautomatiseerde normatieve informatie-uitwisseling

2009-36 Marco Kalz (OUN)
Placement Support for Learners in Learning Networks

2009-37 Hendrik Drachsler (OUN)
Navigation Support for Learners in Informal Learning Networks

2009-38 Riina Vuorikari (OU)
Tags and self-organisation: a metadata ecology for learning resources in a multilingual context

2009-39 Christian Stahl (TUE, Humboldt-Universitaet zu Berlin)
Service Substitution – A Behavioral Approach Based on Petri Nets

2009-40 Stephan Raaijmakers (UvT)
Multinomial Language Learning: Investigations into the Geometry of Language

2009-41 Igor Berezhnyy (UvT)
Digital Analysis of Paintings

2009-42 Toine Bogers
Recommender Systems for Social Bookmarking

2009-43 Virginia Nunes Leal Franqueira (UT)
Finding Multi-step Attacks in Computer Networks using Heuristic Search and Mobile Ambients

2009-44 Roberto Santana Tapia (UT)
Assessing Business-IT Alignment in Networked Organizations

2009-45 Jilles Vreeken (UU)
Making Pattern Mining Useful

2009-46 Loredana Afanasiev (UvA)
Querying XML: Benchmarks and Recursion

====
2010
====

2010-01 Matthijs van Leeuwen (UU)
Patterns that Matter

2010-02 Ingo Wassink (UT)
Work flows in Life Science

2010-03 Joost Geurts (CWI)
A Document Engineering Model and Processing Framework for Multimedia documents

2010-04 Olga Kulyk (UT)
Do You Know What I Know? Situational Awareness of Co-located Teams in Multidisplay Environments

2010-05 Claudia Hauff (UT)
Predicting the Effectiveness of Queries and Retrieval Systems

2010-06 Sander Bakkes (UvT)
Rapid Adaptation of Video Game AI

2010-07 Wim Fikkert (UT)
Gesture interaction at a Distance

2010-08 Krzysztof Siewicz (UL)
Towards an Improved Regulatory Framework of Free Software. Protecting user freedoms in a world of software communities and eGovernments

2010-09 Hugo Kielman (UL)
A Politiele gegevensverwerking en Privacy, Naar een effectieve waarborging

2010-10 Rebecca Ong (UL)
Mobile Communication and Protection of Children

2010-11 Adriaan Ter Mors (TUD)
The world according to MARP: Multi-Agent Route Planning

2010-12 Susan van den Braak (UU)
Sensemaking software for crime analysis

2010-13 Gianluigi Folino (RUN)
High Performance Data Mining using Bio-inspired techniques

2010-14 Sander van Splunter (VU)
Automated Web Service Reconfiguration

2010-15 Lianne Bodenstaff (UT)
Managing Dependency Relations in Inter-Organizational Models

2010-16 Sicco Verwer (TUD)
Efficient Identification of Timed Automata, theory and practice

2010-17 Spyros Kotoulas (VU)
Scalable Discovery of Networked Resources: Algorithms, Infrastructure, Applications

2010-18 Charlotte Gerritsen (VU)
Caught in the Act: Investigating Crime by Agent-Based Simulation

2010-19 Henriette Cramer (UvA)
People's Responses to Autonomous and Adaptive Systems

2010-20 Ivo Swartjes (UT)
Whose Story Is It Anyway? How Improv Informs Agency and Authorship of Emergent Narrative

2010-21 Harold van Heerde (UT)

Privacy-aware data management by means of data degradation

2010-22 Michiel Hildebrand (CWI)
End-user Support for Access to Heterogeneous Linked Data

2010-23 Bas Steunebrink (UU)
The Logical Structure of Emotions

2010-24 Dmytro Tykhonov
Designing Generic and Efficient Negotiation Strategies

2010-25 Zulfiqar Ali Memon (VU)
Modelling Human-Awareness for Ambient Agents: A Human Mindreading Perspective

