Published by Agora Center, University of Jyväskylä

ISSN: 1795-6889

HUMAN TECHNOLOGY

An Interdisciplinary Journal on Humans in ICT Environments

Volume 6, Number 2, November 2010

Pertti Saariluoma, Editor in Chief

www.humantechnology.jyu.fi

HUMAN TECHNOLOGY

An Interdisciplinary Journal on Humans in ICT Environments Volume 6, Number 2, November 2010

Contents

Human Technology is an interdisciplinary, scholarly journal that presents innovative, peer-reviewed articles exploring the issues and challenges surrounding human-technology interaction and the human role in all areas of our ICT-infused societies.

Human Technology is published by the Agora Center, University of Jyväskylä and distributed without a charge online.

ISSN: 1795-6889
Submissions and contact: humantechnology@jyu.fi
Managing Editor: Barbara Crawford
Technical Support: Laura Fadjukoff

www.humantechnology.jyu

HUMAN TECHNOLOGY

An Interdisciplinary Journal on Humans in ICT Environments ISSN: 1795-6889

www.humantechnology.jyu.fi Volume 6 (2), November 2010, 151–154

From the Editor in Chief

SCIENTIFIC AND DESIGN STANCES

Pertti Saariluoma
Cognitive Science, Department of Computer Science and Information Systems
University of Jyväskylä, Finland

Human technology interaction is a strange field of expertise, because both academics and industry are interested in it. And yet, every now and then, it becomes apparent that academics and industry do not always see eye to eye (Carroll, 1997). They seem to think in different manner. While scientists look for how things are, industry mostly seeks out how things should be. Indeed, sometimes two very different stances behind the basic thinking of the two important human–technology interaction (HTI) communities surface.

Scientists primarily are interested in general laws and principles, even eternal truths with no exceptions. They want to identify general laws and use them to explain individual phenomena. As an analogy, they are not satisfied with the simple assessment that a car is not working, but would prefer rather to say that the carburetor of a car broke because freezing water expands as it changes its state (Hempel, 1965). Scientists equally are concerned about finding deterministic or stochastic laws, which are valid in all circumstances (Bunge, 1967) Thus, much of scientific thinking is built upon the idea that the function of science is to produce generalizations. This way of thinking can be termed in this editorial as *scientific stance*.

In solving HTI problems, general principles regarding the human mind are very valuable. Consider the notion of limited capacity (Broadbent, 1958; Miller, 1956). When interaction problems are to be solved, the ergonomic and human factor dimensions are evident. Every cognitive ergonomist knows that it is essential to decrease mental workload and organize matters so that people can use chunking, for example.

Programming paradigms provide a good example. We have no other reason for constructing computer languages and paradigms such as structures programming or object oriented programming except to decrease mental workload by chunking. The problem is not the machine but the mind. A somewhat polemical person may point out that the complexity of the code for a machine is precisely the number of the symbols in a program; any other measure is always constructed from human perspective. The number of functions, or meaningful reserved words, for example, makes sense only to people. They have no meanings to the machines because machines do not have any meanings. Nevertheless, the importance of functions and meanings can be explained on the grounds of human's limited working memory capacity and its laws (Miller, 1956).

URN: NBN:fi:jyu-201011173089

One may ask here, where is the problem if we have general psychological laws such as picture superiority effect, which, for example, explains why graphic user interfaces make sense. The problem is that the study general psychological laws do not directly lead to useful technologies: The laws do not tell us what kind of technologies should be developed for people. This means that there must be something else hidden HTI-thinking than scientific stance.

As stated above, the difference between scientists and industry people can be seen in where they put their emphasis, and industry people place their primary attention on making something that works. Edison designed the electric lamp that worked, but also thought through all of the related infrastructure needed for the technology (Millard, 1990). He understood that many things were needed to advance the technology, while an academic of Edison's time commented that he expected the world would never hear about the device again once the electric lamp exhibit closed at the Paris World Exhibition (Cerf & Navatsky, 1984). Presumably, this person looked the electric lamp without the infrastructure that Edison was able to envision. The difference between how Edison and his academic critic thought was that Edison innovated by thinking constructively. He did not pay attention to the obstacles and difficulties, but how to remove them. This constructive attitude and way of thinking is typical of the *design stance.*

The main criterion for design thinking is not necessarily what is universally true, but what works in practice. A good example was given to me by an experienced industrial designer. He told me about a huge computer program that suddenly achieved everything they hoped it would do. His team did not fully comprehend why it worked, but the case was closed nonetheless. They decided that no one should touch the code, and they just went on. Surely this is not the only case of this kind in the world, but rather the way industry has to work. Nevertheless, it shows how proving truth and constructing technology have different criteria for success. To get something to work is the very core of the design stance.

However, design thinking cannot neglect the laws of nature nor say that the principles are meaningless. In fact, if a product or process contradicts some of law of nature, it will not work. So while a technology could be ignorant of natural laws or the laws of the human mind, it cannot break them. This is why the principles created by scientists are valuable for the designers, even if they possess different approaches to and position on the principles.

Design thinking seldom relies on a single law. Any construction can be viewed as an enormous set of solved problems but the problems can be subsumed under several types of law. This means that while scientists analytically strive to generate one law or principle at a time, designers strive to combine them under one single working idea. A design stance leads us to a specific way of constructive thinking that is typical in industry. It is also something that may be difficult to understand from academic point of view.

The goal of design is innovation. All small problem-solving processes characteristic to design industry should be combined under a single coherent frame, for example, a machine or a Web service, which then can be used by people to improve the quality of their life. In this work, some general principles of how the human mind works are more rational than others in finding solutions to perplexing problems or obvious needs. This means that general principles also can explain why one potential solution for a design problem will work better than another, which is the main characteristic of *explanatory design thinking* (Saariluoma & Oulasvirta, 2010).

Interestingly, very little explanatory thinking is applied in human technology interaction design! When we look at the field of engineering, for instance, it is very common in

mechanical and software engineering for designs to be founded on the laws of nature or principles of mathematics. However, in user interface or general interaction design, solutions are generally intuitive and corrected through testing. Nevertheless, explanatory thinking would aid in bridging the gulf between scientific and design stances.

In this issue, we have a number of design-oriented publications. To very strict adherers of the scientific stance some aspects of the papers in this issue may look somewhat loose, but we still think that it is important to foster discussion and publish these papers with many very original ideas. Indeed, if we do not present design-oriented thinking, we can hardly think and rethink the issues: We simply do not see the issues. Let's think, for example, of Nielsen's (2000) famous principle of five subjects, which states that only five subjects are sufficient to test industrial usability. This principle has received much attention and criticism (Bevan et al., 2003). However, if Nielsen had not called our attention to the issue, we would have today a much poorer understanding of how to construct usability experiments. Indeed, we can see here that design problems can pave the way to scientific problems, analyses, discussion, and theories. The interaction between design and science is not a one-way street.

We begin our issue with a paper by **Aguierre-Urreta and Marakas**, who investigate the role of gender in technology adoption. In particular, they look at the psychological mechanisms that impact technology acceptance and do so through the novel use of a choice between viable technologies. Next, **Solves Pujol and Umemuro** present a new stream of research focused on affective technology, that is, technologies that support and encourage emotional interaction via technology. Their focus is on love, specifically productive love, embodied in eight principles that can guide technology development. They provide a pretest of one such technology as an illustration of how theoretically and empirically derived principles can support technology development aimed at promoting productive love. **Bergvall-Kåreborn and Ståhlbröst** demonstrate how user expressions regarding a service can be translated through qualitative research into requirements for a particular technology. Drawing on focus group data, these authors found that user requirements differed, depending upon the users' need *of* the service as compared to needs *in* the service.

Our fourth paper in this issue addresses the topic of tagging video or photographic materials online, specifically how to motivate and facilitate the consumers of these media in contributing tags that, among other things, assist in the indexing of digital materials. **Melenhorst and van Velsen** tested four tagging input mechanisms to see which process resulted in more individuals tagging consumed videos. They found that none of the three new mechanisms faired better overall than the standard input box, included as a comparison mechanism. They recommend further study of alternatives way of motivating users—either through education or technologies that are more engaging. The final original paper demonstrates a method for capturing user experiences. The repertory grid technique, a mixture of qualitative and quantitative methods, allows researchers to holistically gather cognitive and emotional aspects of the consumer experience of a technology. **Fallman and Waterworth** take the reader step-by-step through the use of the repertory grid technique, with recommendations on how designers and technology researchers can employ this method at various stages of the design process.

We also include in this issue a book review: **Ignacio Del Arco Herrer**a assesses Antti Hautamäki's *Sustainable Innovation: A New Age of Innovation and Finland's Innovation Policy*. In short, Del Arco Herrera acknowledges Hautamäki's contribution toward the

current transformation in perspectives on innovation policy. Whereas innovation policies traditionally have focused strictly on economic outcomes as measures of success, contemporary thinkers on innovation are advocating more holistic and sustainable outcomes, that is, in Hautamäki's view, policies that acknowledge and support equally the values of the environment and natural resources, the human resources (through, e.g., quality of life and education), and the economic outcomes.

REFERENCES

Bevan, N., Barnum, C., Cockton, G., Nielsen, J., Spool, J., & Wixon, D. (2003). The "magic number 5": Is it enough for web testing? In *CHI '03 extended abstracts on human factors in computing systems* (pp. 698–699). New York: ACM.

Broadbent, D. (1958). *Perception and communication*. London: Pergamon Press.

Bunge, M. (1967). *Scientific research I-II*. Berlin, Germany: Springer.

Caroll, J. (1997). Human-computer interaction. *Annual Review of Psychology*, *48*, 61–93.

Cerf, C. & Navatsky, V. (1984). *Experts speak*. New York: Villard.

Hempel, C. G. (1965). *Aspects of scientific explanation*. New York: Free Press.

Millard, A. (1990). *Edison and the business of innovation*. Baltimore, MD, USA: Johns Hopkins University Press.

Miller, G. E. (1956) The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

Nielsen, J. (2000, March 19). Why you only need to test with 5 users. Retrieved November 23, 2010, from http://www.useit.com/alertbox/20000319.html

Saariluoma, P., & Oulasvirta, A. (in press). User psychology: Re-assessing the boundaries of a discipline. *Psychology*.

Author's Note

All correspondence should be addressed to:
Pertti Saariluoma
University of Jyväskylä
Cognitive Science, Department of Computer Science and Information Systems
P.O. Box 35
FIN-40014 University of Jyväskylä, FINLAND
pertti.saariluoma@jyu.fi

Human Technology: An Interdisciplinary Journal on Humans in ICT Environments
ISSN 1795-6889
www.humantechnology.jyu.fi

HUMAN TECHNOLOGY

An Interdisciplinary Journal on Humans in ICT Environments ISSN: 1795-6889

www.humantechnology.jyu.fi Volume 6 (2), November 2010, 155–190

IS IT REALLY GENDER? AN EMPIRICAL INVESTIGATION INTO GENDER EFFECTS IN TECHNOLOGY ADOPTION THROUGH THE EXAMINATION OF INDIVIDUAL DIFFERENCES

Miguel I. Aguirre-Urreta
School of Accountancy and MIS
College of Commerce
DePaul University
USA

George M. Marakas
School of Business
University of Kansas
USA

Abstract: *A recent development in the technology acceptance literature is the inclusion of gender as a moderator of the relationships between intention and its antecedents, such that some are stronger for men than women, and vice versa. While the effects have been well established, the mechanisms by which they operate, that is, which specific gender differences are in operation and how they affect intention to adopt, have not been thoroughly explored. In this research, psychological constructs with established gender differences, such as core self-evaluations, computer self-efficacy and anxiety, psychological gender-role, and risk-taking propensity, are examined. In addition, this research introduces a novel context for the study of technology adoption in that more than a single alternative is offered to participants, thus requiring a choice among technologies. Results indicate that gender effects are more complex than previously thought, with potentially multiple influences from different facets operating simultaneously.*

Keywords: *technology acceptance, UTAUT, gender, choice.*

INTRODUCTION

Technology acceptance has been one of the most researched streams in the information systems literature. Since the introduction of the technology acceptance model (TAM; Davis, 1989), numerous studies have explored and expanded this theory (Agarwal & Karahanna, 2000; Gefen, Karahanna, & Straub, 2003; Koufaris, 2002). A recent study has proposed a theory of technology acceptance, the unified theory of acceptance and usage of technology (UTAUT), that explains a large proportion of variance in intention to use new technologies (Venkatesh, Morris, Davis, & Davis, 2003). It has been pointed out that, given the significantly high variance explained by UTAUT—unusual for the behavioral sciences—further work should aim at testing the boundary conditions of the model and expanding its real world applicability. That is the objective of the research described here.

URN: NBN:fi:jyu-201011173090

A topic of relatively recent emergence in technology acceptance research is the moderating influence of gender. Building on previous work (Venkatesh & Morris, 2000; Venkatesh, Morris, & Ackerman, 2000), UTAUT presents a moderating effect of gender in the relationships between performance expectancy and behavioral intention, such that it becomes stronger for men; and effort expectancy and behavioral intention, such that it is more significant for women (Venkatesh et al., 2003). Gender differences are useful in that they can propel research into an area by putting in evidence the existence of an underlying dynamic (Halpern, 1992).

One proposition drawn from the observed gender differences is that sensitivity to these differences could have significant impact on technology training and marketing, emphasizing the factors that are more salient to each group (Venkatesh et al., 2000). However, without more precise knowledge of the mechanisms by which these differences between men and women operate, the design and development of such programs is greatly hampered. A somewhat contradictory conclusion is the interpretation that such differences might be temporary and tend to disappear as a young cohort of employees are raised and educated in a technological environment (Venkatesh et al., 2003). Additionally, the usage of gender as a moderator can lead to equivocal results (Ndubisi, 2003). Overall, we need a better understanding of this issue before we can apply our knowledge to actual technology adoption settings. Simply knowing of a gender effect does not allow us to make use of this knowledge. The need to uncover the underlying mechanisms by which these gender differences arise has already been made explicit (Venkatesh et al., 2003). This study proposes and explores a set of variables to account for the observed gender effect that may further our understanding in this area. These constructs were selected as candidates for explaining observed gender effects because (a) these known differences have been exhibited by men and women, (b) these constructs are grounded in previous research, and (c) they could plausibly explain the relationships empirically observed. This study is thus concerned with answering the following research question: What are the underlying factors driving observed gender differences in the context of technology acceptance?

We tested these relationships in a novel context, one involving a choice between competing technologies. With but one known exception, TAM research has been conducted using different technologies in the same product category (Davis, 1989; Mathieson, 1991; Venkatesh & Davis, 1996), later evolving into non-comparable technologies (Venkatesh & Davis, 1996), and then just to single technology considerations (Venkatesh & Davis, 2000), where the decision was a binary choice between adopting the proposed technology or adopting no technology (a notable exception is Szajna, 1994). We believe that, while productive in the development of our understanding of the model and its elemental constructs, such scenarios are not representative of real-world technology adoption exercises. In such cases, it is rare that a decision to adopt a given technology is made without comparison to members of a refined choice set or without a mandate to actually choose one of the alternatives for adoption (absent any material weaknesses associated with the members of the final choice set). In other words, simply choosing to accept or reject a single technology in a vacuum is not representative of the conditions under which technologies are evaluated and adopted in an organizational setting.

Building upon this foundation, this research presents participants with an explicit consideration of and choice between alternatives, framed in an actual technology selection and adoption setting, using subjects professionally trained and employed in the domain in which the chosen technology will be used. We believe that this scenario presents a set of externally valid conditions that will further our understanding of UTAUT and its applicability to the domain of

practice, and introduces a refinement and measurable extension to the most accepted and researched model of technology acceptance in the information systems literature.

The next two sections review the current state of research in this area and the development of the hypotheses that define this study. Research design and variable operationalization are presented next. Finally, results and implications for future research are discussed.

THEORETICAL BACKGROUND

Technology Acceptance Research

The TAM, as originally proposed by Davis (1989), was a derivation of the theory of reasoned action (TRA; Fishbein & Ajzen, 1975) that was tailored to the domain of acceptance of information systems. TAM proposes that two beliefs—perceived usefulness and perceived ease of use—are the primary determinants of acceptance behavior, and that the two constructs mediate any other external variables. Following from TRA, TAM postulates that behavioral intention is the main determinant of usage, in turn driven jointly by attitude toward using and perceived usefulness (Davis, Bagozzi, & Warshaw, 1989). Departing from TRA, TAM did not include subjective norm as a determinant of behavioral intention; this construct, however, was added at a later time in an extension to the model (Venkatesh & Davis, 2000).

The appearance of other models attempting to explain technology acceptance, based on motivation, diffusion, and social cognitive theories, led to the formulation of UTAUT (Venkatesh et al., 2003; see Figure 1). The UTAUT postulates that three constructs, performance expectancy, effort expectancy, and social influence, will drive behavioral intention, which serves as an antecedent to use behavior, together with facilitating conditions. While proposed as an encompassing theory of eight competing models, a closer look at UTAUT reveals that TAM is still at the core of the model, with the four moderator variables having been identified in previous TAM research: experience and voluntariness (Venkatesh & Davis, 2000), age (Venkatesh & Morris, 2000) and gender (Venkatesh & Morris, 2000; Venkatesh et al., 2000). Additionally, the two TAM constructs, perceived usefulness and perceived ease of use, form the root components of performance and effort expectancy, respectively.

Past research in technology acceptance has used gender to mean the biological sex of the participants in the study (i.e., men or women). In other areas of research, gender takes on a psychological or socially-constructed meaning. In order to be consistent throughout our discussion, we use gender or sex to refer to the biological sex of individuals, thus keeping the usage from prior information systems studies, and qualify other uses of the term where required (e.g., psychological gender-role when discussing gender as an individual's own construction of femininity or masculinity).

In empirical tests, UTAUT accounted for 70% of the variance in intention to use; substantially higher than competing models and highly significant for the behavioral sciences in general. Given these results, small increases in the predictive power would be obtained only at the expense of increased complexity in the model. A more fruitful avenue of research would result from exploring the different situations and conditions in which UTAUT is applicable (Venkatesh et al., 2003).

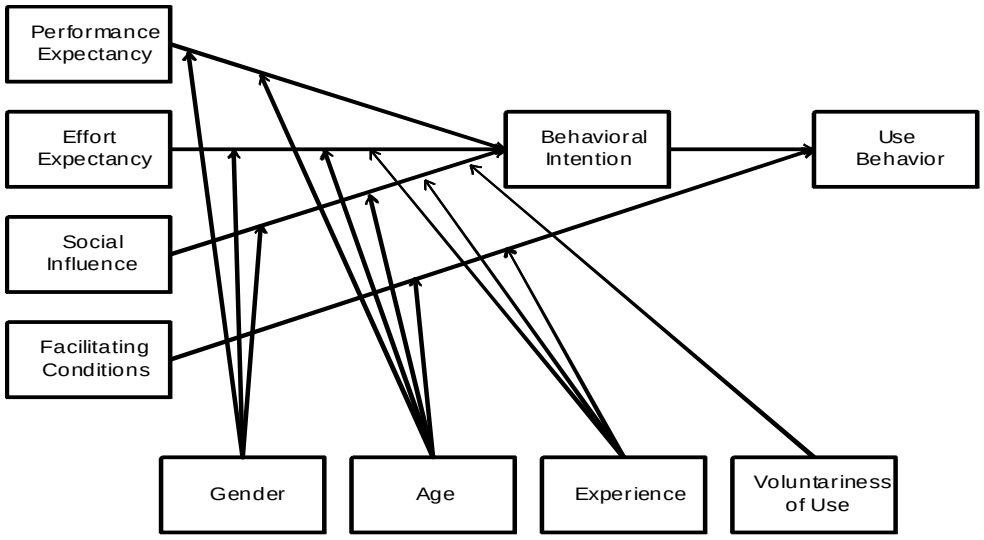


Figure 1. Unified theory of acceptance and usage of technology.
(Figure 3 from V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User Acceptance of Information Technology: Toward a Unified View," MIS Quarterly (27:3), 2003, p. 447.

Gender Differences

Research on gender differences has received the most extensive focus in the personality and social psychology literatures, as well as in the disciplines specializing in these subjects. Comparisons have been conducted in a variety of domains, including verbal and spatial cognitive skills, personality traits and dispositions, and social behaviors (Deaux, 1984, 1985). Theories as to the origin of these differences are grouped into two categories. The biological theories propose that sex-related differences arise from innate temperamental differences, evolved by natural selection (Costa, Terracciano, & McCrae, 2001). Research in studying heritability in twins and correlations with hormonal-chemical substances or physiological measures has suggested there is a strong biological basis underlying differences in personality traits (Feingold, 1994).

An alternative group of theories propose that gender differences arise from social and cultural factors affecting the way each sex develops through socialization. There are three variants of this proposition. The social role model developed by Eagly (Eagly & Wood, 1991) posits that gender differences in behavior arise from gender roles, which dictate appropriate behaviors for men and women. The expectancy model contends that social and cultural factors evolve in gender stereotypes that are reinforced because holders of these beliefs treat others in ways that result in one's conforming to the prejudices of the perceivers (Costa et al., 2001). Lastly, the artifact model proposes that sociocultural factors result in men and women holding different values about the importance of possessing various traits and that these differences bias self-reports of characteristics (Feingold, 1994).

Various studies have attempted to shed light on which of these alternative explanations for the emergence and persistence of gender-based differences work, although the argument is far from settled, if that is even possible. Costa et al. (2001), for example, noted that gender differences were generally modest in magnitude, but also consistent with gender stereotypes and these differences are replicable across cultures. Surprisingly, gender differences were found to be more pronounced

in countries with more progressive sex role ideologies (e.g., Western, individualistic countries). This finding goes counter to arguments from the social role model, whereas one would expect that these cultures would reflect smaller gender differences. It also goes against evolutionary explanations, since these would posit gender differences to be rather uniform within the human species, and not be influenced by particular cultures. Schmitt, Realo, Voracek, and Allik (2008) report a similar finding (see also McCrae & Terracciano, 2005), which counters the sex roles and evolutionary explanations. The authors, however, propose a novel rationale for these findings: More developed societies placed few constraints on human development and basic needs, thus providing more room for basic tendencies within individuals to flourish and diverge, whereas societies in which the lack of good health care, economic hardship, and limited access to education are prevalent, development of an individual's inherent personality is more constrained.

Given the above and varied characterizations of gender differences, it seems reasonable to assume that gender differences presenting themselves as a result of a dichotomous, biological representation of the construct fall short of explaining the underlying causal effects creating such differences. If we are to operationalize our understanding of technology acceptance, we need to understand the previously identified gender effects beyond simplistic biological assignment. We do this through the identification of a number of psychological constructs known to exhibit gender differences, and investigate whether those differences may be responsible for the observed gender effect in the technology acceptance literature.

RESEARCH MODEL AND HYPOTHESES DEVELOPMENT

Figure 2 provides a graphical representation of the research model employed to answer the research question posed above. We conceptualize this model in three distinct parts. The basic acceptance model is depicted along with a number of moderating factors as alternative conceptualizations to the previously observed gender effects derived from the gender literature. In testing multiple moderating effects, this research follows the strategy employed by McKeen, Guimaraes, and Wetherbe (1994) of individually testing the effects of each proposed variable. Finally, past research on antecedents to effort expectancy is replicated for validation purposes.

UTAUT Model

The UTAUT model proposed by Venkatesh et al. (2003) serves as the underlying framework for this research. We chose this theory for two reasons. First, it represents the most current theoretical and empirical synthesis of research in this stream of literature. The theory arose from the many conceptual and empirical similarities present in various models employed to investigate the phenomenon (e.g., TAM, the theories of reasoned action and planned behavior, innovation diffusion theory, etc.) and was empirically validated through extensive longitudinal testing. Second, while research conducted under some of the earlier conceptual frameworks had already identified gender effects (e.g., Venkatesh & Morris, 2000), the UTAUT integrates these effects, which are the central focus of attention in this research, into a comprehensive model of technology acceptance and usage. As a result, we employ the UTAUT as the underlying theoretical framework in this study, and, in more detail, examine one of the effects postulated there: the finding that the gender of the adopter has a moderating effect on the relationship between intention to adopt and its determinants.

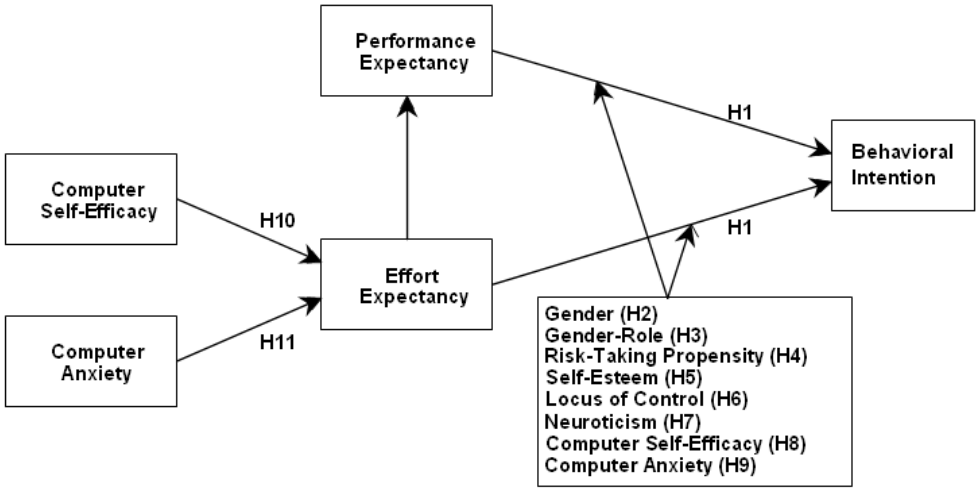


Figure 2. Research model for this study.

It should also be noted that these determinants of intention include three different constructs: Performance expectancy (defined as the degree to which the potential adopter believes using the focal technology will help her[1] increase job performance), effort expectancy (defined as the degree of ease associated with using the system), and social influence (the degree to which the individual perceives that important others believe she should use the technology). In the research model shown in Figure 2, however, only performance and effort expectancy are depicted as determinants of intention. While social influence is certainly an important determinant of intentions, we believe that the hypothetical setting in which the research was conducted limited the ability of participants to form realistic expectations about what important others would believe they should do. As a result, social influence is not included in the research model examined here. This issue is further discussed in the section dealing with the limitations to this research.

In addition to its focal research question, this study will provide a replication of the relevant portion of the UTAUT as a manipulation check. Thus, the following hypotheses will be tested:

H1(a): Performance expectancy will be a significant predictor of behavioral intention, such that increases in the former will result in increases in the latter.

H1(b): Effort expectancy will be a significant predictor of behavioral intention, such that increases in the former will result in increases in the latter.

H2(a): The relationship between performance expectancy and behavioral intention will be moderated by gender.

H2(b): The relationship between effort expectancy and behavioral intention will be moderated by gender.

Psychological Gender Role

Recent related research (e.g., Venkatesh, Morris, Sykes, & Ackerman, 2004) has examined gender as a psychological construct: a set of associations formed throughout human development that is not directly dependent on the natural or physiological gender. The authors

examined the role of psychological gender in technology acceptance and usage, employing the theory of planned behavior (Ajzen, 1991) as the underlying framework and found masculine individuals were significantly influenced only by attitude, while the opposite was the case for feminine subjects (only subjective norm and perceived behavioral control were significant predictors of behavioral intention). These results, while difficult to map in a one-to-one correspondence with those of Venkatesh et al. (2003), certainly parallel them and provide support for the role of psychological gender as a moderator of the relationships of interest. Thus, to further increase the validity of this research, the following is hypothesized:

> H3(a): The relationship between performance expectancy and behavioral intention will be moderated by psychological gender-role.
>
> H3(b): The relationship between effort expectancy and behavioral intention will be moderated by psychological gender-role.

Risk-Taking Propensity

Another demonstrated difference between men and women found in the literature is in their attitude toward risk. A meta-analytic review of studies regarding gender and risk taking found that the majority of reviewed research supported the idea of greater risk taking on the part of males. In particular, risk propensity is defined as an individual's tendency to take or avoid risks, and is conceptualized as a trait that can potentially change over time (Sitkin & Weingart, 1995). Potential explanations for this occurrence include overconfidence on the part of men and double standards of parental monitoring that place more restrictions on girls than on boys (Byrnes, Miller, & Schafer, 1999). Research concerning financial risk taking shows systematic risk-averse behavior by women, even when accounting for changes in total wealth (Jianakoplos & Bernasek, 1998). A study on decision making in a laboratory setting found women to be less risk seeking than men, with men choosing the risky option across other within-subjects differences (Lauriola & Levin, 2001). This study proposes that the decision to adopt an information system presents characteristics similar to those existing in the reviewed literature regarding uncertainty of outcome and consequences. The following hypotheses are thus put forward:

> H4(a): The relationship between performance expectancy and behavioral intention will be moderated by risk-taking propensity.
>
> H4(b): The relationship between effort expectancy and behavioral intention will be moderated by risk-taking propensity.

Personality Traits

Gender differences in personality traits have been documented in many empirical studies (Costa et al., 2001). In the late 1970s, the popularization of meta-analytic techniques allowed researchers to aggregate research findings. Feingold's (1994) review of the seminal research of Maccoby and Jacklin (1974), found that men, compared to women, were higher in self-esteem, more assertive, more internally controlled, and less anxious. Since then, multiple other studies—many with very large samples and across cultures—have confirmed the presence of differences in personality traits between men and women. In a study with self-reported data from 26 national cultures (N = 23,301), Costa et al. (2001) found that women report themselves higher than men in neuroticism,

agreeableness, warmth, and openness to feelings, whereas men were higher in assertiveness and openness to ideas. In another large data collection effort, Schmitt et al. (2008) obtained data from 55 nations (N = 17,637) and found women to report higher levels of neuroticism, extraversion, agreeableness, and conscientiousness than did men. More recently, analysis of a very large, cross-cultural dataset (N > 200,000) confirmed those results (Lippa, 2010).

While the number of personality traits researched in the past is significant, two distinct models have emerged, each presenting a core set of traits that can be used to subsume differences in personality. The first one is the Big Five—neuroticism, extraversion, openness, agreeableness, and conscientiousness (Langston & Sykes, 1997). An alternative categorization, the Core Self-Evaluations, proposes self-esteem, generalized self-efficacy, locus of control, and emotional stability as determinants of an individual's perspective of oneself and her relationship with her environment (Judge, Locke, Durham, & Kluger, 1998). Judge and colleagues defined the individual evaluations as follows: Self-esteem is the basic appraisal people make of themselves, locus of control concerns the degree to which individuals believe that they control events in their lives (as compared to the environment or fate), and neuroticism as constituting the negative pole of self-esteem. Generalized self-efficacy, instantiated here within the computer domain, can be defined as "an individual's perception of efficacy in performing specific computer-related tasks within the domain of general computing" (Johnson, Marakas, & Palmer, 2006; Marakas, Yi, & Johnson, 1998).

All components of the core self-evaluation set have been shown to present significant differences when evaluated in men and women (Feingold, 1994; Johnson et al., 2006; Marakas et al., 1998). This perspective is the one adopted for the purpose of this research. Although considered a member of the core self-evaluations constructs, hypothesis development for computer self-efficacy will be presented in the next section, when discussing its relationship to user acceptance and computer anxiety. Consistent with prior research, it is here proposed that core self-evaluations will be related to the main relationships under study, and thus the following hypotheses are presented:

H5(a): The relationship between performance expectancy and behavioral intention will be moderated by self-esteem.

H5(b): The relationship between effort expectancy and behavioral intention will be moderated by self-esteem.

H6(a): The relationship between performance expectancy and behavioral intention will be moderated by locus of control.

H6(b): The relationship between effort expectancy and behavioral intention will be moderated by locus of control.

H7(a): The relationship between performance expectancy and behavioral intention will be moderated by neuroticism.

H7(b): The relationship between effort expectancy and behavioral intention will be moderated by neuroticism.

Computer Self-Efficacy and Computer Anxiety

Past research has argued for, and strongly supported, the lack of a direct effect of both computer self-efficacy and computer anxiety on intention to adopt a new technology (Venkatesh et al., 2003). In this research, these two constructs are argued to influence behavioral intention through moderating the effects of performance and effort expectancy on the former. There is strong support in the literature for the notion that, *ceteris paribus*, women generally exhibit a lower initial level of general computer self-efficacy (Busch, 1995, 1996; Hartzel, 2003; Marakas et al., 1998), and higher levels of computer anxiety (Busch, 1995; Harrison & Rainer, 1992; Heinsenn, Glass, & Knight, 1987). Following from the above exposition, the following hypotheses are advanced, expressed in terms consistent with the formulation of UTAUT:

H8(a): The relationship between performance expectancy and behavioral intention will be moderated by computer self-efficacy.

H8(b): The relationship between effort expectancy and behavioral intention will be moderated by computer self-efficacy.

H9(a): The relationship between performance expectancy and behavioral intention will be moderated by computer anxiety.

H9(b): The relationship between effort expectancy and behavioral intention will be moderated by computer anxiety.

Another explanation for the observed gender differences advanced by previous research refers to the characterization of perceived ease of use (effort expectancy in UTAUT) as a hurdle to user acceptance (Venkatesh & Morris, 2000). In this conception, users anchor their perceptions of ease of use to their computer self-efficacy and adjust those perceptions according to the objective usability of the system after hands-on experience. Thus, systems whose perceived usability falls beneath the threshold of the user's computer self-efficacy are more likely to be rejected (Venkatesh & Davis, 1996). Research into antecedents of perceived ease of use has found significant results for both computer self-efficacy (Agarwal & Karahanna, 2000; Venkatesh, 2000; Venkatesh & Davis, 1996) and computer anxiety (Venkatesh, 2000). The proposition previously advanced is that lower levels of computer self-efficacy and higher levels of computer anxiety among women lead to lowering their perceptions of ease of use, and thus low perceptions of this construct increase its salience in forming the intention to adopt (Venkatesh & Morris, 2000). Consistent with past research (e.g., Venkatesh, 2000; Venkatesh & Davis, 1996), the following hypotheses are advanced, in an attempt to replicate past findings:

H10: Computer self-efficacy will have a positive effect on effort expectancy.

H11: Computer anxiety will have a negative effect on effort expectancy.

VARIABLE OPERATIONALIZATION AND MEASUREMENT

This section discusses in more detail the different instruments used to measure the different constructs of interest. All scales were drawn from existing research and have been employed and validated in various contexts.

Psychological Gender-role

A shortened version of the Bem Sex-Role Inventory (BSRI; Bem, 1974, 1981; Campbell, 1997; Powell & Butterfield, 2003) was used to measure the psychological gender-role of individual participants. While the original version of the BSRI instrument comprised 60 items, a shorter set was developed by Bem to facilitate its use in research settings without sacrificing its underlying characteristics. The scores of two sets of 10 items are totaled and subtracted one from the other to arrive at a difference score that measures gender traits. An important advantage of this form of measurement is that it generates a continuous variable, theoretically ranging between minus 60 and plus 60, although the actual observed range is generally narrower. Thus, it is not necessary to categorize individuals as masculine or feminine in order to analyze the effects of psychological gender-role on the outcomes of interest.

Core Self-Evaluations

These constructs were measured using the Core Self-Evaluation instrument developed by Johnson et al. (2006). In some studies (e.g., Judge, Thoresent, Pucik, & Welbourne, 1999), the various core self-evaluation traits are combined into one single factor, and then the predictive validity of the latter is examined. The current research, however, distinguishes between the traits and analyzes their potential effects independently.

Computer Anxiety

This construct has been measured in a variety of ways ever since computers were introduced in the workplace. Many implementations of the concept can be traced back to the fear facet of the original rating scale by Heinsenn et al. (1987), the Computer Anxiety Rating Scale (CARS), which used a 5-point *strongly agree–strongly disagree* format. An alternative scale is used by Venkatesh (2000), composed of nine items in a 7-point Likert scale of similar format. The items employed in this study are a subset of those originally developed by Heinsenn et al. (1987), after removing those items that are no longer representative of the current technological context. Higher scores are an indication of increased anxiety toward computers.

UTAUT constructs

The core constructs of UTAUT were measured following the guidelines set in the original study.

Risk-Taking Propensity

Two major approaches regarding the measurement of attitudes toward risk can be found in the relevant literature: Those derived from the employment of the expected utility framework, and those resulting from using psychometric scales that ask participants to rate their agreement with a set of relevant statements, where the former appear to be better predictors of actual behavior (Penning & Smidts, 2000). This research used two measures to operationalize expected utility and capture the construct of risk-taking propensity. The first measure was constructed within the expected utility (e.g., "lottery") approach following the guidelines set forth by Lauriola and Levin

(2000, 2001). For the second measure (e.g., "Lottery measure B"), the decision was between two risky propositions, where the first involved less outcome variability (e.g., 60/40) and second more outcome variability (e.g., 25/75), while still holding expected value between options equal. In both cases, participants choosing the first alternative were deemed to be more risk-averse, while participants choosing the second alternative, more risk taking.

RESEARCH DESIGN AND DATA COLLECTION

Participants

Sixty-four business professionals participated in this study, drawn mostly from large public accounting firms in the Midwest United States. All subjects were employed at firms that supported a curricular advisory body, and were recruited by contacting representatives of this body requesting they distribute a call for participation to other employees of their firm. Of the original sample, 56 provided evaluations of the two technologies as well as answered questions regarding their intention to hypothetically adopt them in a business organization. Of these, 40 participants explicitly chose one of the two technologies under consideration, and these form the final sample for analysis. The remaining subjects could not decide between the two alternatives presented to them and were thus removed from further analysis. Table 1 displays the demographic and employment characteristics of the final subject pool.

Design

Data for this research were collected via a secure Website that participants could access at their convenience. After agreeing to participate in the study and providing basic demographic information, participants answered a set of questions that captured the constructs of interest by selecting the desired option from drop-down boxes located next to the statement prompting

Table 1. Sample Characteristics (N = 40).

Gender	**Count**	**%**		**Count**	**%**
Male	23	57.5	**Income (annual)**		
Female	17	42.5	$40,000 - $60,000	8	20.0
Age			$60,000 - $80,000	7	17.5
18 – 25	11	27.5	$80,000 - $100,000	4	10.0
26 – 35	15	37.5	$100,000 - $150,000	13	32.5
36 – 45	8	20.0	$150,000 or more	8	20.0
46 – 55	5	12.5	**Position**		
56 – 65	1	2.5	Exec / Senior Mgmt.	7	17.5
Education level			Middle mgmt.	10	25.0
Some college	1	2.5	Supervisory	10	25.0
Graduated college	9	22.5	Admin. / Clerical	3	7.5
Post-graduate studies	30	75.0	Technical	10	25.0

a response. Where appropriate, items were randomized across different measures. All scales were validated and refined during a series of pilot studies using techniques appropriate for the nature of the scales and in keeping with the tenets set forth by Straub (1989) and Boudreau, Gefen and Straub (2001).

Figure 3 shows the entire sequence of data collection and assignment to the appropriate research condition as was experienced by the participating subjects. Data about the proposed moderating variables were collected before participants had access to the experimental materials, whereas data about their technology evaluations and intentions (e.g., data for performance expectancy, effort expectancy, and intention to adopt for each technology) were collected afterwards. Finally, participants were thanked for their time and dismissed. While participants were informed of the general nature of the study, focused on the decision-making process behind technology adoption decisions, they were not made aware of the focus on gender effects in this area. This was done in an effort to prevent participants from considering how their responses to the questionnaire may be construed in light of their gender, and thus allow us to obtain data that was less subject to self-presentation bias. A complete list of all items presented in the questionnaire, organized by measure and including, where necessary, response instructions, are included in Appendix A. Sources for these measures were discussed in the previous section.

Participants were randomly assigned based on their domain of training and employment as either accountants or marketing professionals. Subjects were asked to review and evaluate two technologies for potential adoption in a hypothetical organization. In half of the cells, the two technologies were accounts receivable packages, in the remaining, with appropriate modification of the framing, coupon management software. All participants were presented with a hypothetical framing: Their organization was undergoing the evaluation and selection process for a new technology, and they had been selected as members of the committee tasked with such endeavor. After prior screening by their Information Technology department, two candidate software packages had been identified as potential candidates.

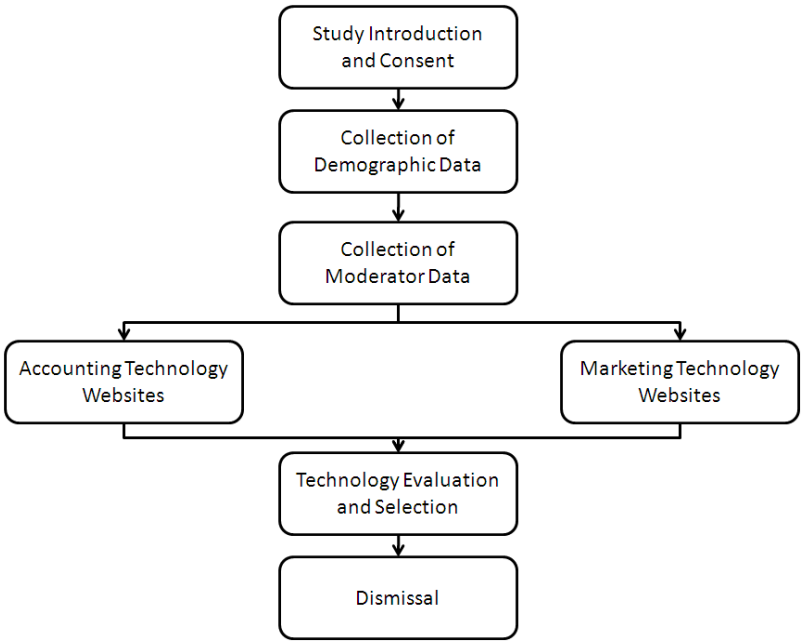


Figure 3. This study's complete sequence of events and data collection.