2010-26 Ying Zhang (CWI)
XRPC: Efficient Distributed Query Processing on Heterogeneous XQuery Engines

2010-27 Marten Voulon (UL)
Automatisch contracteren

2010-28 Arne Koopman (UU)
Characteristic Relational Patterns

2010-29 Stratos Idreos(CWI)
Database Cracking: Towards Auto-tuning Database Kernels

2010-30 Marieke van Erp (UvT)
Accessing Natural History - Discoveries in data cleaning, structuring, and retrieval

2010-31 Victor de Boer (UVA)
Ontology Enrichment from Heterogeneous Sources on the Web

2010-32 Marcel Hiel (UvT)
An Adaptive Service Oriented Architecture: Automatically solving Interoperability Problems

2010-33 Robin Aly (UT)
Modeling Representation Uncertainty in Concept-Based Multimedia Retrieval

2010-34 Teduh Dirgahayu (UT)
Interaction Design in Service Compositions

2010-35 Dolf Trieschnigg (UT)
Proof of Concept: Concept-based Biomedical Information Retrieval

2010-36 Jose Janssen (OU)
Paving the Way for Lifelong Learning; Facilitating competence development through a learning path specification

2010-37 Niels Lohmann (TUE)
Correctness of services and their composition

2010-38 Dirk Fahland (TUE)
From Scenarios to components

2010-39 Ghazanfar Farooq Siddiqui (VU)
Integrative modeling of emotions in virtual agents

2010-40 Mark van Assem (VU)
Converting and Integrating Vocabularies for the Semantic Web

2010-41 Guillaume Chaslot (UM)
Monte-Carlo Tree Search

2010-42 Sybren de Kinderen (VU)
Needs-driven service bundling in a multi-supplier setting - the computational e3-service approach

2010-43 Peter van Kranenburg (UU)
A Computational Approach to Content-Based Retrieval of Folk Song Melodies

2010-44 Pieter Bellekens (TUE)
An Approach towards Context-sensitive and User-adapted Access to Heterogeneous Data Sources, Illustrated in the Television Domain

2010-45 Vasilios Andrikopoulos (UvT)
A theory and model for the evolution of software services

2010-46 Vincent Pijpers (VU)
e3alignment: Exploring Inter-Organizational Business-ICT Alignment

2010-47 Chen Li (UT)
Mining Process Model Variants: Challenges, Techniques, Examples

2010-48 Milan Lovric (EUR)
Behavioral Finance and Agent-Based Artificial Markets

2010-49 Jahn-Takeshi Saito (UM)
Solving difficult game positions

2010-50 Bouke Huurnink (UVA)
Search in Audiovisual Broadcast Archives

2010-51 Alia Khairia Amin (CWI)
Understanding and supporting information seeking tasks in multiple sources

2010-52 Peter-Paul van Maanen (VU)
Adaptive Support for Human-Computer Teams: Exploring the Use of Cognitive Models of Trust and Attention

2010-53 Edgar Meij (UVA)
Combining Concepts and Language Models for Information Access

====
2011
====

2011-01 Botond Cseke (RUN)
Variational Algorithms for Bayesian Inference in Latent Gaussian Models

2011-02 Nick Tinnemeier(UU)
Organizing Agent Organizations. Syntax and Operational Semantics of an Organization-Oriented Programming Language

2011-03 Jan Martijn van der Werf (TUE)
Compositional Design and Verification of Component-Based Information Systems

2011-04 Hado van Hasselt (UU)
Insights in Reinforcement Learning; Formal analysis and empirical evaluation of temporal-difference learning algorithms

2011-05 Base van der Raadt (VU)
Enterprise Architecture Coming of Age - Increasing the Performance of an Emerging Discipline.