Participants could access modified vendor Websites for each technology. While the Websites included in this research retained the look and feel of the actual vendors of these technologies (including color, layout, and logos), they were modified by the authors both to remove elements extraneous to this research, such as contact information, links to other products offered by the same vendor, and so on, and to shorten the number of features to reduce the load on the participants. Sample screenshots of the materials are included in Appendix B. Results of the pilot studies revealed no perceived loss of functionality relevant to the selection process as a result of the reduction of listed functions originally supplied by the vendors. The data collection system was designed to ensure that no subject could participate more than once and no subject could suspend their participation and return at a later time.

ANALYSIS METHODS

Data modeling and analysis for this research was conducted using Partial Least Squares (specifically, SmartPLS 2.0 M3; Ringle, Wende, & Will, 2005). The PLS methodology was selected for its ability to handle small samples, such as the one employed in this study, and the existence of prescriptive literature on the modeling of interaction effects with latent variables (e.g., Chin, Marcolin, & Newsted, 2003). Given the comparative nature of this study, perceptions for the different technologies were grouped into those that had been chosen by the participant, and those that were not, with an eye toward assessing the possibly differential effects of the moderating variables for these two groups of technologies. However, when limiting the items in each latent variable to those that loaded highly and significantly in their intended construct (e.g., Gefen & Straub, 2005), it was realized that the intended moderator variables would not necessarily be represented by the same set of indicators, raising questions about the comparability of the effects across chosen and not-chosen technologies. Thus, an alternative approach was devised in order to test the hypothesized relationships. An example using computer anxiety is depicted in Figure 4.

By modeling latent variables in this fashion, and retaining only those items that significantly loaded on the intended moderating variable, two objectives were fulfilled. First, comparability of the moderator effects between the two groups was made possible, since the same set of indicators represented the latent variable in both cases. To further constrain this to be the case, all moderating effects presented in this section were tested jointly with both technologies present, as shown in Figure 4. Second, this allowed for the direct effect of the proposed moderator variables to be included in the model before any interaction effects were assessed (Jaccard, Turrisi, & Wan, 1990). In particular, interaction effects were modeled and analyzed as follows.

First, a base model containing perceptions of effort and performance for each technology, as well as the direct effect of the focal moderating variable, was estimated using PLS; Figure 4 represents an example of this first step when examining the moderating effects of computer anxiety. Results from this analysis are referred to as the *base model* in the next section. Next, interaction effects were added to this base model. The product-indicator approach recommended by Chin et al. (2003) was employed to model the interaction effects, with the indicators being standardized prior to the multiplication. Following the recommendations of Chin and colleagues, as many significantly-loading indicators were retained as allowed by the sample size, given the importance of this factor in the appropriate detection of interaction effects. The proportion of

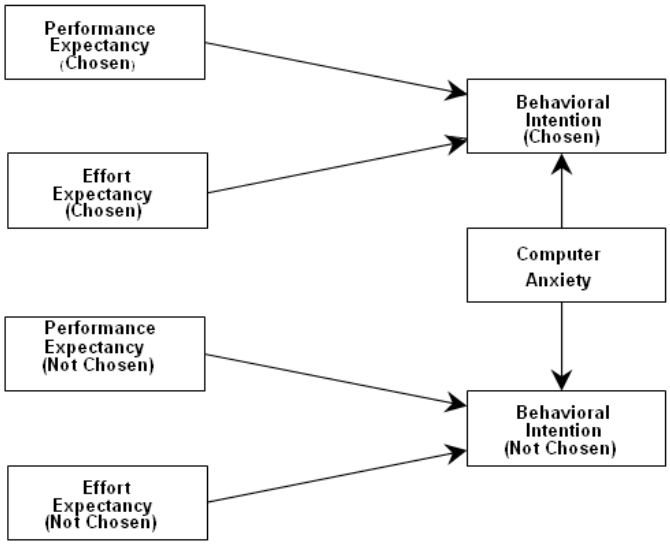


Figure 4. Two-group modeling approach.

variance explained in the dependent variable by the *full model*, containing the interaction terms, was compared to that of the base model, which contained only the direct effects. The statistical significance of this increase in variance explained was then assessed. The approach is analogous to the hierarchical testing of moderating effects in multiple linear regressions, but employing PLS as the underlying technique. Pavlou and El Sawy (2006) employed a similar approach.

Given the statistical limitations imposed by the number of participants in this research (e.g., the heuristic of 10 cases per effect on any endogenous variable), interactions were tested for performance and effort expectancy separately, as detailed below. Despite not being the main focus of this study, additional validation of the research framework employed was obtained by modeling the intentions to adopt for each of the two technologies evaluated by the participants as antecedents to a dummy-coded variable indicating the actual choice made. The results strongly support the comparative nature of this research, with both paths strongly significant (at the $p < 0.0001$ level) and the variance explained in the choice variable just short of 68%.

Convergent and discriminant validity were assessed following the extant procedures outlined by Gefen and Straub (2005). Only those indicators that loaded significantly in their latent variable were retained in the final model. An examination of the loading patterns revealed no cross-loadings of any important magnitude, and in all cases the square root of the average variance extracted was larger than any correlations among pairs of latent constructs. Composite reliabilities were also above recommended thresholds.

HYPOTHESIS TESTING AND RESULTS

Tables 2a[2] and 2b contain the results of the testing of H1a and H1b. As can be seen from the results, both performance and effort expectancy are significantly associated with behavioral intention for both the chosen and the not chosen technologies ($p < 0.05$). The standardized betas shown in Table 2b also indicate significance with regard to the relationship between performance and behavioral intention and effort expectancy and behavioral intention ($p < 0.05$). These results

Table 2a. Measurement Model – Base Models.

	CR	BI (CH)	PE (CH)	EE (CH)	BI (NCH)	PE (NCH)	EE (NCH)
BI (CH)	0.8681	**0.833**					
PE (CH)	0.8678	0.452**	**0.790**				
EE (CH)	0.9593	0.456**	0.429**	**0.925**			
BI (NCH)	0.9519				**0.932**		
PE (NCH)	0.9629				0.467**	**0.931**	
EE (NCH)	0.9738				0.459**	0.480**	**0.950**

Note: Models were estimated independently of each other. Elements in the diagonal are the square root of the average variance extracted (AVE); off-diagonal elements are correlations between the latent constructs. CH = Chosen, NCH = Not Chosen, CR = Composite Reliability, BI = Behavioral Intention, PE = Performance Expectancy, EE = Effort Expectancy.
*Correlation significant at the 0.05 level (two-tailed), **Correlation significant at the 0.01 level (two-tailed).

Table 2b. Base Models.

Block	Term	Behavioral Intention (Chosen)		Behavioral Intention (Not Chosen)	
		B	R^2	B	R^2
Base Model	PE	0.314*	0.288	0.321*	0.290
	EE	0.321*		0.305*	

Note: Models for the chosen and not-chosen technologies were estimated independently of each other. PE = Performance Expectancy, EE = Effort Expectancy.
*$p < 0.05$.

provide clear support for H1a and H1b and are in keeping with previous results obtained for UTAUT suggesting validity of the measurement models (Venkatesh et al., 2003).

H2a and H2b focus on the moderating effects of gender as reported by prior studies. As can be seen from Table 3, the results parallel those of prior studies with the observed gender effect negatively related to performance expectancy (PE) and positively related to effort expectancy (EE). Based on the coding of gender employed in this research, these results suggest that the effects of PE on behavioral intention (BI) are stronger for men than are for women, while the converse is true for the effects of EE on BI (which are stronger for women than for men). This is evidenced by the negative path coefficient from PE to BI, indicating that women place less importance than men on the level of expected performance derived from use of the focal technology, and by the positive path emanating from EE to BI, suggesting in this case that women place more of an emphasis on levels of ease of use associated with the technology under consideration than men do. These results are significant only for the chosen technology, although the coefficients are of the expected sign for the not-chosen technology. This provides support for H2a and H2b and replicates prior work.

H3a and H3b focus on the proposed relationships between psychological gender-role and BI. We find little evidence of this relationship; significance for these coefficients was found only

in the moderating relationships for the not-chosen technology. The signs of the coefficients parallel those obtained in the testing of H2, however. As such, we can find no support for H3a but we find some support for H3b. Other research that has examined these relationships, albeit using a different theoretical basis (Venkatesh et al., 2004) indicates that masculine individuals form their intentions based on utilitarian attitudes toward technology, whereas more feminine individuals emphasize their ability to use the technology more. These results are robust to the gender of the individual, thus showing that psychological gender-role provides additional variance beyond the dichotomous classification of participants into male and female, thus increasing the explanatory power of the model. When viewed in conjunction with the results obtained for H2, and in keeping with earlier findings related to this construct, we find support for gender (either biological or role; see Table 4) as a moderator within the model.

Table 3. Moderating Effects of Biological Gender.

Block	Term	Behavioral Intention (Chosen)			Behavioral Intention (Not Chosen)		
		B	R^2	ΔR^2	B	R^2	ΔR^2
PE Interaction only	PE	0.275*	0.321	0.031	0.331*	0.378	0.010
	EE	0.322*			0.330*		
	GENDER	-0.032			-0.276*		
	PE x GENDER	-0.159$^+$			-0.132		
EE Interaction only	PE	0.278$^+$	0.346	0.086	0.420*	0.368	0.000
	EE	0.329*			0.310*		
	GENDER	-0.023			-0.318**		
	EE x GENDER	0.258*			0.003		

Note: Models for the chosen and not-chosen technologies were estimated independently of each other. Changes in R^2 for the interaction terms are calculated using the base model with the direct effect of the moderator variable as the reference. PE = Performance Expectancy, EE = Effort Expectancy.
*$p < 0.05$, + $p < 0.10$, **$p < 0.01$.

Table 4. Moderating Effects of Psychological Gender-Role (BSRI).

Block	Term	Behavioral Intention (Chosen)			Behavioral Intention (Not Chosen)		
		B	R^2	ΔR^2	B	R^2	ΔR^2
PE Interaction only	PE	0.263$^+$	0.304	0.015	0.213$^+$	0.340	0.041
	EE	0.336*			0.364*		
	BSRI	0.091			-0.076		
	PE x BSRI	-0.126			-0.236$^+$		
EE Interaction only	PE	0.320*	0.297	0.008	0.420**	0.368	0.069
	EE	0.294$^+$			0.218$^+$		
	BSRI	0.060			-0.253*		
	EE x BSRI	-0.080			0.319**		

Note: Models for the chosen and not-chosen technologies were estimated independently of each other. Changes in R^2 for the interaction terms are calculated using the base model with the direct effect of the moderator variable as the reference. PE = Performance Expectancy, EE = Effort Expectancy, BSRI = Bem Sex Role Index.
*$p < 0.05$, + p <0.10, **$p < 0.01$.

To provide a thorough investigation into the various forms of risk-taking behavior, recall that two different approaches were employed: (a) Lottery A, looking at sure gain versus a risky proposition, and (b) Lottery B, a choice between two risky propositions. The results for each of these measures are shown in Tables 5a and 5b.

Lottery A (sure gain vs. risky proposition) displays somewhat equivocal results with regard to its potential moderating effect. As shown in Table 5a, the construct displays a negative moderating effect for both PE and EE for both technologies but is significant only for EE in the chosen technology and PE for the not-chosen technology ($p < 0.05$). For Lottery B, we find a slightly different set of relationships. The results in Table 5b indicate the construct provides a negative moderation for PE and EE in the chosen technology and EE in the not-chosen technology, but a positive moderation for PE in the not-chosen technology. Further, significance

Table 5a. Moderating Effects of Risk Propensity (Measure Lottery A).

Block	Term	Behavioral Intention (Chosen)			Behavioral Intention (Not Chosen)		
		B	R^2	ΔR^2	B	R^2	ΔR^2
PE Interaction only	PE	0.303	0.410	0.002	0.224*	0.401	0.084
	EE	0.326*			0.228$^+$		
	RP-A	-0.351**			0.251*		
	PE x RP-A	-0.059			-0.324*		
EE Interaction only	PE	0.362**	0.509	0.101	0.295	0.323	0.006
	EE	0.121			0.157		
	RP-A	-0.290*			0.210$^+$		
	EE x RP-A	-0.396*			-0.147		

Note: Models for the chosen and not-chosen technologies were estimated independently of each other. Changes in R^2 for the interaction terms are calculated using the base model with the direct effect of the moderator variable as the reference. PE = Performance Expectancy, EE = Effort Expectancy, RP-A = Risk Propensity, Lottery Measure A.
*$p < 0.05$, + p <0.10, **$p < 0.01$.

Table 5b. Moderating Effects of Risk Propensity (Measure Lottery B).

Block	Term	Behavioral Intention (Chosen)			Behavioral Intention (Not Chosen)		
		B	R^2	ΔR^2	B	R^2	ΔR^2
PE Interaction only	PE	0.341*	0.335	0.037	0.306**	0.395	0.060
	EE	0.283$^+$			0.299*		
	RP-B	-0.079			0.209*		
	PE x RP-B	-0.158			0.254*		
EE Interaction only	PE	0.409**	0.343	0.045	0.146	0.418	0.083
	EE	0.195			0.280$^+$		
	RP-B	-0.139			0.237*		
	EE x RP-B	-0.280*			-0.337*		

Note: Models for the chosen and not-chosen technologies were estimated independently of each other. Changes in R^2 for the interaction terms are calculated using the base model with the direct effect of the moderator variable as the reference. PE = Performance Expectancy, EE = Effort Expectancy, RP-B = Risk Propensity, Lottery Measure B.
*$p < 0.05$, + p <0.10, **$p < 0.01$.

is found only for PE in the chosen technology and EE in both technologies. While similarities exist between the two tests of risk propensity, the differences are notable. When considered together, the results provide support for H4a and H4b, suggesting that risk-taking propensity is a moderator for both PE and EE and their relationship to BI.

The results for the test of self-esteem (one of the core self-evaluation constructs) are found in Table 6. The results suggest that self-esteem (SE) plays an important moderating role with regard to the PE– and EE–BI relationships. Interestingly, SE serves as a positive moderator for the chosen technology and in a negative capacity for the not-chosen technology. It would seem that SE provided the subjects with a form of enhancement of the differences between the two technologies (this will be discussed further immediately below). Also, given the prior relationships between SE and gender reported in the literature, SE appears to be a strong candidate to better explain the previously reported gender moderation in UTAUT. Given these results, we find support for H5a, and H5b.

Continuing with our tests of the individual components within the core self-evaluation construct (see Table 7), we find locus of control (LC) to be a significant moderator within the model. While significant for both PE and EE for both technologies, LC appears to positively moderate PE while negatively moderating EE for the chosen technology and negatively moderating both variables for the not-chosen technology. While the reasoning behind these findings requires further thought and discussion (and, given the exploratory nature of this research, possibly further study), the results obtained provide clear support for H6a, and H6b.

In much of the psychology literature, neuroticism is characterized as an opposing core self-evaluation to self-esteem. When viewed in combination with the results obtained for self-esteem, we see continued evidence of this characterization. A review of Table 8 indicates neuroticism to be a potential moderator within UTAUT but more clearly for PE than for EE. As such, we find clear support for H7a, with limited support for H7b.

The final component in the core self-evaluation construct is computer self-efficacy (CSE). Recall this variable was measured at the general domain level (GCSE) as conceptualized by Marakas et al. (1998) and operationalized by Johnson & Marakas (2000). As shown in Table 9, GCSE is a significant moderator for both PE and EE with regard to the

Table 6. Moderating Effects of Self-Esteem.

Block	Term	Behavioral Intention (Chosen)			Behavioral Intention (Not Chosen)		
		B	R^2	ΔR^2	B	R^2	ΔR^2
PE Interaction only	PE	0.235^{+}	0.347	0.031	0.238*	0.368	0.074
	EE	0.305*			0.226^{+}		
	SE	-0.155^{+}			-0.165^{+}		
	PE x SE	0.211^{+}			-0.344*		
EE Interaction only	PE	0.399**	0.387	0.071	0.236^{+}	0.323	0.029
	EE	0.095			0.326*		
	SE	-0.309*			-0.087		
	EE x SE	0.346*			-0.200^{+}		

Note: Models for the chosen and not-chosen technologies were estimated independently of each other. Changes in R^2 for the interaction terms are calculated using the base model with the direct effect of the moderator variable as the reference. PE = Performance Expectancy, EE = Effort Expectancy, SE = Self-Esteem.
*$p < 0.05$, + p <0.10, ** $p < 0.01$.

Table 7. Moderating Effects of Locus of Control.

Block	Term	Behavioral Intention (Chosen)			Behavioral Intention (Not Chosen)		
		B	R^2	ΔR^2	B	R^2	ΔR^2
PE Interaction only	PE	0.279	0.368	0.043	0.221*	0.318	0.026
	EE	0.189$^+$			0.335$^+$		
	LC	-0.232$^+$			0.007		
	PE x LC	0.239*			-0.204$^+$		
EE Interaction only	PE	0.281*	0.381	0.056	0.220$^+$	0.336	0.044
	EE	0.220$^+$			0.301*		
	LC	-0.256*			0.020		
	EE x LC	-0.266*			-0.253*		

Note: Models for the chosen and not-chosen technologies were estimated independently of each other. Changes in R^2 for the interaction terms are calculated using the base model with the direct effect of the moderator variable as the reference. PE = Performance Expectancy, EE = Effort Expectancy, LC = Locus of Control.
*$p < 0.05$, + p <0.10, **$p < 0.01$.

Table 8. Moderating Effects of Neuroticism.

Block	Term	Behavioral Intention (Chosen)			Behavioral Intention (Not Chosen)		
		B	R^2	ΔR^2	B	R^2	ΔR^2
PE Interaction only	PE	0.228$^+$	0.329	0.038	0.222*	0.384	0.056
	EE	0.312*			0.298*		
	NE	-0.047			0.210$^+$		
	PE x NE	0.223$^+$			0.258$^+$		
EE Interaction only	PE	0.388**	0.436	0.145	0.341*	0.333	0.005
	EE	0.144			0.236$^+$		
	NE	-0.122			0.237*		
	EE x NE	0.429**			0.080		

Note: Models for the chosen and not-chosen technologies were estimated independently of each other. Changes in R^2 for the interaction terms are calculated using the base model with the direct effect of the moderator variable as the reference. PE = Performance Expectancy, EE = Effort Expectancy, NE = Neuroticism.
*$p < 0.05$, + p <0.10, **$p < 0.01$.

chosen technology, suggesting that higher levels of GCSE contribute to the formation of PE and EE perceptions and to the choice process. In addition, the results suggest that higher levels of GCSE in a choice setting will have a greater effect on the formation of EE perceptions than on PE perceptions. Further, the results suggest that GCSE is not a salient moderator with regard to the not-chosen technology. Here again, we see evidence of a type of enhancement in differentiating between the two technologies brought forth by the subject's GCSE perceptions. Given these results, we find clear support for H8a, and H8b.

Computer anxiety is generally characterized as a deterrent to forming sound perceptions regarding a technology. A review of Table 10 suggests this characterization to be salient in its moderating effects in UTAUT. Consistent with the results obtained with self-esteem, these findings clearly position higher levels of computer anxiety as a negative moderator to forming PE and EE perceptions and their relationships to BI. Given these results, we find support for H9a, and H9b.

Table 9. Moderating Effects of Generalized Computer Self-Efficacy.

Block	Term	Behavioral Intention (Chosen)			Behavioral Intention (Not Chosen)		
		B	R^2	ΔR^2	B	R^2	ΔR^2
PE Interaction only	PE	0.343*	0.451	0.113	0.280*	0.353	0.018
	EE	0.369*			0.300*		
	GCSE	-0.367**			-0.240*		
	PE x GCSE	0.363**			0.145		
EE Interaction only	PE	0.236*	0.589	0.251	0.336*	0.348	0.013
	EE	0.473**			0.280*		
	GCSE	-0.162$^+$			-0.216*		
	EE x GCSE	0.529***			-0.124		

Note: Models for the chosen and not chosen technologies were estimated independently of each other. Changes in R^2 for the interaction terms are calculated using the base model with the direct effect of the moderator variable as the reference. PE = Performance Expectancy, EE = Effort Expectancy, GCSE = Generalized Computer Self-Efficacy.
** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

Table 10. Moderating Effects of Computer Anxiety.

Block	Term	Behavioral Intention (Chosen)			Behavioral Intention (Not Chosen)		
		B	R^2	ΔR^2	B	R^2	ΔR^2
PE Interaction only	PE	0.356*	0.371	0.083	0.273*	0.438	0.119
	EE	0.222			0.268*		
	CANX	-0.060			0.300*		
	PE x CANX	-0.292*			0.403**		
EE Interaction only	PE	0.371**	0.433	0.145	0.314*	0.328	0.009
	EE	0.270*			0.363*		
	CANX	0.008			0.211$^+$		
	EE x CANX	-0.405**			0.102		

Note: Models for the chosen and not-chosen technologies were estimated independently of each other. Changes in R^2 for the interaction terms are calculated using the base model with the direct effect of the moderator variable as the reference. PE = Performance Expectancy, EE = Effort Expectancy, CANX = Computer Anxiety.
*$p < 0.05$, + $p < 0.10$, **$p < 0.01$.