2011-06 Yiwen Wang (TUE)
Semantically-Enhanced Recommendations in Cultural Heritage

2011-07 Yujia Cao (UT)
Multimodal Information Presentation for High Load Human Computer Interaction

2011-08 Nieske Vergunst (UU)
BDI-based Generation of Robust Task-Oriented Dialogues 2011-09 Tim de Jong (OU)
Contextualised Mobile Media for Learning

2011-10 Bart Bogaert (UvT)
Cloud Content Contention

2011-11 Dhaval Vyas (UT)
Designing for Awareness: An Experience-focused HCI Perspective

2011-12 Carmen Bratosin (TUE)
Grid Architecture for Distributed Process Mining

2011-13 Xiaoyu Mao (UvT)
Airport under Control. Multiagent Scheduling for Airport Ground Handling

2011-14 Milan Lovric (EUR)
Behavioral Finance and Agent-Based Artificial Markets

2011-15 Marijn Koolen (UvA)
The Meaning of Structure: the Value of Link Evidence for Information Retrieval

2011-16 Maarten Schadd (UM)
Selective Search in Games of Different Complexity

2011-17 Jiyin He (UVA)
Exploring Topic Structure: Coherence, Diversity and Relatedness

2011-18 Mark Ponsen (UM)
Strategic Decision-Making in complex games

2011-19 Ellen Rusman (OU)
The Mind ' s Eye on Personal Profiles

2011-20 Qing Gu (VU)
Guiding service-oriented software engineering - A view-based approach

2011-21 Linda Terlouw (TUD)
Modularization and Specification of Service-Oriented Systems

2011-22 Junte Zhang (UVA)
System Evaluation of Archival Description and Access

2011-23 Wouter Weerkamp (UVA)
Finding People and their Utterances in Social Media

2011-24 Herwin van Welbergen (UT)
Behavior Generation for Interpersonal Coordination with Virtual Humans On Specifying, Scheduling and Realizing Multimodal Virtual Human Behavior

2011-25 Syed Waqar ul Qounain Jaffry (VU)
Analysis and Validation of Models for Trust Dynamics

2011-26 Matthijs Aart Pontier (VU)
Virtual Agents for Human Communication - Emotion Regulation and Involvement-Distance Trade-Offs in Embodied Conversational Agents and Robots

2011-27 Aniel Bhulai (VU)
Dynamic website optimization through autonomous management of design patterns

2011-28 Rianne Kaptein(UVA)
Effective Focused Retrieval by Exploiting Query Context and Document Structure

2011-29 Faisal Kamiran (TUE)
Discrimination-aware Classification

2011-30 Egon van den Broek (UT)
Affective Signal Processing (ASP): Unraveling the mystery of emotions

2011-31 Ludo Waltman (EUR)
Computational and Game-Theoretic Approaches for Modeling Bounded Rationality

2011-32 Nees-Jan van Eck (EUR)
Methodological Advances in Bibliometric Mapping of Science

2011-33 Tom van der Weide (UU)
Arguing to Motivate Decisions

2011-34 Paolo Turrini (UU)
Strategic Reasoning in Interdependence: Logical and Game-theoretical Investigations

2011-35 Maaike Harbers (UU)
Explaining Agent Behavior in Virtual Training

2011-36 Erik van der Spek (UU)
Experiments in serious game design: a cognitive approach

2011-37 Adriana Burlutiu (RUN)
Machine Learning for Pairwise Data, Applications for Preference Learning and Supervised Network Inference

2011-38 Nyree Lemmens (UM)
Bee-inspired Distributed Optimization

2011-39 Joost Westra (UU)
Organizing Adaptation using Agents in Serious Games

2011-40 Viktor Clerc (VU)
Architectural Knowledge Management in Global Software Development

2011-41 Luan Ibraimi (UT)
Cryptographically Enforced Distributed Data Access Control

2011-42 Michal Sindlar (UU)
Explaining Behavior through Mental State Attribution

2011-43 Henk van der Schuur (UU)
Process Improvement through Software Operation Knowledge

2011-44 Boris Reuderink (UT)
Robust Brain-Computer Interfaces

2011-45 Herman Stehouwer (UvT)
Statistical Language Models for Alternative Sequence Selection