Hypotheses 10 and 11 propose both GCSE and computer anxiety (CANX) will have an antecedent relationship to the formation of effort expectancy perceptions. Table 11 contains the results obtained with regard to the testing of these hypotheses. Tested separately, both GCSE and CANX display significant direct effects with EE for the chosen technology, suggesting that higher levels of GCSE and lower levels of CANX will directly affect perceptions of ease of use. When tested together, however, the effect of CANX on the formation of effort perceptions appears to supplant the effects of GCSE. Given that CANX often has been positioned as an antecedent to GCSE (Marakas et al., 1998), these results suggest that, in the presence of high levels of CANX, a person's GCSE perception is less important than her feelings of concern with regard to forming a perception of effort expectancy. Given these results, we find support for H10 and H11.

Table 11. Antecedents to Effort Expectancy.

Block	Term	Effort Expectancy (Chosen)		Effort Expectancy (Not Chosen)	
		B	R^2	B	R^2
GCSE only	GCSE	0.290**	0.084	0.129	0.017
CANX only	CANX	-0.361**	0.130	-0.350**	0.123
Both GCSE and CANX (no relationship between GCSE and CANX)	GCSE CANX	0.131 -0.291^{+}	0.142	-0.091 -0.398*	0.128
Both GCSE and CANX (GCSE and CANX related)	GCSE CANX	0.101 -0.308*	0.139	-0.083 -0.389*	0.123

Note: Models for the chosen and not-chosen technologies were estimated independently of each other. The same indicators used in the estimation of the interaction effects were used in these models in order to maintain consistency. GCSE = Generalized Computer Self-Efficacy, CANX = Computer Anxiety.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

LIMITATIONS TO THIS STUDY

As with all empirical investigations, certain limitations must be acknowledged when interpreting the results. First, the sample size of this study was admittedly, albeit necessarily, smaller than ideal. Early in the research design, we chose to focus only on subjects who were actively employed in the accounting domain to increase the external validity of the study. We believe this constraint contributed positively to the results obtained and the conclusions derived thereof, but resulted in challenges associated with finding professionals who were willing to give of their valuable time to participate in the manipulation. The subject recruitment process took over 6 months with several subjects from a wide variety of Midwestern accounting firms (both Big 4[3] and independent) ultimately taking part.

While we are logically comfortable with the test power for those results reaching the $p <$ .05 level or below, the relatively small sample size obtained may be a contributing factor to several of the relationships being significant at the more liberal 0.10 level (statistically suggesting the results would have reached greater significance with a slightly larger sample). Further, the smaller sample size precluded us from the best practice of testing all moderators in unison. In addition, the proportion of men and women in the sample data collected was not completely balanced, although the imbalance was not severe (see Table 1). When researchers employ moderated regression approaches to testing the effects of dichotomous variables as possible moderators in a relationship between continuous variables, as was done here, unequal proportions of participants in each group leads to an increase in the likelihood of committing a Type II error (that is, a decrease in statistical power to detect a significant difference). Although the small sample size is a limitation that overall affects the research presented here, this particular issue of unbalanced groups is most directly of importance for the results presented in Table 3, where the moderating effects of gender were assessed. That said, we believe the skill set and perspective brought to this exercise by the business professionals (as opposed to random subjects or convenience samples, such as students) contributes both to the external validity of the study and the generalizability of the results and conclusions.

Another possible limitation to consider lies with the method by which the data were collected. While clearly falling into the experimental category of methods, the use of a voluntary Web-delivered vehicle for data collection raises questions of possible loss of experimental control. Given our desire to use accounting professionals as subjects, we determined that bringing them to a laboratory setting would prove inconvenient and further exacerbate the challenges in reaching a suitable sample size for analysis. Further, by allowing the subjects to participate while in their natural work setting, we believe any possible concerns or anxieties associated with a more formal experimental setting were reduced. Subjects were clearly instructed to complete the exercise in one sitting and to not begin the exercise unless they felt reasonably confident they had a minimum of 1 hour uninterrupted in which to complete the project. Start time and completion time for each subject was analyzed to ascertain the extent to which these criterion were met. In all cases, subjects participated in the exercise during normal business hours with no subject's completion time being statistically different than the mean completion time for the exercise. Given this, we believe minimal loss of experimental control occurred.

The research model tested in this study, shown in Figure 2, did not include the important construct of social influence, which is a direct determinant of intention to adopt, and whose relationship with the latter is also affected by gender (Venkatesh et al., 2003). Due to the constrained nature of the research design employed here, where participants were asked to make hypothetical adoption decisions, their ability to form valid perceptions of social influence was surely limited. Past research examining these effects found that they appear to be more relevant in contexts where mandatory usage of the specific technology is required, but not directly significant when operating in contexts where technology usage is under the control of the individual (Venkatesh & Davis, 2000). Even in mandatory settings, the effect of social influence on intentions appears to be limited to the early stages of adoption and usage. All this should not be taken to mean that we believe the construct not to be worthy of careful examination; to the contrary, we believe social factors play an important role in technology adoption within organizations. However, we believe that, due to the inherently social nature of the construct, in order for these investigations to be meaningful, they should be conducted in field settings where these effects are important to the individual adopter.

Finally, we must consider the closed set of two technologies as a possible limitation. We believe the setting for this study to be novel in the sense that it represents more than one technology under consideration, the use of actual and available technologies, and the use of subjects professionally engaged in the same domain in which the study was framed. The extent to which choices made in hypothetical scenarios, such as the ones employed here compare to those in real-life adoption settings, is related, at least partially, to the degree to which both the decision makers and technology alternatives compare to those in actual settings. In this research, the participants involved in the evaluation and selection of technologies were professionals in the field of practice from which the technologies were drawn, which we believe to be representative of the community of users who would be involved in these processes in organizational adoption scenarios. As well, the technologies chosen for this research were commercially available products. On the other hand, participants were aware that this was a hypothetical scenario that had been constructed for research purposes, and that was a likely influence on their behavior. While we cannot know the participants' state of mind while they were completing the research, the time taken by the participants to complete the tasks,

which we obtained by accessing the logs of the Website used to set up the research, provides some evidence that thought was given to the research scenario presented to them.

Nonetheless, the selection and adoption of a technology such as an organization-wide accounting package would clearly entail the review of multiple candidate packages before a final pair of two could be compared. Further, it is probable that many hours of discussion among the selection committee would occur with regard to the functional requirements upon which the final selection will be based. Given this, it is possible that the framing of the subject to simply compare and select among a choice set of two candidates may limit the richness of the true choice process. We believe future research needs to investigate this issue to determine the extent to which multiple candidates affect the choice process.

DISCUSSION

A number of recent studies in the stream of literature examining user acceptance of information technology have shown the presence of a moderating effect of the gender of the user, such that certain relationships are stronger for men than for women, and vice versa. Gender effects such as this one are useful in that they put in evidence the presence of an underlying dynamic that affects relationships of interest; however, they provide neither an explanation for the occurrence of those effects, nor a lever that can be incorporated into design considerations such that it would be possible to develop technologies enjoying wider acceptance.

The present research set out to investigate a number of different potential explanations for the observed gender effects. In particular, we identified a number of individual traits that exhibit gender differences and could plausibly be responsible for the moderating influences that have hitherto been identified as related to the gender of the users. Through an analysis of data collected from business professionals employing commercially available technologies within their professional discipline, we uncovered a number of interesting effects that we believe can form the basis for future investigations in this area. Results from our analyses are summarized in Table 12.

In light of the limitations discussed in the previous section, it is clear that our results should be regarded as preliminary and in need of replication. We believe, however, that our results contribute to a better clarification of the underlying dynamics of the observed gender effect or, at the very least, provide interesting directions for future research. We see the current status of research in this area as limiting for one major reason. While there is no doubt as to the existence of a gender effect in all of the central relationships in our models explaining user acceptance of technology, there is little that can be done, from an applied standpoint, with knowledge of such an effect. Thus, designers and marketers are presented with several moral, societal, and possibly legal constraints. Understanding how such an effect operates, on the other hand, may potentially provide both researchers and practitioners with a better understanding of the adoption process, ultimately leading to increased success in the adoption of technology.

We see our findings, shown in Table 12, as belonging to three separate groups. First, Hypotheses 1, 2, 3, 10 and 11 were included with an eye toward replicating past research and thus establishing the adequacy of our research design to examine an effect that can be repeatedly found in the extant literature. While not designed as a test of the UTAUT, which has been successfully replicated many times since it was first published, we deemed it necessary to show that our research model worked as expected according to the theory on which it was based.

Table 12. Summary of Hypotheses Testing.

	Hypothesis	Results
1	Replication of UTAUT	Supported
2	Moderating effect of biological gender	Support only for chosen technology
3	Moderating effect of psychological gender	Support only for not-chosen technology
4	Moderating effect of risk propensity	Partially supported
5	Moderating effect of self-esteem	Supported
6	Moderating effect of locus of control	Supported
7	Moderating effect of neuroticism	Support only for chosen technology
8	Moderating effect of computer self-efficacy	Support only for chosen technology
9	Moderating effect of computer anxiety	Supported
10	Computer self-efficacy as antecedent of effort expectancy	Supported
11	Computer anxiety as antecedent of effort expectancy	Supported

Results from these hypotheses confirm this, as well as the presence of some effect related to gender of the participants (biological or psychological) in the relationships. Finally, we replicated past findings about the role of computer self-efficacy and computer anxiety as determinants of perceptions of the amount of effort required to use the technology.

In the second group of hypotheses (from 4 to 7), we investigated potential candidates for the observed gender effect that can be deemed to be largely invariant over the life of the individual, such as risk propensity and personality traits. While almost by definition these cannot be manipulated or changed in any way, and may thus be deemed of more limited applicability by both researchers and practitioners, we believe knowing of their existence and importance is nonetheless valuable. At the very least, researchers can control for these constructs in future investigations and thus reduce any potential confounds, as well as better highlight the value and contribution of their research against the findings reported here. These personality traits, particularly neuroticism, seem to be involved in moderating the relationships between PE and EE, and BIs toward new technologies.

Finally, we investigated the roles that computer SE and CANX may play in moderating these relationships. Interaction effects involving these constructs showed large effect sizes when explaining variance in the dependent variable of interest, adoption intention. These large effects, in addition to the extensive literature dealing with interventions able to improve those perceptions, make these two variables particularly attractive as targets for further research. While we believe that further research, likely in the form of a research program, is required before these findings (or any others in the technology acceptance literature) can be practically applied in the design and development of technology artifacts, we do believe these results have direct implications for technology implementation and change management programs.

Hypotheses tests associated with these two variables, reported in Tables 9 and 10 in their role as moderators and in Table 11 in their role as antecedents, are very clear in their significance and direction: Both constructs have dual effects on intentions to adopt. First, higher levels of computer SE lead to higher PE associated with using the application, which in turn has a positive effect on the intention to adopt it. Furthermore, that last relationship is also strengthened for those users with higher levels of computer SE, leading to even more

positive intentions toward the technology for any levels of EE. Through these two channels, computer SE significantly impacts technology adoption. Opposite effects can be seen for CANX: Users with higher levels of CANX perceive applications as being harder to use, which leads in turn to a more limited intention to adopt them in the future. As well, CANX negatively affects that relationship, such that potential adopters with higher levels of CANX are even less likely to adopt the technology.

These findings are even more relevant when considering the existence of extensive literature bearing on the modification of these two important constructs, largely based on the seminal work of Bandura (1986, 1997). There is also extensive work published on different intervention methods in the psychology, education, and management disciplines, and even within the information systems domain itself, directly concerned with computer self-efficacy (Davis & Yi, 2004; Johnson & Marakas, 2000; Yi & Davis, 2003). As a result, we believe the design and development of implementation and change management programs associated with the introduction of new technologies in the workplace could draw from these findings and others in this domain to incorporate those in the future.

One possible issue that may limit the contribution of this research is the degree of permanence of gender effects observed in technology acceptance research. Indeed, Venkatesh et al. (2003) interpreted some of their findings as indicating that gender differences in the use of information technology may be transitory and may possibly disappear as younger generations of users are raised in an environment where technology is pervasive. If that were the case, gender differences with respect to technology use may represent an area of research that, while certainly interesting, will slowly decrease in importance as those differences disappear over time. In this scenario, the value of our findings, which were obtained from a sample of business professionals, would be diminished. We believe, however, this not to be the case, for multiple reasons.

First, to the extent that gender differences in the use of information technology and other areas of life are the result of innate biological differences between the sexes, these are by definition permanent in the timespan in which social science researchers operate. Alternatively, if those differences are the result of one or more of the social and cultural factors affecting development discussed above, those would have to had changed drastically for the younger generations (now and in the future) for these differences to be transitory. As much as societies have changed in the last few decades in this regard, this is unfortunately not the case in many areas of the world, across countries of different economic conditions and societal values and traditions. For example, research conducted in five U.S. universities (Goh, Ogan, Ahuja, Herring, & Robinson, 2007) shows that the gender of a mentor has an effect on the extent to which students develop their computer SE, where students with male mentors exhibited higher levels of the construct than students with female mentors. In particular, women students who worked with male mentors reported higher levels of computer SE than women students who worked with female mentors. We take these findings as evidence that some of the culture-based gender issues discussed above still have an important impact on how students (and, later, professionals) of both genders develop their attitudes toward technology. Indeed, Goh and colleagues concluded that, "Possibly the most important implication of this study is that IT-related programs that are committed to attracting and retaining women need to address deeply-seated stereotypes and praxis surrounding the roles of women in these departments" (p. 36).

Finally, there is evidence that, contrary to expectations that these differences may disappear or be tempered as younger generations are raised in a technology-pervasive environment, young individuals today still exhibit both gender differences in this regard, as well as difficulties using technology. The research just cited (Goh et al., 2007), as well as work by Mcilroy, Sadler, and Boojawon (2007) in the U.K., provide some evidence of this. In the first case, and in addition to the findings discussed above, the sex of the students significantly predicted their levels of computer SE, whereas age did not. In the study by McIlroy et al. (2007), between 33% and 41% of students surveyed exhibited some degree of computer phobia, as measured by two separate scales. Significantly, approximately 20% of the students exhibited moderate to high levels of computer phobia, an important minority. Moreover, the authors indicated these findings are in line with prior research going back more than 10 years; thus, the issue does not seem to have abated. Results from both studies are even more striking when considering data were collected from young populations of college students in developed countries, which one would expect, based on arguments by Venkatesh et al. (2003), to exhibit little of these difficulties. Altogether, we take these as evidence that the issue of gender differences related to information technology remains a worthwhile area of research.

CONTRIBUTION AND FUTURE RESEARCH

We believe this research makes several contributions to the rich stream of investigation into technology adoption in general and UTAUT specifically. The use of multiple technologies from which the selection was made combined with the use of actual technologies available within the domain of the professional subjects is, to our best knowledge, a first in the UTAUT literature. We also believe this study represents one of the first to meet the mandate brought forth by Davis, Venkatesh, and others to begin focusing our attention on practical applicability of the model rather than on investigating possibilities of additional explanatory power. To that end, we believe we have demonstrated UTAUT in an actual technology adoption setting and have furthered our understanding of its value thereof.

This research also represents a novel approach to modeling the relationships between the constructs of interest in order to further the comparability and consistency of the obtained results—by simultaneously including both the chosen and not-chosen technologies in the same model and constraining indicators to those that significantly loaded on their intended construct when direct effects on both intentions were present. The fact that the pattern of loadings was different between chosen and not-chosen technologies (particularly for the CANX construct) may in itself be a fruitful area of future research. It may indeed be the case that facets of the same concept play different roles in a context where comparisons between technologies are made.

It is important to note that the alternative constructs to gender tested herein displayed moderating effects with significant explanatory power over previously observed gender effects, both statistically and conceptually (i.e., they provide the "why" behind the differences). While some of these moderators are largely stable over a lifetime (i.e., neuroticism), others are more malleable (i.e., computer SE, CANX) and thus provide for actionable mechanisms by which to influence technology selection (as gender provides social, and possibly legal challenges in this regard).

We believe the differential results obtained with regard to the chosen versus not-chosen technologies are a fruitful area for further investigation. More research is needed to understand the mechanisms and reasons for these disparate effects. While the underlying UTAUT model held very well in both cases (indicating a common approach to evaluation), the proposed moderators did not play a consistent role in the comparison. Another direction for future research may involve disentangling those factors that affect the overall ability to choose from those that have effects on only the chosen or only the not-chosen technology.

This research can be considered both replicative and exploratory in nature. Given this, future research should focus on explicitly investigating the alternative choice behaviors under consideration, rather than the more traditional focus solely on the chosen technology. In the case of technology selection, behavioral alternatives should include other possible technologies in the same choice set. In the case of individual acceptance of a technology already selected for use, alternatives might be related to resistance and thus use different evaluation models and/or approaches to arrive at a specific behavioral intention.

Further, alternative research methods that can capture the richness present in field settings where technology adoption decisions happen are strongly needed. This need goes beyond conducting survey research in field settings; rather, triangulation, verification, and enrichment of these results by qualitative means should also be a focus of attention. We believe conducting this research would allow researchers to uncover other factors involved in the multidimensional and complex nature of user acceptance of technologies that may help further our understanding of the phenomena and, possibly, have important design implications.

In closing, we believe the results of this research present an opportunity for both the academic and applied research communities to further explore the nature of the technology acceptance process such that its processes can be understood in a manner that allows for prescriptive actions to be taken to improve its outcomes. It is our hope that the relevant research communities will embrace this direction.

ENDNOTES

1. Personal pronoun use is intended to be inclusive.

2.Table 2a represents the PLS measurement for the base research model under study. For ease of exposition, we have chosen to exclude representation of the measurement models for the additional variables and relationships under study. They are available from the authors upon request.

3. These represent the four largest accountancy organizations in the world (Wikipedia, 2010).

REFERENCES

Agarwal, R., & Karahanna, E. (2000). Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage. *MIS Quarterly, 24*, 665–694.

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

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ, USA: Prentice Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: H. W. Freeman and Company.

Bem, S. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42*, 155–162.

Bem, S. (1981). *Bem Sex Role Inventory: Professional manual*. Palo Alto, CA, USA: Consulting Psychologists Press.

Boudreau, M., Gefen, D., & Straub, D. (2001). Validation in IS research: A state-of-the-art assessment. *MIS Quarterly, 24*, 1-16.

Busch, T. (1995). Gender differences in self-efficacy and attitudes toward computers. *Journal of Educational Computing Research, 12*, 147–158.

Busch, T. (1996). Gender, group composition, cooperation, and self-efficacy in computer studies. *Journal of Educational Computing Research, 15*, 125–135.

Byrnes, J., Miller, D., & Schafer, W. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin, 125*, 367–383.

Campbell, T. (1997). The structure of the Bem Sex Role Inventory (BSRI): Confirmatory analysis of long and short forms. *Educational and Psychological Measurement, 57*, 118–124.

Chin, W., Marcolin, B., & Newsted, P. (2003). A partial least squares latent variable modeling approach for measuring interaction eEffects: Results from a Monte Carlo simulation study and an electronic email emotion/adoption study. *Information Systems Research, 14*, 189–17.

Costa, P., Terracciano, A., & McCrae, R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology, 81*, 322–331.

Davis, F. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*, 318–339.

Davis, F., Bagozzi, R., & Warshaw, P. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science, 35*, 982–1003.

Davis, F., & Yi, M. (2004). Improving computer skill training: Behavior modeling, symbolic mental rehearsal, and the role of knowledge structures. *Journal of Applied Psychology, 89*, 509–523.

Deaux, K. (1984). From individual differences to social categories: Analysis of a decade's research on gender. *American Psychologist, 39*, 105–116.

Deaux, K. (1985). Sex and gender. *Annual Review of Psychology, 36*, 49–81.

Eagly, A., & Wood, N. (1991). Explaining sex differences in social behavior: A meta-analytic perspective. *Personality and Social Psychology Bulletin, 17*, 306–315.

Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin, 116*, 429–456.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading, MA, USA: Addison-Wesley.

Gefen, D., Karahanna, E., & Straub, D. (2003). Trust and TAM in online shopping: An integrated model. *MIS Quarterly, 27*, 51–90.

Gefen, D., & Straub, D. (2005). A practical guide to factorial validity using PLS-Graph: Tutorial and annotated example. *Communications of the AIS, 16*, 91–109.

Goh, D., Ogan, C., Ahuja, M., Herring, S., & Robinson, J. (2007). Being the same isn't enough: Impact of male and female mentors on computer self-efficacy of cCollege students in IT-related fields. *Journal of Educational Computing Research, 37*, 19–40.

Halpern, D. (1992). *Sex differences in cognitive abilities*. Hillsdale, NJ, USA: Erlbaum.

Harrison, A., & Rainer, R. (1992). The influence of individual differences on skill in end-user computing. *Journal of Management Information Systems, 9*(1), 93–111.

Hartzel, K. (2003). How self-efficacy and gender issues affect software adoption and use. *Communications of the ACM, 6*(9), 167–171.

Heinsenn, R., Glass, C., & Knight, L. (1987). Assessing computer anxiety: Development and validation of the computer anxiety rating scale. *Computers in Human Behavior, 3*, 49–59.

Jaccard, J., Turrisi, R., & Wan, C. (1990). *Interaction effects in multiple regression*. Newbury Park, CA, USA: Sage Publications.

Jianakoplos, N., & Bernasek, A. (1998). Are women more risk averse? *Economic Inquiry, 36*, 620–630.

Johnson, R., & Marakas, G. (2000). The role of behavioral modeling in computer skills acquisition: Toward refinement of the model. *Information Systems Research, 11*, 402–417.

Johnson, R., Marakas, G., & Palmer, J. (2006). Differential social attributions toward computing technology: An empirical investigation. *International Journal of Human-Computer Studies, 64*, 446-–460.

Judge, T., Locke, E., Durham, C., & Kluger, A. (1998). Dispositional effects on job and life satisfaction: The role of core evaluations. *Journal of Applied Psychology, 83*, 17–34.

Judge, T., Thoresen, C., Pucik, V., & Welbourne, T. (1999). Managerial coping with organizational change: A dispositional perspective. *Journal of Applied Psychology, 4*, 107–122.

Koufaris, M. (2002). Applying the technology acceptance model and flow theory to online consumer behavior. *Information Systems Research, 13*, 205–223.

Langston, C., & Sykes, W. (1997). Beliefs and the Big Five. *Journal of Research in Personality, 31*, 142–165.

Lauriola, M., & Levin, I. (2000). Relating individual differences in attitude toward ambiguity to risky choices. *Journal of Behavioral Decision Making, 14*, 107–122.

Lauriola, M., & Levin, I. (2001). Personality traits and risky decision making in a controlled experimental task: An exploratory study. *Personality and Individual Differences, 31*, 215–226

Lippa, R. (2010). Sex differences in personality traits and gender-related occupational preferences across 53 nations: Testing evolutionary and social-environmental theories. *Archives of Sexual Behavior, 39*, 619–636.

Maccoby, E., & Jacklin, C. (1974). *The psychology of sex differences.* Stanford, CA, USA: Stanford University Press.

Marakas, G., Yi, M., & Johnson, R. (1998). The multilevel and multifaceted character of computer self-efficacy: Toward clarification of the construct and an integrative framework for research. *Information Systems Research, 9*, 126–163.

Mathieson, K. (1991). Predicting user intentions: Comparing the technology acceptance model with the theory of planned behavior. *Information Systems Research, 2*, 173–191.

McCrae, R., & Terracciano, A. (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology, 88*, 547–561.

Mcilroy, D., Sadler, C., & Boojawon, N. (2007). Computer phobia and computer self-efficacy: Their association with undergraduates' use of university computer facilities. *Computers in Human Behavior, 23*, 1285–1299.

McKeen, J., Guimaraes, T., & Wetherbe, J. (1994). The relationship between user participation and user satisfaction: An investigation of four contingency factors. *MIS Quarterly, 18*, 427–451.

Ndubisi, N. (2003). Women entrepreneurs and IT usage: The impact of traits. *Journal of Business in Developing Nations, 7*, 112–147.

Pavlou, P., & El Sawy, O. (2006). From IT lLeveraging competence to competitive advantage in turbulent environments: The case of new product development. *Information Systems Research, 17*, 198–227.

Penning, J., & Smidts, A. (2000). Assessing the construct validity of risk attitude. *Management Science, 46*, 1337–1348.

Powell, G., & Butterfield, A. (2003). Gender, gender identity, and aspirations to top management. *Women in Management Review, 18*, 88–96.

Ringle, C. M., Wende, S., & Will, A. (2005). SmartPLS 2.0 (beta) [Computer software], University of Hamburg. Retrieved November 1, 2008, from www.smartpls.de

Schmitt, D., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *Journal of Personality and Social Psychology, 94*, 168–182.

Sitkin, S., & Weingart, L. (1995). Determinants of risky decision-making behavior: A test of the mediating role of risk perceptions and propensity. *Academy of Management Journal, 38*, 1573–1592.

Straub, D. (1989). Validating instruments in IS research. *MIS Quarterly, 13*, 147–167.

Szajna, B. (1994). Software evaluation and choice: Predictive validation of the technology acceptance instrument. *MIS Quarterly, 18*, 319–324.

Venkatesh, V. (2000). Determinants of perceived ease of use: Integrating control, intrinsic motivation and emotion into the technology acceptance model. *Information Systems Research, 11*, 342–65.

Venkatesh, V., & Davis, F. (1996). A model of the antecedents of perceived ease of use: Development and test. *Decision Sciences, 27*, 451–481.

Venkatesh, V., & Davis, F. (2000). A theoretical extension of the technology acceptance model: Four longitudinal studies. *Management Science, 46*, 186–204.

Venkatesh, V., & Morris, M. (2000). Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior. *MIS Quarterly, 24*, 115–139.

Venkatesh, V., Morris, M., & Ackerman, P. (2000). A longitudinal field investigation of gender differences in individual technology adoption decision-making processes. *Organizational Behavior and Human Decision Processes, 83*, 33–60.

Venkatesh, V., Morris, M., Davis, G., & Davis, F. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*, 425–478.

Venkatesh, V., Morris, M., Sykes, T., & Ackerman, P. (2004). Individual reactions to new technologies in the workplace: The role of gender as a psychological construct. *Journal of Applied Social Psychology, 34*, 445–467.

Wikipedia (2010). Big Four (audit firms). Retrieved October 14, 2010, from http://en.wikipedia.org/wiki/Big_Four_(audit_firms)

Yi, M., & Davis, F. (2003). Developing and validating an observational learning model of computer software training and skill acquisition. *Information Systems Research, 14*, 146–169.

Authors' Note

The authors wish to acknowledge the support of the Advisory Council of the Department of Accounting and Information Systems, School of Business, University of Kansas. Data for this study were collected while the first author was a doctoral student at the University of Kansas.

All correspondence should be addressed to
Miguel I. Aguirre-Urreta,
School of Accountancy and Management Information Systems,
College of Commerce, DePaul University,
1 E. Jackson Blvd.,
Chicago, IL 60604
maguirr6@depaul.edu

APPENDIX A – LIST OF SURVEY ITEMS BY MEASURE

Demographics

What is your gender?

Male Female

What is your age group?

- 18 - 25
- 26 - 35
- 36 - 45
- 46 - 55
- 56 - 65
- 66 or older

What is your highest level of education?

- Some high school or less
- Graduated high school
- Vocational/technical school
- Some college
- Graduated college
- Post-graduate study

Which of the following best represents your approximate household income?

- Less than $20,000
- $20,000 - $40,000
- $40,000 - $60,000
- $60,000 - $80,000
- $80,000 - $100,000
- $100,000 - $150,000
- $150,000 or more

Which of the following best describes your job level?

- Executive / Top Management
- Middle Management
- Supervisory
- Administrative / Clerical
- Technical

BSRI

For this section, please answer the questions as to how the term describes you best according to the following scale:

1 = Never or almost never true
2 = Usually not true
3 = Sometimes but infrequently true
4 = Occasionally true
5 = Often true
6 = Usually true
7 = Always or almost always true

Adaptable
Truthful
Affectionate
Compassionate
Eager to soothe hurt feelings
Gentle
Loves children
Sensitive to the needs of others
Sympathetic
Tender
Understanding
Conceited
Warm
Aggressive
Assertive
Defends own beliefs
Dominant
Forceful
Has leadership abilities
Independent
Strong personality
Willing to take a stand
Conscientious
Willing to take risks
Conventional
Jealous
Moody
Reliable
Secretive
Tactful

Computer Anxiety

Please indicate the extent to which you agree with each of the following statements, using a 7-point scale, where,

1 = Strongly disagree
2 = Disagree
3 = Somewhat disagree
4 = Neither agree nor disagree
5 = Somewhat agree
6 = Agree
7 = Strongly agree

I look forward to using a computer on my job
I hesitate to use a computer for fear of making mistakes that I cannot correct
If given the opportunity, I would like to learn about and use computers
I have avoided computers because they are unfamiliar and somewhat intimidating to me
I feel computers are necessary tools in both educational and work settings
The challenge of learning about computers is exciting
I am confident that I can learn computer skills
Anyone can learn to use a computer if they are patient and motivated
Learning to operate computers is like learning any new skill – the more you practice, the better you become
I am afraid that if I begin to use computers I will become dependent upon them and lose some of my reasoning skills
I feel apprehensive about using computers
I have difficulty in understanding the technical aspects of computers
It scares me to think that I could cause the computer to destroy a large amount of information by hitting the wrong key

[*Note*: Reverse-coded items are in italics]

Locus of Control

My life is determined by my own actions
My life is chiefly controlled by powerful others
It's chiefly a matter of fate whether or not I have a few friends or many friends
Often there is no chance of protecting my personal interest from bad luck happenings
I am usually able to protect my personal interests
I can pretty much determine what will happen in my life
When I make plans, I am almost certain to make them work
Even if I were a good leader, I would not be made a leader unless I play up to those in positions of power
It's not always wise for me to plan too far ahead because many things turn out to be a matter of good or bad fortune
I feel like what happens in my life is mostly determined by powerful people
Whether or not I get to be a leader depends on whether or not I'm lucky enough to be in the right place at the right time

[*Note*: Reverse-coded items are in italics; same response instructions as the previous measure]

Neuroticism

My mood often goes up and down
Sometimes I feel miserable for no reason
I am an irritable person
My feelings are easily hurt
I often feel "fed up"
I am often tense or high strung
I often worry too long after an embarrassing experience
I often feel lonely
I am often troubled by feelings of guilt

[Note: Same response instructions as the previous measure]

Performance Expectancy

The system would be useful for the job under analysis
Usage of the system will allow tasks to be completed more quickly
Usage of the system will lead to increased productivity
Using the system can increase the quality of output on the job

[Note: Same response instructions as the previous measure]

Effort Expectancy

Interactions with the system would be clear and understandable
It would be easy to become skillful at using the system
The system would be easy to use
Learning to operate the system would be easy

[*Note*: Same response instructions as the previous measure]

Self-esteem

I feel that I am a person of worth, at least on an equal plane with others
I feel that I have a number of good qualities
I am able to do things as well as most people
All in all, I am inclined to feel that I am a failure
I feel that I do not have much to be proud of
On the whole, I am satisfied with myself
I wish I could have more respect for myself

[*Note*: Reverse-coded items are in italics; same response instructions as the previous measure]

Behavioral Intention

I would choose this software to be implemented in my organization
If this software package were available, I would likely recommend it for adoption in my organization
I would propose this package as a good candidate for the needs of my organization

[*Note*: Same response instructions as the previous measure]

Generalized Computer Self-Efficacy

For each of the following questions please select YES or NO. If yes, indicate how confident you are with your ability (100-point scale).

I believe I have the ability to describe how a computer works
I believe I have the ability to install new software applications on a computer
I believe I have the ability to identify and correct common operational problems with a computer
I believe I have the ability to unpack and set up a new computer
I believe I have the ability to remove information from a computer that I no longer need
I believe I have the ability to understand common operational problems with a computer
I believe I have the ability to use a computer to display or present information in a desired manner

Risk Propensity

The following 10 questions present two alternative options to choose from. Both options are equivalent in terms of their expected value. Please indicate your selection without making any mental calculations.

100 chances out of 100 to GAIN 10,000 dollars or 25 chances out of 100 to GAIN 40,000 dollars / 75 chances out of 100 to GAIN nothing

100 chances out of 100 to GAIN 20,000 dollars or 20 chances out of 100 to GAIN 100,000 dollars / 80 chances out of 100 to GAIN nothing

100 chances out of 100 to GAIN 30,000 dollars or 10 chances out of 100 to GAIN 300,000 dollars / 90 chances out of 100 to GAIN nothing

100 chances out of 100 to GAIN 40,000 dollars or 30 chances out of 100 to GAIN 133,333 dollars / 70 chances out of 100 to GAIN nothing

100 chances out of 100 to GAIN 50,000 dollars or 40 chances out of 100 to GAIN 125,000 dollars / 60 chances out of 100 to GAIN nothing

50 chances out of 100 to GAIN 15,000 dollars / 50 chances out of 100 to GAIN 5,000 dollars or 30 chances out of 100 to GAIN 33,333 dollars / 70 chances out of 100 to GAIN nothing

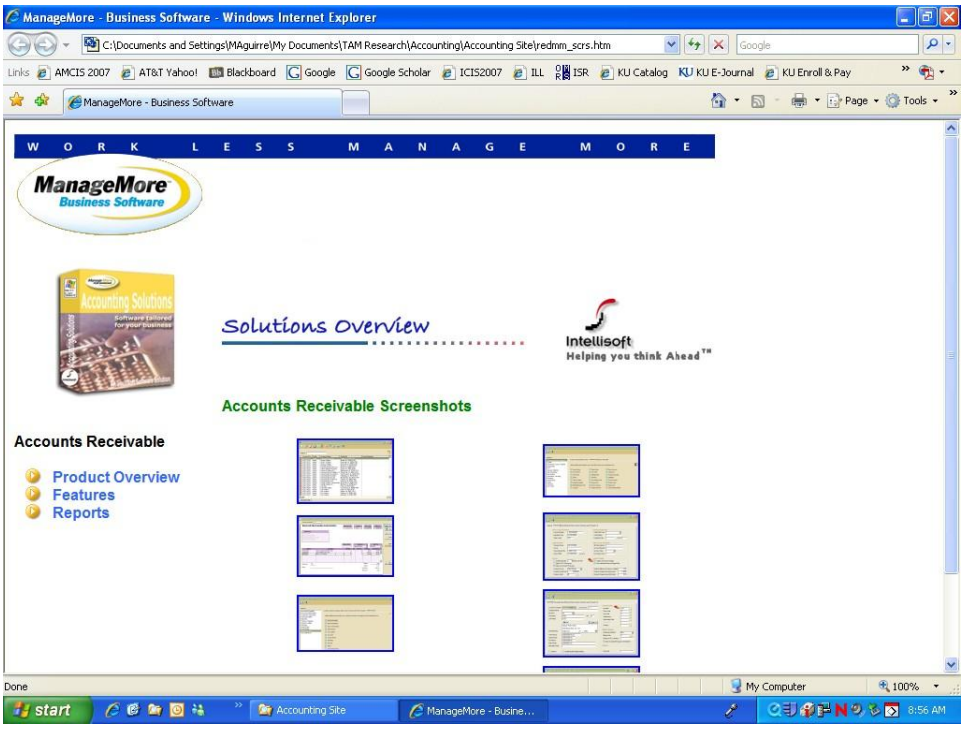
60 chances out of 100 to GAIN 25,000 dollars / 40 chances out of 100 to GAIN 12,500 dollars or 25 chances out of 100 to GAIN 80,000 dollars / 75 chances out of 100 to GAIN nothing

50 chances out of 100 to GAIN 40,000 dollars / 50 chances out of 100 to GAIN 20,000 dollars or 25 chances out of 100 to GAIN 120,000 dollars / 75 chances out of 100 to GAIN nothing

40 chances out of 100 to GAIN 50,000 dollars / 60 chances out of 100 to GAIN 33,333 dollars or 20 chances out of 100 to GAIN 200,000 dollars / 80 chances out of 100 to GAIN nothing

50 chances out of 100 to GAIN 75,000 dollars / 50 chances out of 100 to GAIN 25,000 dollars or 30 chances out of 100 to GAIN 250,000 dollars / 70 chances out of 100 to GAIN nothing

APPENDIX B – SCREENSHOTS OF EXPERIMENTAL MATERIALS


ManageMore - Business Software - Windows Internet Explorer
C:\Documents and Settings\MAguirre\My Documents\TAM Research\Accounting\Accounting Site\redmm_scrs.htm
ManageMore - Business Software
W O R K L E S S M A N A G E M O R E
ManageMore
Business Software
Solutions Overview
Intellisoft
Helping you think Ahead™
Accounts Receivable Screenshots
Accounts Receivable
Product Overview
Features
Reports
Done
My Computer
100%
start
Accounting Site
ManageMore - Busine...
8:56 AM

HUMAN TECHNOLOGY

An Interdisciplinary Journal on Humans in ICT Environments ISSN: 1795-6889

www.humantechnology.jyu.fi Volume 6 (2), November 2010, 191–211

PRODUCTIVE LOVE PROMOTION VIA AFFECTIVE TECHNOLOGY: AN APPROACH BASED ON SOCIAL PSYCHOLOGY AND PHILOSOPHY

Ramon Solves Pujol
Department of Industrial Engineering and Management
Tokyo Institute of Technology, Japan

Hiroyuki Umemuro
Department of Industrial Engineering and Management
Tokyo Institute of Technology, Japan

Abstract: *This paper proposes the use of social psychological and philosophical foundations for designing affective technology that promotes the experience of love. The adopted theoretical basis is the concept of productive love, which is heavily based on Enrich Fromm but also includes theories and scientific findings of numerous psychoanalysts, social psychologists, and philosophers. We conducted a review of the theory about the nature of love and found that social psychological and philosophical approaches differ regarding people's understandings. The findings were used to elaborate eight principles of productive love. Based on these principles, we derived criteria for designing affective technology when the objective is to promote productive love. We reviewed the existent studies on affective technologies and implemented the criteria into a system design, the Pictures' Call. A prototype of the system was pretested to illustrate how productive love technology could be based on established criteria.*

Keywords: *affective technology, productive and receptive love, care, responsibility, respect, knowledge.*

INTRODUCTION

An emerging trend in information technologies aims to support personal relationships. Studies usually approach the topic under titles such as intimacy, connectedness, awareness, or social presence. Most of these studies are predominately based on people's habits and opinions about their relationships. For instance, Kaye and Goulding (2004) based their designs on couples in stable, long-distance relationships. Van der Hoog, Keller, and Stappers (2004) used participatory designs to find out what people miss in distant relationships. Hindus, Mainwaring, Leduc, Hagström, & Bayley (2001), as well as Vetere et al. (2005), assessed users'

URN:NBN:fi:jyu-201011173091

self-reports about their activities related to their relationships. Furthermore, Vetere et al. listed several research methods that have been commonly used: online questionnaires, data logs, longitudinal focus groups, interviews, and written reflections.

Conversely, some psychologists, psychoanalysts, sociologists and philosophers claim that the average person's natural behavior may not be perfect, and suggest that scholars have a broader understanding of love than the average person, which offers the possibility to improve loving relationships. Still, consideration of sociopsychological and philosophical studies about love appears to be lacking when implementing technology in order to promote the experience of love.

The purpose of this paper, therefore, is to introduce a new research field that has been mostly unobserved to date: to use the existent theories and scientific findings on love as a basis for designing *affective technologies.* Umemuro (2009) defined affective technologies as "products to make owners pleased and proud of their owning, products that are comfortable and enjoyable in use, and/or products that provide remarkable affective experience such as excitement and deep satisfaction" (p. 3).

THE NATURE OF LOVE

Levin and Kaplan (2010) point out that a body of theoretical writing has emerged concerning love, along with efforts to validate measurement instruments. However, Levin and Kaplan note that, among scientific disciplines, only social psychology had directed systematic attention to love. Moreover, they remark, a consensus or global definition of love is not forthcoming. The following paragraphs provide an idea regarding the diverse categorizations of love within the literature.

Sorokin (1954) differentiated seven forms of love. Religious love is the love of a god or the absolute, while ethical love represents the identification of love with values such as goodness, truth, and beauty. Ontological love reflects the instrumentality of love or loving to unify, harmonize, elevate, enrich, and empower. Physical love is affirmation of the unifying, integrating, and orderings energies of the universe and biological love is love expressed sexually, romantically, and through passion. Finally, psychological love is love experienced emotionally through giving, or through receiving empathy, sympathy, kindness, and benevolence; and social love manifests in meaningful interactions or relationship with others, as driven by sharing, helping and altruism.

Later, Newcomb (1960, cited in Rubin, 1970) placed love alongside the varieties of personal attraction, such as liking, admiration, and respect. Further, Rubin (1970) compiled speculations about the nature of love, finding that love was seen as related to physical attraction, idealization, predisposition to help, the desire to share emotions and experiences, feelings of exclusiveness and absorption, felt affiliative and dependent needs, and the relative unimportance of universalistic norms in the relationship. Similarly, Averill (1985, cited in Dion & Dion, 1996) proposed four features of romantic love: idealization of the romantic partner, suddenness of onset, physiological arousal, and commitment to the well-being of the loved person. Finally, Weinstein (2007) suggested that love seems to underline terms such as empathy, compassion, acceptance, joining, reflecting, positive feedback, holding and containing environments, meeting mutual needs, and corrective emotional experience.