2011-46 Beibei Hu (TUD)
Towards Contextualized Information Delivery: A Rule-based Architecture for the Domain of Mobile Police Work

2011-47 Azizi Bin Ab Aziz(VU)
Exploring Computational Models for Intelligent Support of Persons with Depression

2011-48 Mark Ter Maat (UT)
Response Selection and Turn-taking for a Sensitive Artificial Listening Agent

2011-49 Andreea Niculescu (UT)
Conversational interfaces for task-oriented spoken dialogues: design aspects influencing interaction quality

====
2011
====

2012-01 Terry Kakeeto (UvT)
Relationship Marketing for SMEs in Uganda

2012-02 Muhammad Umair(VU)
Adaptivity, emotion, and Rationality in Human and Ambient Agent Models

2012-03 Adam Vanya (VU)
Supporting Architecture Evolution by Mining Software Repositories

2012-04 Jurriaan Souer (UU)
Development of Content Management System-based Web Applications

2012-05 Marijn Plomp (UU)
Maturing Interorganisational Information Systems

2012-06 Wolfgang Reinhardt (OU)
Awareness Support for Knowledge Workers in Research Networks

2012-07 Rianne van Lambalgen (VU)
When the Going Gets Tough: Exploring Agent-based Models of Human Performance under Demanding Conditions

2012-08 Gerben de Vries (UVA)
Kernel Methods for Vessel Trajectories

2012-09 Ricardo Neisse (UT)
Trust and Privacy Management Support for Context-Aware Service Platforms

2012-10 David Smits (TUE)
Towards a Generic Distributed Adaptive Hypermedia Environment

2012-11 J.C.B. Rantham Prabhakara (TUE) Process Mining in the Large: Preprocessing, Discovery, and Diagnostics

2012-12 Kees van der Sluijs (TUE)
Model Driven Design and Data Integration in Semantic Web Information Systems

2012-13 Suleman Shahid (UvT)
Fun and Face: Exploring non-verbal expressions of emotion during playful interactions

2012-14 Evgeny Knutov(TUE)
Generic Adaptation Framework for Unifying Adaptive Web-based Systems

2012-15 Natalie van der Wal (VU)
Social Agents. Agent-Based Modelling of Integrated Internal and Social Dynamics of Cognitive and Affective Processes

2012-16 Fiemke Both (VU)
Helping people by understanding them - Ambient Agents supporting task execution and depression treatment

2012-17 Amal Elgammal (UvT)
Towards a Comprehensive Framework for Business Process Compliance

2012-18 Eltjo Poort (VU)
Improving Solution Architecting Practices

2012-19 Helen Schonenberg (TUE)
What's Next? Operational Support for Business Process Execution

2012-20 Ali Bahramisharif (RUN)
Covert Visual Spatial Attention, a Robust Paradigm for Brain-Computer Interfacing

2012-21 Roberto Cornacchia (TUD)
Querying Sparse Matrices for Information Retrieval

2012-22 Thijs Vis (UvT)
Intelligence, politie en veiligheidsdienst: verenigbare grootheden?

2012-23 Christian Muehl (UT)
Toward Affective Brain-Computer Interfaces: Exploring the Neurophysiology of Affect during Human Media Interaction

2012-25 Laurens van der Werff (UT)
Evaluation of Noisy Transcripts for Spoken Document Retrieval

2012-25 Silja Eckartz (UT)
Managing the Business Case Development in Inter-Organizational IT Projects: A Methodology and its Application

2012-26 Emile de Maat (UvA)
Making Sense of Legal Texts

2012-27 Hayrettin Gürkök (UT)
Mind the Sheep! User Experience Evaluation & Brain-Computer Interface Games

2012-28 Nancy Pascall (UvT)
Engendering Technology Empowering Women

2012-29 Almer Tigelaar (UT)
Peer-to-Peer Information Retrieval