The most noteworthy measurements of love include several versions of the Love Attitudes Scale by Hendrick and Hendrick (1986, 1990, 1998), based on Lee's (1973, 1988) multidimensional theory of love styles, termed *eros, ludus, storge, mania, agape,* and *pragma*. And Sternberg's (1997) triangular love scale measured the intensity levels of three components: intimacy, passion, and commitment.

In a review of the categorizations of love, Weiss (2006) concluded, "Comparison shows that love styles and systems overlap to a considerable degree" (p. 214). As for commonalities among the diverse categorizations, Murstein (1988, cited in Levin & Kaplan, 2010) pointed out that, depending on upon the researcher, love had been conceptualized as an affect, attitude, behavior, or form of cognition. Further, one line of thought tends to simplify the distinction of love into two categories—rational and irrational. This perspective is well summarized by Burston (2007):

> Throughout the ages, there have been two schools of thought on the nature of love. One holds that erotic love is an involuntary passion that springs from an inner sense of lack, and thrives on illusions. Plato, Schopenhauer, Nietzsche, Freud and Lacan all subscribe to this point of view. By this account, love is "blind", and therefore, the adversary of reason, or the sober realism that characterizes the "lover of wisdom." The other school of thought, represented by Soren Kierkegaard, Max Scheler, Martin Buber and Erich Fromm claims that genuine love always includes an element of volition, is a creature of abundance, and bestows insight into the beloved that is impossible to achieve in any other way. (p. 199)

Fromm (1956) explained that the assumption that there is nothing to be learned about love is led, in part, by the generally shared confusion between the initial experience of falling in love and the permanent state of being in love. Fromm made a comprehensive bipolar categorization of love. Fromm named the first category in three ways: immature love, symbiotic union and pseudo love, indicating passive and irrational love and corresponding to the person whose character has not developed further than the receptive orientation. Fromm named the second category in two ways: mature love and genuine love, which refers to active and rational love, and explained it to be attributable to the person who has developed a productive character or orientation.

Several philosophies have paralleled the idea of irrational and rational love under different designations. For instance, Maslow (1968, 1970, cited in Le, 2005) named the idea of irrational love as Deficiency love (D-love) and rational love as Being love (B-love), while Murstein (1990, cited in Le, 2005) described benevolent love as a form of rational love. Moreover, Giddens (1992, p. 38) pointed out that "passionate love is a more or less universal phenomenon and should be differentiated… from romantic love, which is more culturally specific." LaFollette (1996, p. 194) suspected that "marginal relationships fail because they are founded on rigid love," which "is tied to a particular organism, not to a particular person with specific, embodied characteristics," and where "the lover is likely less sensitive to the beloved's interests, needs, and desires." Lastly, Bauman, (2003, p. 9) described love as "the wish to care and to preserve the object of care, a centrifugal impulse, unlike centripetal desire."

Nevertheless, measurement instruments support the bipolar understanding of love. Rubin (1970) was the first to provide an empirical measure of love, distinguishing romantic from friendship. Le (2005) developed a measurement of love in its immature form, while Sprecher and Fehr (2005) developed a compassionate love scale that can be experienced for family,

friends, peripheral ties, and all of humanity. Finally, Levin and Kaplan (2010), in the development and validation of a love scale based on Sorokin's (1954) conceptual model of love, found strong correlations between six of the seven forms of love: Only biological love, which is the love expressed sexually and romantically and through passion, was distinct among Sorokin's typology.

In accord with the theories, rational love has been empirically validated as a higher form of love. Lin and Huddleston-Casas (2005) positively correlated agape love with relationship satisfaction. Sprecher and Fehr (2005) found compassionate love to be more encompassing and experienced among family, friends, social acquaintances, and humanity. It positively associates with prosocial behavior directed toward both close others and all of humanity, while compassionate love for a specific close other was associated with the provision of social support for that person. Finally, Sprecher and Fehr (2006) found that people perceived that their self-esteem, positive mood, self-awareness, spirituality, and closeness to the other(s) increased as a result of feeling compassionate love toward other(s).

Because love means different things to different investigators, depending upon their worldviews and theoretical perspectives, Levin and Kaplan (2010) advise prospective researchers to settle on a precise operational definition that is appropriate for their specific study. Based on the aforementioned attributes of rational love, and the fact that it offers the best possibility to be learned and improved, our approach toward the design of affective technology that promotes love will be based on rational love.

In the following section we deepen our operational definition of rational love by drawing on Fromm's (1956) understanding of love, as well what other theorists have defined as the attributes of rational love. Weiss (2006, p. 324) noted that a universally accepted vocabulary on the subject of love has not yet been found. For those reasons we currently name the forms of love inspired by the receptive and productive orientations as explained by Fromm: The passive and irrational love is defined as *receptive love,* and the active and rational love is defined as *productive love.*

PRODUCTIVE LOVE: OPPORTUNITY FOR IMPROVEMENT

Based on the idea that productive love is not an irrational passion but a voluntary action that can be learned and maintained, the aim of this study is to improve productive love relationships in couples, the family, or among friends. In order to establish a consistent basis for using productive love in technology design, productive love attributes need to be identified, as well as a means to differentiate it from what has been described as its antithesis, receptive love. Moreover, it is valuable to assess whether the productive love principles are understandable and applicable for contemporary individuals. A literature review on receptive love and productive love has been carried out and the results of a brainstorming discussion are presented and analyzed.

Elements of Receptive Love

Fromm (1956) described immature love as a symbiotic union resulting from the biological pattern in the relationship between the pregnant mother and the fetus, and is represented in adult

relationships in the form of masochism and submission. Fromm explained that, consequently, the receptive character focuses on acquiring and possessing the other person or the other's love.

Likewise, Maslow (1968, 1970, as cited in Le, 2005) conceived D-love to reflect a lower love in the service of needs. Furthermore, for Rubin (1970), the concept of love as just involuntary passion belongs to a restricted view, which is the understanding of love as an emotion, a need, or a set of behaviors. The linkage to a particular target implies a narrower perspective than that held by those who regard love as an aspect of the individual's personality or experience, which transcends particular persons and situations. Gelbond (1979) explained that D-love is needful or selfish love, as with all forms of self-centered love in which two people love one another only in the sense that each meets the deficiencies or needs of the other in some way.

Similarly, Loy (2002, as cited in Le, 2005) explained that love becomes a means to ground oneself and fulfill one's sense of something otherwise lacking, which is rooted in the ego's need to ground itself through objectification of self and others. Unless one is able to transcend the self–object duality, any love or attempt of love will never be completely satisfying or adequate. Accordingly, Pickering (2009) recalled the perspectives of Spinoza, Freud, and Grostein. Spinoza pointed out, "Erotic passion may give rise to frustration, anger, and hate as inevitable corollaries of egocentric desire" (in Pickering, 2009, p. 27). Freud saw the "overvaluation of the beloved and denigration of oneself as a form of displaced primary narcissism," which is but self-love (in Pickering, 2009, p. 213). And Grostein explained that "the real person and the real relationship are only disappointing because we have failed to keep our appointment with the other's reality" (in Pickering, 2009, p. 12).

Finally, Le (2005, p.75) explained immature love "remains ego centered and is dependent on self–other distinctions and relationships." Le's measurement of love in its immature form contained these items: (a) love under the condition of being loved, (b) love under the condition of being pleased, (c) the value of receiving love over giving love, (d) expectations of some return for one's love, (e) giving value to commitment and security, (f) belief that to love someone needs practice, (g) belief that it is easier to love someone with good qualities, (h) belief that nonreciprocity of love is less satisfying, (i) the feeling of love without reason, and (j) loving a spouse and children because they are part of oneself.

In conclusion, our study found the support of several philosophers and social psychology scientists for the proposed receptive love concept as the antithesis of productive love.

Elements of Productive Love

Sorokin (1954, p.13) regarded the social aspect of love as "the meaningful interaction–or relationship–between two or more persons where the aspirations and aims of one person are shared and helped in their realization by other persons." Fromm (1956) explained that mature love is a voluntary action, rather than a pleasant sensational experience as a matter of chance. Love, in its productive character, is the union under the condition of preserving one's integrity and individuality, and the active striving for the growth and happiness of the loved person. For Fromm, giving is the foremost basic element of all forms of love. Moreover, Fromm (1956, p. 20) also cited Marx's (1844) comment, "…you can exchange love only for love,…" as indicating that there is no need to care about the fairness of a relationship. Fromm pointed out that mature love depends on the character development of the person; that is, the overcoming of one's narcissism is the condition for mature love. The opposite pole of narcissism is objectivity, which is the capacity to

see things as they are. Similarly Pickering (2009, p. 83) saw that "erroneous views of each other form the greatest impediment to love." Lastly, Fromm's mature love goes beyond the element of giving: The active character of love always implies certain and mutually interdependent basic elements, common to all forms of love. These elements are care, responsibility, respect, and knowledge—a syndrome of attitudes that are found as well in the mature persons.

In the same vein, Maslow's B-love (1968, 1970, as cited in Le, 2005) is the appreciation of others and the appreciation of the experience of love per se. Maslow stated that self-actualized persons are freer from dependency and thus they are able to engage in B-love. Maslow pointed out as well that B-love is a richer, higher-level, and more valuable subjective experience than D-love, which all B-lovers have also experienced. B-love was further explained by Gelbond (1979, p. 75) as "the love for the very being or presence of another person, for the qualities, gifts, acts, and aspirations of that person. In B-love one gives of oneself without necessarily expecting any return." Additionally, Gelbond explained B-love as the tendency toward more and more complete spontaneity, the dropping of defenses and roles, and growth in intimacy, honesty, and self-expression. Maslow stressed as well the following aspects: the absence of jealousy, eagerness for the growth of the other, essential affirmation and respect for the other's individuality, and enjoyment that includes fun, exuberance, gaiety, and the absence of anxiety. In the same way, Gelbond explained that May (1969) defined love as "a delight in the presence of the other person, and affirming of his value and development as much as one's own" (in Gelbond, 1979, p. 75). Murstein (1990, as cited in Le, 2005) defined benevolent love as the intention to help and to give to another person, without shades of self-interest. This form of love includes spontaneity, motivation by selflessness, impartiality, and creativity. Similarly, Shinebourne (2006) drew on the definitions of Kierkegaard and Levinas. Kierkegaard (1995, cited in Shinebourne, 2006) argued that the person who loves does not seek his/her own, because he gives in precisely such a way that it looks as if the gift were the recipient's property. And Levinas (2001, cited in Shinebourne, 2006) suggested that the relation is always nonreciprocal: Love exists without worrying about being loved. Pickering reviewed Levinas and Steiner as well. Levinas saw "the principle of seeking to serve another without thought of reciprocation as the most fundamental starting point for ethical relations" (in Pickering, 2009, p. 26), while Steiner saw "in the luminosity of authentic love there is a sense of flourishing and emerging into the fullness of our enlightened being" (in Pickering, 2009, p. 8). Finally, Pickering added, "When we move to mature love based on appreciation of others then love grows exponentially" (p. 212).

The presented reviews of productive love, as well as its antithesis, receptive love, serve as the foundation for the elaboration of the components of productive love. Further, Fromm's basic elements of love—care, responsibility, respect, and knowledge—are studied in detail in the following section.

Socio-Psychological Approach Versus People's Actual Understanding

As previously discussed, Fromm (1956) proposed four basic elements common to all forms of love: care, responsibility, respect, and knowledge. In order to find out what these concepts mean in today's world and so that we could compare it with the theory, a brainstorming discussion on each of these elements was conducted. The participants were asked to freely talk about what they think it means to care, to be responsible, to respect, and to share knowledge within a relationship. The given time was 20 minutes to talk about each element. The

participants were recruited within the university; they were a university professor and four graduate students from various cultures: Two participants were Japanese, one was Chinese-American, one Ecuadorian, and one was Spanish. One was female and four were males; they were aged between 22 and 44 years old. (M = 32, SD = 7.28). The results demonstrated that the participants' views about the meaning of the elements were similar to each other, suggesting that their views could represent people´s actual understanding. Furthermore, in order to better understand the discrepancies between the theories and the brainstorming, the results are accompanied by some other theories and findings of other theorists and researchers.

Fromm's Elements of Love: Care

In defining the notion of care, the participants included the care for the elderly or the control of a parent over an adolescent. However, most of the ideas were oriented toward caring within a relationship, such as a couple within an environment of equality. This duality of understanding is in line with Graham (1983) and Ungerson (1983), both cited in Ungerson (2005), who made a basic distinction for the use of the word *care*. They differentiated between caring about, defined within feelings terms, and caring for, defined as task-oriented activity and, hence, most closely defined by *work*. The latter seems close to Fromm's (1956) proposal that the essence of love is to labor for something: "the active concern for the life and growth of that which we love" (p. 22). Similarly, Mayeroff (1972, p. 1) suggested, "To care for another person, in the most significant sense, is to help him grow and actualize himself." In addition, Mayeroff described eight major components for caring: knowing, alternating rhythms, patience, honesty, trust, humility, hope, and courage.

Fromm's Elements of Love: Responsibility

The participants interpreted the concept of responsibility as sharing the blame, keeping promises, and standing on someone's side. Likewise, Fromm (1956) pointed out that responsibility is often meant to denote duty, something imposed from the outside. However to Fromm, responsibility is implied by care and concern, a voluntary act as being able and ready to respond: "is my response to the needs, expressed or unexpressed, of another human being" (p. 22). In the case of a mother and infant, it refers mainly to the care for physical needs. In the love between adults, it refers mainly to the psychic needs of the other person, which can be expressed or unexpressed. Consequently, it is necessary to stress that responsibility is a response as well as an ability to respond. As a result, this element demonstrates that great differences could be obtained between an empirical assessment and a philosophical perspective. As for the importance of responsibility in relationship with love, Buber (1958, cited in Shinebourne, 2006) conceived love as responsibility for the other: "Love is the responsibility of the I for thou" (p. 29).

Fromm's Elements of Love: Respect

Most of the ideas provided in the brainstorming discussion focused on respect in terms of not interfering in the other's ways. This is close to Fromm (1956), who defined the concept, "in accordance with the root of the word (respicere, to look at), the ability to see a person as he is, to be aware of his unique individuality" (p. 22).

On the other hand, some participants pointed out the possibility that too much respect can imply too much credit, and thus less commitment. This does not seem to be in accord with Fromm's understanding, which explained the importance of respect as preventing responsibility from deteriorating into domination and possessiveness. "Respect is not fear and awe… respect is possible only if I have achieved independence; If I can stand and walk without needing crutches, without having to dominate and exploit anyone else" (Fromm, 1956 p. 22). Accordingly, Pickering (2009) pointed that "a good relationship is predicated on 'the capacity to be alone'… as well as capacity to be together" (p. 212).

As for researchers' understanding today, Hendrick and Hendrick (2006) suggested that respect should be viewed in both structure and content. Structurally, respect can be seen as an attitude. Respect as an attitude consists of affect, cognition, and behavioral tendencies. Respect as content can be viewed as having two primary components: equality/mutuality and caring/supportiveness; the latter seems to be in line with Fromm's idea. Additionally, Hendrick and Hendrick proposed respect to be positively correlated with eros, storge, agape, satisfaction, commitment, and, except for older couples, self-disclosure. Moreover, respect as well correlated negatively with ludus, permissiveness, and instrumentality.

Fromm's Elements of Love: Knowledge

The participants tended to talk about knowledge that seemed more significant or valuable, which is in accord with Fromm (1956). Among the layers of knowledge, he posits, the one that is an aspect for love is the one that does not stay at the periphery, but penetrates the core. To Fromm, knowledge, "is possible only when I can transcend the concern for myself and see the other person in his own terms" (p. 22–23). However, objective knowledge is something that was not commented on by the participants. Besides, Fromm pointed that "knowledge would be empty if it were not motivated by concern" (p. 23).

In a study of the organization of partners' beliefs, Showers and Limke (2006) suggested that there are different ways to organize the beliefs of a partner that are activated in a particular situation. These can be organized in two types that fall on a continuum from compartmentalized knowledge (i.e., positive and negative beliefs are segregated into separate categories of partner knowledge) to integrated knowledge (i.e., positive and negative beliefs frequently appear within the same categories of knowledge). Compartmentalized knowledge is more efficient and thus easier to maintain. It may be used for a partner with many possible attributes, may be more optimistic, and may result in liking the partner more and being more satisfied with the relationship. Integrated knowledge requires more effort and may be used in stressful situations. Integrated knowledge seems closer to Fromm's idea.

Fromm's Elements of Love: Conclusion

The brainstorming discussion we conducted demonstrated that people's understanding of concepts related to love today differ considerably to Fromm's (1956) theories, illustrating how social psychology and philosophical methods lead to different and better understandings of love. Further, the brainstorming results were perceived as a warning regarding making assumptions and avoiding misinterpretation of the productive love principles in a contemporary environment.

Summary and Final Productive Love Principles

We brought together all the reviewed theories for a summation of the principles of productive love, taking into account the differences between today's understanding and Fromm's (1956) elucidation of the four basic elements of love. The literature and the empirical data allow us, then, to define eight principles of productive love.

The first principle, *giving, not exchange or egoism,* embraces doing things for the other without expecting a return. The second principle, *care, not involuntary love,* is the most important thing that can be done for the other, since it implies assisting the life and growth of the other. In order to be able to care one has to undertake the third principle, *responsibility, not irrationality,* which is to listen and respond to the other's needs. The fourth principle, *respect, not exploitation,* is needed in order to prevent responsibility from deteriorating into domination and possessiveness. The fifth principle, *realistic knowledge, not delusion,* is essential to guide care and responsibility, while the sixth principle, e*njoyment, not evaluation,* is included in order to motivate concern for learning about the other. The seventh principle, *freedom, not a feeling of duty*, is a condition to experience without restraint the previous principles. The eighth principle, *self-growth, not dependency,* is the base condition which makes possible the rest of the principles in a larger or smaller degree.

PRODUCTIVE LOVE CRITERIA FOR TECHNOLOGY DESIGN

Although many modern technologies may distract us from active caring, responsibility or loving, our objective is to create new technologies that move the attention away from the technology itself and refocus it on the person's ability to love. That could be done through two different settings: computer mediated communication (CMC) or face-to-face (FtF) communication.

The transmission of direct information about inner feelings may be the fundamental reason for getting together and talking intimately. However sometimes FtF communication is not possible, such as in the case of long-distance relationships. Furthermore, for some people or some cultures, it is possible that FtF communication is difficult, for instance, some people may find it easier to express their feelings in a letter than in person. On that direction, Briggle (2008) explained that filtration cues in computer-mediated communication have been viewed positively. McKenna et al. (2002, cited in Briggle, 2008, p. 225) pointed out that "many may feel less vulnerable in mediated situations—outside the gaze of the other—and thus find it easier to express their real 'selves,' including their intimate feelings of love and care."

However, productive love technology, that is, affective technology with the objective to promote productive love, should not simulate the other's presence, or replace a genuine encounter if it leads to FtF communication being substituted by CMC. On the other hand productive love technology should assist the existent communication, for instance, by *reconfiguration*. Briggle (2008, p. 226) explains reconfiguration as the case where "the technology mediates the experience by making visible what was previously hidden." This mediated vision seems not limited to CMC outcomes, but could be applied for enhancing FtF relationships as well.

In order to make available the final productive love principles for designing productive love technology, this section presents an initial proposal of design criteria for the use of the eight principles of productive love listed above. Additionally, we provide examples of how these principles could be taken into account in order to explicitly design tools that target the promotion of love.

Giving, Not Exchange or Egoism

A productive love technology can support actions such as buying a present, sending a greeting, writing a poem, and so on. Moreover, the technology can automatically remind us of tasks that we may want to do for the other, such as giving a birthday present or visiting grandparents once in a while. On the other hand, the productive love technology should avoid putting a premium on the user actions by obtaining points or evaluation, since that may divert focus to the return rather than the sentiments expressed by the action

Care, Not Involuntary Love

In FtF encounters, productive love technology can motivate care by promoting communication and thus can contribute to making the people know what the other may need. This could be done, for example, by suggesting topics to talk about, activities to do together, or games to play together.

Through CMC, productive love technology could provide mutual information, such as surfacing ideas or realizations about the other's dreams, dislikes, or moods, or even sharing virtual common spaces. Being in contact with each other's reality may let us see the other's needs and motivate care as a voluntary act. Suggesting actions to take or providing information without the people's voluntary action may seem at first contrary to the spirit of active caring that is essential to productive love. However, the automatic action of the device should not be the activity or action of the caring, but just a reminder. As an example, one has to remember to water flowers to care for them. Thus, setting an alarm or putting the flowers in a visible place is a way that may reinforce the active caring.

Responsibility, Not Irrationality

In FtF encounters, productive love technology could support responsibility by providing environments where people are able to relax and bring more attention into the other, and thus feel concern about the other's possible problems. This could be accomplished through using relaxing music or nature sounds, or projecting peaceful scenes within the shared environment. Moreover investigating recipes to cook together, games to play, or prompting yoga and meditation-like exercises could support relaxation.

In CMC, technology can support responsibility by facilitating an answer when one receives some information about the other and cares about it. For instance, it can support writing, voice, and videoconference. Also actions such as buying something that one sees that the other needs, or planning to meet up are responsibility acts that could be supported. In that sense, a real-time technology could be more supportive for the response, just as a phone call

can facilitate the response ability versus sending a letter. Furthermore, being informed about the person allows a user to better respond to technology-enhanced opportunities.

Respect, Not Exploitation

If the technology could provide information that supports the person—in both FtF encounters and CMC—seeing the other realistically, as he or she is, it then supports the meaning of respect that is adopted here. Such support could be generated through new or enhanced knowledge about the other, or by facilitating the attention toward the other, as seen in the first three principles.

In addition, productive love technology should avoid situations in which the other person is "acquired," or where the user is getting something from the other, such as obtaining personal favors or completing tasks through the other. Such processes could create a situation where the user feels obliged to the other. Furthermore, productive love technology should not facilitate differences in rank between the people, which may lead to situations of domination.

Realistic Knowledge, Not Delusion

In both FtF encounters and CMC, knowing about the other person could be facilitated by providing personal information through conversation, answering questions, writing, or by sharing personal images or objects. Moreover, information could be collected automatically, for instance, by using sensory technology that may collect images, sound, movement, presence, and so on. Importantly, however, such automatic information gathering should not invade the person's privacy. It is possible as well to estimate the person's activity or feelings from data collected in unobtrusive ways (Eguez Guevara & Umemuro, 2010).

On the other hand, and in particular in CMC, productive love technology should avoid showing an unrealistic or partial image of the partner, such as highlighting only good points or showing too many signs of affection through, for instance, exaggerated emoticons. Biased or incomplete perspectives could create an idealization of the partner, which may lead to "'hyperpersonal communication,' or the state in which CMC becomes more desirable than FtF interactions" (Walther, 1992; Walther et al., 1994, in Briggle, 2008, p. 225)

Enjoyment, Not Evaluation

Although applications similar to a game might provide initial mutual interest among users, engagement with the system ultimately should contribute to engagement with the other person. In order to move the pleasure of interaction with the device into the personal relationship, the actions carried out within the productive love technology should be as close to reality as possible, and as distant as possible from fantasy, such as activities carried out by fictional characters in fictional contexts in most videogames. Information regarding what the other does, what he/she is interested in, and so on, may prove a stimulus for thinking about the object of affection.

Importance is also attached to having productive love technology diminish the differences between people regarding rank, status, comparisons or competition, and personal scoring, all of which may promote evaluation and criticism of the other person. Therefore, if actions are assessed in any way, a high number of actions or measures should not be

evaluated as better or worse. Likewise, the nature of the actions, such as buying a present or asking "How are you?" should not be established as having different value. All possible actions should be shown as valuable to the receiver and his/her understanding of the specific contexts and appreciation for them.

Freedom, Not a Feeling of Duty

In order to facilitate every user feeling it easy to act and express freely using the productive love technology, the technology should accept a wide range of actions, as opposed to fixed and predetermined ones. Additionally, there should not be rules determining "good" or "bad" actions, which may limit the users' expression. For instance, no topic for discussion should be considered inherently bad. Nor should the user feel obligated toward stereotyped actions, such as using emoticons, or toward imposed duties, such as defined tasks that may not represent either the user or the receiver.

Moreover, respect for privacy is requisite for not limiting freedom. This can be sustained by keeping personal information private and by discouraging the use of devices (e. g., microphones or cameras) if those would invade one's privacy or lead to (perceived or actual) control over a person

Self-growth, Not Dependency

Acquiring maturity is not a simple process. Still, productive love technology could target and enhance it through local elements, such as mirroring the person's own changes and improvements. For instance, simply reviewing their own pictures may make people reflect on the emotions of the moment, what changes have occurred, and how they feel about them now. Moreover, if the user has some particular habits related to the use of the productive love technology, those habits could be tracked and illustrated; for instance, the time spent communicating with others. Furthermore, productive love technology could target self-improvement as a whole by supporting techniques that have been demonstrated to lead to self-actualization, such as mediation. See Sorokin (1954) for possibly the finest available summary of techniques of altruistic transformation.

On the other hand, productive love technology should not lead to dependency that is in opposition to personal growth. For instance, as a very basic example, a device that helps the user to wake up another person in the morning would be better if it also facilitates the receiver learning how to do it on his/her own.

RELATED STUDIES ON AFFECTIVE TECHNOLOGIES

In order to apply the reviewed findings and design criteria within a practical technological device, we briefly review systems that provide interactions related to loving relationships. We discuss the types of interaction, including the identification of the gaps related to the viewpoint of productive love, and suggestions for improvement.

Several systems aim to promote connectedness. For example, Hindus et al. (2001) proposed a simple and lightweight means of distance communication in the *Casablanca*

project. The project included the *Lampshade* and the *Intentional Presence*, which glowed when remote users manually indicate their presence, while the *Pulling the Curtain IPL* depicted the user as a flower in the remote location.

Awareness systems with a higher degree of intimacy have been explored through metaphoric representations. For instance, Strong and Gaver (1996) proposed three systems: the *Feather*, which lets a plume float on a transparent tube, and the *Scent*, which lets a fragrance vaporize into the room, with both systems activated when a distant partner touched a frame of a picture of the couple; and the *Shaker*, which transmitted a vibration while maintaining timing and amplitude of movements. Hindus et al. (2001) created the *In Touch*, which transmitted touch into glowing light, warmth, or vibration. Chang, Resner, Koerner, Wang, and Ishii (2001) proposed the *LumiTouch* picture frame, which lighted when the remote user touched a picture. Lastly, Chung, Lee, and Selker (2006) created the *Lover's Cups*, which transmitted the movement of the cup into illumination of another cup in a remote location. These devices support awareness through several kinds of actions and representations. However, metaphoric representations can be ambiguous in their interpretation; therefore, they may fail to provide objective knowledge about the partner, which is needed as a basis for growing productive love.

Some proposals reproduce a companion's actions in several nonmetaphoric fashions. Gibbs, Vetere, Bunyan, and Howard (2005) created two systems, the *Secret Touch*, which allowed sharing tactile impulses within pockets, and the *Hug Over a Distance*, where jackets allowed exchanging a virtual hug. Similarly, the *iFeel_IM*, by Tsetserukou et al. (2009), provided realistic hugs over distance accompanied by butterflies in the stomach and shivers in the body's spine using an augmented reality vest. Other devices allow sharing personal information. For instance, Hindus et al. (2001) enabled two houses to share a writing surface with the *Scanboard*. Vetere et al. (2005) allowed leaving messages around to be found serendipitously with the *i. fuzz*. Gibbs et al. (2005) allowed the exchanging of messages while they were being composed though the *Synchromate*. And more recently, Romero et al. (2007) created the *ToTell List*, where pictures or messages acted as a postcard that functioned as a reminder of interesting moments and experiences to talk about. Each of these devices provided more intimate or objective information about the partner through a voluntary action of the user.

Remote location and activity can be informed automatically as well. Brown et al. (2007) demonstrated the *Whereabouts Clock* to serve as positioning representation in which icons of family members are plotted based on the location of their cell phones. A more intimate approach was the *Sensing Beds* by Goodman and Misilim (2003), which transmitted the remote user's position by heating a parallel spot. Hindus et al. (2001) visually showed activity from a remote location by turning on and increasing the brightness of the *Presence Light* and showed general noise levels at a remote location through the synchronized *CommuteBoard*. Siio, Rowan, Mima, and Mynatt (2003) used the *Coffee Aroma Generator* as a clear and natural representation of coffee that is being made in a distant location. Moreover, Yashikida and Umemuro (2008) presented the *Close to you*, which transmitted prepared sounds or smells that suggested several actions of the counterpart. Huijnen, IJsselsteijn, Markopoulos, and de Ruyter (2004) achieved social presence by displaying a processed visual representation or a full video of a remote friend watching the same television program. Furthermore, Takashima and Umemuro (2008) displayed the same program in a submonitor next to the main television in order to activate the communication among family members. Van der Hoog et al. (2004) placed the *Gustbowl* at the

home entrance, which sent to the remote individual pictures of things like keys when dropped into it. Lastly, Sorakubo and Umemuro (2008) created the *Two-nearly,* which allowed the house of the family members who are living in a physically distant location to be seen through an analogy of a window at the present location. Although these devices succeed in automatically transmitting information about location and activities of the companion, under the point of view of productive love, the output of the devices are restricted regarding the transmission of the person's inner feelings, which would be highly valuable in promoting productive love.

Finally some devices automatically transmit intimate information about the other person's condition. For example, Hindus et al. (2001) simply connected two remote locations with high quality audio through the *RoomLink,* allowing listening to the other person´s activities. Mynatt, Rowan, Craighill, and Jacobs (2001) provided abstract visualizations of information about the well-being of an elderly relative with the *Digital Family Portrait*. Finally, Kaye and Goulding (2004) transmitted the heartbeat and hand warmth via the *Hand Holding* device.

The technology research above, to varying extents, is useful toward the objectives of productive love technology. However most of these systems were designed based on the self-perceived beliefs of either the designers or the users about love and related relationships, rather than on explicit theoretic and scientific considerations on what is necessary for creating, supporting, or enhancing the experience of love. Thus the principles of productive love that our study follows might provide a significant contribution toward the design of productive love technology.

PICTURES' CALL SYSTEM PROTOTYPE

The objective of productive love technology is to create an environment where the users can experience the principles related to productive love. The described design criteria offer space for many types of technologies, for instance, in aiding FtF communication, for supporting distant relationships, or even for helping personal development. This section describes a prototype of a system, *Pictures' Call*, which intends to embrace the proposed principles for productive love in a case where the users live separately and spend some time without meeting each other, whether it is just some days, or several months. The empirical study presented here is intended to illustrate and model how technology development can fulfill specified research-based criteria.

Pictures' Call System Description

As seen in the productive love principles and productive love design criteria sections, objective knowledge is essential to support care, respect, and responsibility as basic principles of productive love. Consequently, the focus of supporting the exchange of valued and valuable information was established as the key priority of the system intended to develop the qualities of productive love through the Pictures' Call system. What is more, we wanted to automatically initiate activities that can promote productive love. Pictures' Call is a bidirectional, dual-component system that automatically takes, sends, and displays everyday images between two users. Pictures' Call is built in the Java environment, connected via the Internet, and is configured for tablet PCs at the site of each user. Details of the process and how the device meets the rest of the principles are explained as follows.

Pictures' Call automatically takes pictures of each of the users and sends to the other within a certain time frame. The automated system frees the users from a task that could be seen as a duty, and perhaps abandoned. The automatism intends to provide better engagement with the system and the interaction with the other person in a fun and nonjudgmental way, therefore supporting the sixth principle of productive love—enjoyment, not evaluation.

The first device holds the capture system, which is designed for placement at home or other habitual environment. When movement is detected, the system takes several pictures at different times of the day and at random intervals: In this way, the photographed person appears natural, doing daily tasks and not posing. Moreover the user has no control over the picture that may be taken; neither have the option to look at what has been taken before it is transmitted. Such a simple and natural process eliminates the need for the photographed individual from having to choose the best smiling picture to be sent; therefore, the user's condition is represented realistically, which is the objective of the Pictures' Call device. Users who wish to capture specific actions or happenings can use their digital cameras and video conference systems, activities which are distinct from, but may complement, the Pictures' Call device.

The capture system accomplishes the seventh principle—freedom, not a feeling of duty—firstly by displaying a mirror image when in operation mode. The mirror function informs the user of the type of view of him/her that may be taken and thus preventing the users from feeling that they are being secretly observed. Secondly, the users are able to turn the device off or to move it to locations around the home that they like or that do not interfere with their privacy. Moreover, the pictures are sent after a security delay that allows the users to erase the images within 2 hours and to stop the system for 2 hours by pressing the privacy button (see Figure 1). The fact that the privacy button takes over from the real-time picture exchange was a significant issue during the design process. While a real-time transfer and viewing of pictures could be a high motivator for responding to them, and thus support responsibility and enjoyment through engagement, there is a downside to this as well. Because the seventh principle—freedom, not a feeling of duty—is a condition for being able to experience without restraining the other principles, we agreed that it was most important to support this principle through technology that did not confine the user.

The second device holds the receiver system. It takes delivery of the pictures from a distant location and displays them as a slide show as they become available. The newest pictures supplant

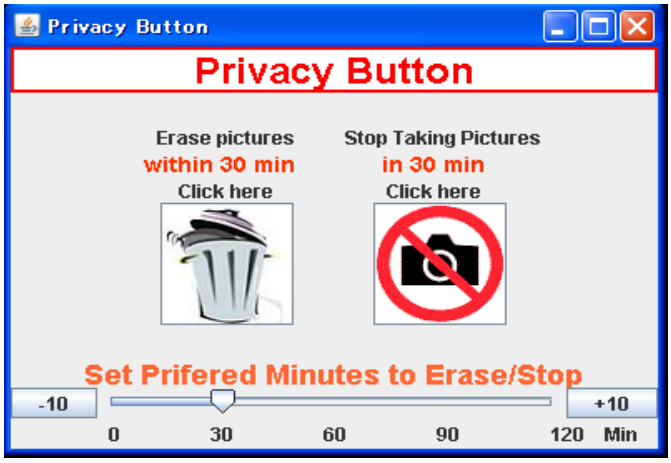


Figure 1. Privacy button of the Pictures' Call capture system.

the oldest ones, keeping a slideshow of up to 20 pictures. Displaying realistic pictures aims to let the receiver see the partner as he/she is, which intends to contribute to the fifth principle—realistic knowledge, not delusion—as well as the fourth principle—respect, not exploitation.

Additionally, the receiver system allows commenting on the pictures through an edit function on the touch panel display, and then sends them back to the original person in real-time (see Figure 2 for a screen shot of the edit function). This encourages the person to take an action related to the loved one who appears in the picture, which may motivate the second principle—care, not involuntary love—and the third principle—responsibility, not irrationality. Moreover, taking an action for the other, as a response to the possible clues conveyed by the picture, can be a selfless act, which strives for the first principle—giving, not exchange or egoism. Nevertheless, the nature of the comments is completely up to the user. Therefore it does not limit the seventh principle—freedom, not a feeling of duty.

Furthermore, Pictures' Call does not generate any obligation on the receiver. Because the user has not seen his/her own pictures, she/he should not expect any specific response. Moreover, the number of sent pictures is not clearly defined but can range from zero to seven in a day, depending on the time that the user spends in front of the capture system's camera. The undefined number of pictures a day keeps the user away from knowing that the other user should have received already a determined number of pictures. This intends to decrease the expectations of responses of the user who appears in the pictures, and the feeling of obligation to respond to the pictures from the receiver, contributing as well to the seventh principle—freedom, not a feeling of duty.

Finally, receiving commented pictures of oneself, which the user has not seen yet, serves as a mirror that supports personal awareness. This can support the eighth principle—self-growth, not dependency. The commented pictures are not automatically erased but remain available for further review. However, the receiver of the commented picture can delete the edit, or edit the edit and resend it, which encourages further communication as well as the first principle—giving, not exchange or egoism, second principle—care, not involuntary love—and the third principle—responsibility, not irrationality.

Figure 2. Edit function of the Pictures' Call receiver system.

Pictures' Call System, Tentative Evaluation

This section intends to provide a brief illustration of how productive love technology can be evaluated in principle, as well as provide a preliminary feedback from six users specifically on the Pictures' Call system. The users were two females and four males, aged between 21 and 55 years old. (M = 36.16, SD = 11.99). Three of them were Japanese: a married couple who through Picture's Call remotely connected with their grandson. The rest involved a Spanish citizen living in Japan who was connected to his sister and a close friend living in Spain.

The testing took two weeks for each group of users; the instructions given to the users were to place the capture system in a place where they would feel comfortable to share images of themselves and to spend some time seeing the other´s pictures at the receiver system and try to send comments. No restrictions were given. The users became familiar with the system from the very beginning; however the system failed to send the pictures several times during the testing, which created some concern in the users about their correct usage of the system.

The users answered three questionnaires about productive love and three questionnaires about system use; the questionnaires were administrated in Japanese and in Spanish, depending of each user mother tongue. Any quotes drawn from these questionnaires for this paper have been translated by the authors.

In terms of productive love, one questionnaire was completed before system usage and assessed how much they valued the principles of productive love. A second questionnaire was completed before and after using the system and assessed if the users experienced changes in their relationships in terms of productive love. The third questionnaire was completed after the system use and assessed if the system had promoted their productive love. All three questionnaires contained 24 items each, representing the eight principles of productive love.

Regarding the system use, the fourth questionnaire assessed the costs of communication, inspired by the Affective Benefits and Costs of Communication (ABC) questionnaire, created by Romero et al. (2007). However, for this research, we created a new questionnaire with nine items about the creation of expectations, creation of obligations, and privacy invasion. The fifth questionnaire assessed three of the Nielsen's (1993) five criteria of usability: learnability, efficiency, and satisfaction. In the final questionnaire, three open-ended questions asked their impressions about the system.

Although the users participated with someone they loved and had no productive love objectives at first, the results from the first questionnaire showed that all users evaluated positively the productive love principles. However, the second questionnaire indicated that the overall difference in productive love relationships between the users before and after using the system was minimal. This seems attributable to having used the system for only 2 weeks, which may be not long enough for the actions carried out for productive love to be reflected on the users' relationships. On the other hand, the results of the third questionnaire showed that the system succeeded to support all the principles of productive love for all the users.

The open-ended questions showed that all of the users highly enjoyed receiving images of their companion and sending comments. Moreover, the participants used the system in an unpredicted playful way: They made drawings like cartoons, expressed their creativity, and challenged the other. Also, they enjoyed that the pen was not perfectly precise, resulting in childish writings that contrasted well with real images. For further enjoyment, some users suggested additional functions for future development. These included taking a photo when

they want; videoconferencing opportunities; an available keyboard to type comments on pictures; image editing tools, like Photoshop or Illustrator, or the ability to easily export and import from them; the availability of some sound or talk; and predefined and easy to tag messages, like "Congratulations!" From this feedback, it seems that playfulness is an important factor for an enjoyable engagement with this type of system (which is essential for the technology to convey productive love). Moreover, users highly valued sending and receiving handwriting in real time, and handwriting itself, in words of a participant, "led to natural communication of feelings."

The system was generally well evaluated in terms of the creation of expectations and obligations, and privacy invasion. This makes the system valuable for relationships where one of the parties has a higher desire of updated information than the other. For example, grandparents may want to see more of the younger generation, who are sometimes too busy to keep informing them.

The system was well evaluated also in terms of learnability and satisfaction, which indicates good qualities of the system in spite of having stopped several times, which caused low evaluation in terms of efficiency. The comments of the users reflected that they were satisfied principally about the ease of communication. In the words of one user, "The best is to have it all in one at hand: Receive the picture, play with it and send it."

Overall, although the productive love qualities of the users seemed to improve just minimally during the testing periods of 2 weeks, the users experienced the system as successful to promote the productive love objectives. Moreover, the system did not bring unwanted burdens due to communication and had acceptable usability. These results support the idea that an automatic picture exchange can be enjoyed by users and has the potential to support productive love.

DISCUSSION AND CONCLUSION

This is an original study that proposes the use of social psychology and philosophy for designing technology that promotes love. Although numerous philosophers, psychologists and spiritual gurus have tried to teach the nature of love, their success reached a few curious people only. We believe that in the era of ubiquitous technology, where many people feel more excited about using the latest technology than reading the latest books, there exists the possibility to make use of the technology to promote love. A possible criticism of productive love technology could be a point of view similar to Illich (1973, p. 76), who advised, "When overefficient tools are applied to facilitate man's relationship with the physical environment, they can destroy the balance between man and nature." However, Sorokin (1954) and Fromm (1956) claimed that among the forms of love similar to the one presented here, productive love, while not very common nowadays, can help better the world. Moreover, Fromm's (1956) view is that certain cultures, particularly capitalistic ones, hinder productive love.

The proposed design criteria are a first step of an approach where technology incorporates the proposed philosophy of productive love. Therefore any idea that is in accord with the productive love principles could improve the design criteria presented here. Likewise the undertaken philosophical review about love may have potential to be extended. Further, the Pictures' Call device is a first example of how the principles of productive love

can be applied and evaluation methods could be conducted. The device is not intended to be the definitive love-prompting technology but rather aims to raise criticism as well as challenge designers to consider ideas for new productive love-promoting technologies.

Furthermore, the authors foresee the possibility that the design criteria could be incorporated into other existent technologies, which may make them able to support productive love or avoid creating environments that may undermine it. Conveying the productive love principles through technologies can improve not only people's relationships and therefore their happiness, but also make the technologies that surround us more affective and thus contribute on their commercial success.

REFERENCES

Bauman, Z. (2003). *Liquid love: On the frailty of human bonds*. Cambridge, UK: Polity.

Briggle, A. (2008). Love on the Internet: A framework for understanding eros online. *Journal of Information, Communication, and Ethics in Society*, *6*(3), 216–232.

Brown, B., Taylor, A., Izadi, S., Sellen, A., Kaye, J. "J.," & Eardley, R. (2007). Locating family values: A field trial of the whereabouts clock. In J. Krumm, G. D. Abowd, A. Seneviratne, & T. Strang (Eds.), *Lecture Notes in Computer Science: Proceedings of UbiComp 2007* (Vol. 4717; pp. 354–371). Berlin, Germany: Springer-Verlag.

Burston, D. (2007). Reply to Pnina Shinebourne's essay review in Existential Analysis 17. 2 on The Art of Loving by Erich Fromm. *Existential Analysis, 18,* 198–201.

Chang, A., Resner, B., Koerner, B., Wang, X., & Ishii, H. (2001). LumiTouch: An emotional communication device. In *CHI '01 Extended Abstracts on Human Factors in Computing Systems* (pp. 313–314). New York: ACM.

Chung, H., Lee, C. H., & Selker, T. (2006). Lover's cups: Drinking interfaces as new communication channels. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems* (pp. 375–380). New York: ACM.

Dion K. K., & Dion K. L. (1996). Cultural perspectives on romantic love. *Personal Relationships, 3,* 5–17.

Eguez Guevara, J. P., & Umemuro, H. (2010).Unobtrusive estimation of psychological states through human movement observation. *e-Minds, 2*(6), 39–54.

Fromm, E. (1956). *The art of loving.* London: Thorsons.

Gelbond, B. (1979). Self-actualization and unselfish love. *Journal of Religious Humanism, 13,* 74–78.

Gibbs, M. R., Vetere, F. Bunyan, M., & Howard. S. (2005). Synchromate: A phatic technology for mediating intimacy. In *Proceedings of Designing User Experience* (pp. 2–6). New York: AIGA.

Giddens, A. (1992). *The transformation of intimacy: Sexuality, love and eroticism in modern societies.* Cambridge, UK: Polity.

Goodman, E., & Misilim, M. (2003). The sensing beds. In *Proceedings of UbiComp Workshop 2003* (pp. 23–25). New York: ACM.

Hendrick, C., & Hendrick, S. (1986). A theory and method of love. *Journal of Personality and Social Psychology, 50,* 392–402.

Hendrick, C., & Hendrick, S. (1990). A relationship-specific version of the love attitude scale. *Journal of Social Behaviour and Personality, 5,* 239–254.

Hendrick, C., & Hendrick, S. (1998). The Love Attitudes Scale: Short form. *Journal of Social and Personal Relationships, 15,* 147–159.

Hendrick, S. S., & Hendrick, C. (2006). Measuring respect in close relationships. *Journal of Social and Personal Relationships, 23,* 881–889.

Hindus, D., Mainwaring, S. D., Leduc, N., Hagström, A. E., & Bayley, O. (2001). Casablanca: Designing social communication devices for the home. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 325–332). New York: ACM.

van der Hoog, W., Keller, A. I., & Stappers, P. J. (2004). Gustbowl: Technology supporting affective communication through routine ritual interactions. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (pp. 24–29). New York: ACM.

Huijnen, A. IJsselsteijn, W., Markopoulos, P., & de Ruyter, B. (2004). Social presence and group attraction: Exploring the effects of awareness systems in the home. *Cognition Technology and Work, 6*, 41–44.

Illich, I. (1973). *Tools for conviviality.* New York: Harper & Row.

Kaye, J. "J.," & Goulding, L. (2004). Intimate objects. In *Proceedings of the 5th Conference on Designing Interactive Systems: Processes, practices, methods, and techniques* (pp. 341–344). New York: ACM.

LaFollette, H. (1996). *Personal relationships: Love, identity, and morality*, Cambridge, MA, USA: Blackwell Press.

Le, T. N. (2005). A measure of immature love. *Individual Differences Research, 3*, 72–87.

Levin, J., & Kaplan B. H. (2010). The Sorokin multidimensional inventory of love experience (SMILE): Development, validation, and religious determinants. *Review of Religious Research, 51*, 380–401.

Lee, J. A. (1973). *The colors of love: An exploration of the ways of loving*. Don Mills, Ontario, Canada: New Press.

Lee, J. A. (1988). Love styles. In R. L. Sternberg & M. L. Barnes (Eds.), *The psychology of love* (pp. 38–67). New Haven, CT, USA: Yale University Press.

Lin, L., & Huddleston-Casas, C. A. (2005). Agape love in couple relationships. *Marriage and Family Review, 37*, 29–48.

Mayeroff, M. (1972). *On caring*. New York: Harper & Row.

Mynatt, E. D., Rowan, J., Craighill, S., & Jacobs, A. (2001). Digital family portraits: Supporting peace of mind for extended family members. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 333–340). New York: ACM.

Nielsen, J. (1993). *Usability engineering*. San Francisco: Morgan Kaufmann.

Pickering J. (2009). *Being in love: Therapeutic pathways through psychological obstacles to love.* Sussex, UK: Routledge.

Romero, N., Markopoulos, P., Baren, J., Ruyter, B., Ijsselsteijn, W., & Farshchian, B. (2007). Connecting the family with awareness systems. *Personal and Ubiquitous Computing, 11*, 299–312.

Rubin, Z. (1970). Measurement of romantic love. *Journal of Personality and Social Psychology, 16*, 265–273.

Shinebourne, P. (2006). Erich Fromm's *The art of loving*: 50 years on. *Existential Analysis, 17*, 397–408.

Showers, C. J., & Limke, A. (2006). Organization of partner knowledge: Implications for liking, loving, longevity, and change. In K. D. Vohs & E. J. Finkel (Eds.), *Self and relationships: Connecting intrapersonal and interpersonal processes* (pp. 177–192). New York: Guilford Press.

Siio, I., Rowan, J., Mima, N., & Mynatt, E. (2003). Digital decor: Augmented everyday things. In *Proceedings of Graphics Interface, 2003* (pp. 159–166). Mississauga, Ontario, Canada: The Canadian Human-Computer Communications Society.

Sorakubo M., & Umemuro H. (2008). Two-nearly: Moderate communication of awareness to connect families living far apart. In *Proceedings of the 6th Conference of the International Society for Gerontechnology* (on CD-ROM). Eindhoven, the Netherlands: The International Society for Gerontechnology.

Sorokin, P. (1954). *The ways and power of love: Types, factors, and techniques of moral transformation*. Boston: Beacon Press.

Sprecher, S., & Fehr, B. (2005). Compassionate love for close others and humanity. *Journal of Social and Personal Relationships, 22*, 629–652.

Sprecher, S., & Fehr, B. (2006). Enhancement of mood and self-esteem as result of giving and receiving compassionate love. *Current Research in Social Psychology, 11*, 227–242.

Sternberg, R. J. (1997). Construct validation of a triangular love scale. *European Journal of Social Psychology,* 27, 313–335.

Strong, R., & Gaver, B. (1996). Feather, scent and shaker: Supporting simple intimacy. In *Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work* (pp. 29–30). New York: ACM Press.

Takashima K., & Umemuro H. (2008). Activating communications among family members living far apart by sharing common topics through television. In *Proceedings of the 6th Conference of the International Society for Gerontechnology* (on CD-ROM). Eindhoven, the Netherlands: The International Society for Gerontechnology.

Tsetserukou, D., Neviarouskaya, A., Prendinger, H., Kawakami, N., Ishizuka, M., & Tachi S. (2009). iFeel_IM! Emotion enhancing garment for communication in affect sensitive instant messenger, in virtual and mixed reality. In M. J. Smith & G. Salvendy (Eds.), *Lecture Notes in Computer Science: Human interface and the management of information, designing information environments, Part 1* (Vol. 5617; pp. 628–637). Berlin, Germany: Springer-Verlag.

Umemuro, H. (2009). Affective technology, affective management, towards affective society. In J. Jacko (Ed.), *Lecture Notes in Computer Science: Human-computer interaction. Ambient, ubiquitous and intelligent interaction, Part III* (Vol. 5612; pp. 683–692). Berlin, Germany: Springer-Verlag.

Ungerson, C. (2005). Care, work and feeling. *The Sociological Review,* 53, 188–203.

Vetere, F., Gibbs, M., Kjeldskov, J., Howard, S., Mueller, F., Pedell, S., Mecoles, K., & Bunyan, M. (2005). Mediating intimacy: Designing technologies to support strong-tie relationships. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 471–480). New York: ACM.

Weinstein, R. S. (2007). What heals in psychoanalysis? *Psychoanalytic Inquiry*, 27 302–309.

Weiss, K. (2006). Conclusion: The nature and interrelations of theories of love. In R. J. Sternberg & K. E. Weis (Eds.), *The new psychology of love*. (pp. 313–325). New Haven, CT, USA: Yale University Press.

Yashikida J., & Umemuro H. (2008). Close to you: Unobtrusive awareness communication to bring family living far apart closer. In *Proceedings of the 6th Conference of the International Society for Gerontechnology.* (on CD-ROM). Eindhoven, the Netherlands: The International Society for Gerontechnology.

Author's Note

All correspondence should be addressed to
Hiroyuki Umemuro
Tokyo Institute of Technology
2-12-1-W9-67 O-okayama, Meguro-ku
Tokyo, 152-8552 Japan
umemuro.h.aa@m.titech.ac.jp

Human Technology: An Interdisciplinary Journal on Humans in ICT Environments
ISSN 1795-6889
www.humantechnology.jyu.fi

HUMAN TECHNOLOGY

An Interdisciplinary Journal on Humans in ICT Environments ISSN: 1795-6889

www.humantechnology.jyu.fi Volume 6 (2), November 2010, 212–229

USER EXPRESSSIONS TRANSLATED INTO REQUIREMENTS

Birgitta Bergvall-Kåreborn
Social Informatics
Luleå University of Technology
Sweden

Anna Ståhlbröst
Social Informatics
Luleå University of Technology
Sweden

Abstract: *Grounding the development of mobile and ubiquitous services on actual needs and behaviors of users, rather than on designers' intuition, is a well-established tradition. However, gathering data about users in different contexts usually results in large amounts of data that have to be analyzed and translated into requirements. This crucial activity and its outcome are often shaped by the preconceptions of the developers or researchers. Despite this subjectivity, the translation process is seldom transparent. The aim of this paper, therefore, is to contribute to the field by presenting a process for translating user expressions into needs and later into requirements using Reiss' taxonomy of human needs. By adopting this process of translation, we were able to identify two hierarchical levels of needs: needs of a service and needs in a service. These two levels provide a transparent bridge between user expressions and system requirements*.

Keywords: *user needs, mobile services, eGovernment, user involvement, motivators*.

INTRODUCTION

As the demand for innovation increases and the number of services linked to voluntary use in mobile, ubiquitous, or pervasive contexts are increasing, it is possible to discern a growing interest among researchers to understand users' needs, preferences, and everyday behavior. One illustration of this interest is Oulasvirta (2005), who stated that, "innovation, development, and evaluation of design ideas cannot be based only on the designer's intuitions but must be grounded in users' actual needs and behaviors" (p. 60).

Adopting a needs-driven approach to product and service development also provides a number of benefits. Appreciating human needs is a valuable approach since needs last longer than any specific solution and a need can be met with many different products or services (Bergvall-Kåreborn, Holst, & Ståhlbröst, 2008; Ericson & Ståhlbröst, 2005; Kankainen & Oulasvirta, 2003; Patnaik & Becker, 1999); human needs are opportunities waiting to be explored, rather than guesses at the future; (Kankainen & Oulasvirta, 2003; Patnaik & Becker, 1999); human needs provide a roadmap for organizational development (Kankainen &

URN: NBN:fi:jyu-201011173092

Oulasvirta, 2003; Patnaik & Becker, 1999); discovering needs is beneficial for innovating new design ideas (Kankainen & Oulasvirta, 2003); uncovering needs offers product developers a different dynamic for understanding customers (Patnaik & Becker, 1999; Tiitta, 2003); and the empirical data on which needs are identified and interpreted is valuable in later stages of development, such as user interface design (Kankainen & Oulasvirta, 2003).

In order to obtain a richer understanding of various user groups and contexts, some user-centered approaches, combined with data-gathering methods inspired by anthropology and sociology, have been adopted in a number of studies (Esbjörnsson, Juhlin, & Östergren, 2004; Ha, Jung, & Oh, 2006; Holtzblatt, 2005; Kaasinen, 2003; Kankainen & Oulasvirta, 2003; Kankainen, Titta, & Rantanen, 2003; Tiitta, 2003). The gathered data in these studies are then consolidated, translated, and used to create an understanding of important requirements for the new products and services that are to be developed.

This consolidation and translation process plays a key role in systems development processes, yet it has attracted limited attention in the research literature. When it comes to guidance on how needs can be elicited, classified, and translated to systems requirements, there are few detailed guidelines given within the existing literature. Few authors use any particular taxonomy to classify and analyze needs; instead, needs are classified and analyzed from a "common sense" point of view. This can be seen in studies by, for example, Titta (2003), Kaasinen (2003), Kankainen and Oulasvirta (2003), and Oulasvirta (2004).

Kankainen and Oulasvirta (2003) focus on mobile and ubiquitous computing and they identify a number of needs among diverse user groups in the context of everyday activities that occur when people move through places occupied by other people and/or technological devices. No particular taxonomy was employed to classify the needs; instead "needs and motivations driving the behavior described in a narrative form were approached from a 'common sense' point of view" (p. 460). The categories constituted the technical solutions that the needfinding resulted in, but the translation from need to solution appeared as an opaque process. In later writings, Oulasvirta (2004) categorized the identified needs into three types: personal needs, navigational and cognitive needs, and, socially determined needs. Here, the needs also seemed to be categorized using a common sense approach and no taxonomy was presented. Further, not all of the examples given were stated in the form of needs. Examining the given examples of personal needs, some were stated as concerns, such as a concern of losing control over one's money when paying or sharing costs in public places. Others indicated a need but were not clearly communicated as such; for instance, the expression that moving around certain places triggered memories or opinions that the study subjects considered worth preserving. The same was true for the navigational and socially determined needs.

Tiitta (2003) focused on identifying everyday motivational needs concerning communication and mobility by elderly people. He used narratives cited by the participants and analyzed the narratives via contextual design methods to understand user needs. A key question in the analysis was, "Why is he or she doing or saying this?" Through an iterative process, phenomena with similar motivational needs were then clustered together in the same category. This resulted in the identification of 20 motivational needs within the following categories: group coherence, utility and experience, easy traveling, and security. As with Kankainen and Oulasvirta, the needs seemed to be categorized using a common sense approach: No taxonomy was given.

However, Oulasvirta (2005) was not unaware of these problems. Instead, he argued that the concept of user need is inflated by divergent definitions and uses. He also pointed to the weak linkage between the needfinding notion in human–computer interaction and the related discussions and typologies in modern psychology. According to Oulasvirta, this is problematic since attributing needs to users is not a straightforward process and categorizations of user data are inherently laden with the preconceptions of the researcher. Such processes, therefore, need to be based on sound scientific theories and methods. Finally, he argued that the notion of user needs is almost entirely individualistic, and, as such, does not address emergent needs that pertain to groups and organizations of users.

The above examples indicate the necessity for greater clarity regarding needs along with an understanding of how to elicit, analyze, categorize and translate needs into requirements. Hence, the aim of this paper is to present a method for translating user expressions into needs and later into requirements. The method adopts a framework based on psychological motivators and its application is illustrated through a project case study focused on increasing citizens' involvement in municipality matters. Our role in the project was to elicit citizens' needs related to communication with local authorities and translate these into system requirements.

In the next section, Reiss' framework of psychological motivators, which provides an organizing framework for the study, is explained. This is followed by a presentation of the case study, which is centered on a European Union project that aims to increase citizen involvement in municipality matters through the use of information and communication technologies (ICTs). The next section afterward details the translation process based on three illustrative user expressions. Finally, the paper concludes with a discussion and some final remarks.

RELATING USER NEEDS TO A PSYCHOLOGICAL FRAMEWORK

The concept of needs is closely related to the concept of motivation, as can be seen in Herzberg, Mausner, and Snyderman (1959), Madsen (1970), Maslow (1954), and Schein (1970). *Motive* is generally defined as something that stimulates or drives an individual to act in a certain way, and the motive is usually a need or a desire of some kind. Due to this close relation between the two concepts, they are often used nearly synonymously in the literature (see, e.g., Maslow, 1954).

Within the information systems (IS) field, the concept of needs is used in a wide variety of ways, as was illustrated in this paper's Introduction. In this article, we do not clearly separate the related concepts of needs, motives, and desires. Instead we view all of these concepts as underlying rationales that motivate people and, as such, trigger behavior and drive the requirement specification. We do, however, make a clear distinction between needs and requirements, and propose that *needs* are used in relation to humans and *requirements* are used in relation to solutions, products and services. Distinguishing clearly between needs and requirements is also supported by Sharp, Rogers, and Preece (2007).

Further, we aim to address the weaknesses related to the concept confusion identified by Oulasvirta (2005) by presenting and applying Reiss' and Havercamp (1996, 1998) psychological theory of human motivators. Their theory has been developed and continuously validated in studies since 1995, and it is, according to Reiss (2001), one of the first scientific studies that is based on what people value the most. The aim of the theory is to explain what

people experience as meaningful behavior, or what motivates them to act. Their work resulted in a framework consisting of 16 basic desires, or motives (Reiss and Havercamp 1998; Reiss, 2000). According to this theory, nearly all meaningful human behavior is motivated by some compound variation of the 16 basic desires or motives (see Table 1; Reiss, 2001, 2004).

Each of Reiss' 16 basic desires constitutes an end motive. The idea of *end motives* started with Aristotle, who divided motives into ends and means (Reiss, 2000). End motives are things people enjoy for their own sake, whereas means are methods for satisfying end motives. The number of means used to reach an end is limited only by imagination, while the number of ends is limited by human nature (Reiss, 2004).

The 16 motives are satisfied by meaningful behavior usually sought out in relationships, careers, families, sports, and spirituality (Reiss, 2001). We feel secure, for example, when we are in an environment with a degree of stability and order. We experience love when we spend time with our children and satisfy the desire for family. The satisfaction of each basic desire gives rise to a different joy, so we go through life trying to experience 16 different types of intrinsically valued feelings. Once we satisfy a basic desire, the joy soon dissipates and that specific desire reasserts itself. Therefore, we seek activities that make possible repeated satisfactions of our basic desires (Reiss, 2004, 2005).

Each basic desire is a continuum between two extremes, indicating a strong versus weak variation of that desire. Although all people are motivated by each basic desire, they are not motivated to the same extent (Reiss, 2005). The importance that people place on a desire often varies by factors such as personality type and cultural influences. Even from an individual's point of view, the importance of a desire and the form it takes on the continuum is dependent on both context and time. This implies that what motivates one person might not motivate another person, nor the same person in different contexts. Hence, human motivation is fundamentally multifaceted, as people seek to experience different intensities and frequencies

Table 1. Human Motivators, Motives, and Their Intrinsic Feelings (Drawn on Reiss, 2004).

Motivator	**Motive (the driving force)**	**Intrinsic Feeling (the feeling obtained)**
Power	Desire to influence (including leadership)	Efficacy
Curiosity	Desire for knowledge	Wonder
Independence	Desire to be autonomous	Freedom
Status	Desire for social standing (including desire for attention)	Self-importance
Social contact	Desire for peer companionship (including desire to play)	Fun
Vengeance	Desire to get even (including desire to win)	Vindication
Honor	Desire to obey a traditional moral code	Loyalty
Idealism	Desire to improve society (including altruism, justice)	Compassion
Physical exercise	Desire to exercise muscles	Vitality
Romance	Desire for sex (including courting)	Lust
Family	Desire to raise own children	Love
Order	Desire to organize (including the desire for ritual)	Stability
Eating	Desire to eat	Satiation (avoidance of hunger)
Acceptance	Desire for approval	Self-confidence
Tranquility	Desire to avoid anxiety, fear	Safe, relaxed
Saving	Desire to collect, value of frugality	Ownership

of each of the 16 desires in different situations (Reiss, 2005). From this follows that these 16 motives direct almost everything humans do and they constitute the foundation upon which humans become individuals. Because the 16 basic desires are irreducible (Reiss, 2005) they cannot be organized further into supercategories. Nor can they be organized as a hierarchy, where one desire builds upon another.

Motives are also the reasons why people are willing to do things on a voluntary basis (Reiss, 2004). This becomes important in the context of IT development, since knowing what motivates people to use a product or service becomes central, especially when the product or service is aimed at private and voluntary use. In our study, we have used Reiss' motivators as a tool for eliciting and analyzing the data gathered in the focus group interviews within the SMART project.

THE SMART PROJECT

Many eGovernment projects are characterized by a technocentric approach with minimal involvement from citizens (Olphert & Damodaran, 2007). The SMART (2006–2007) project was one of many EU projects aimed at increasing citizen involvement in municipality matters through the use of ICTs. This took place in SMART by exploring the concept of "reaction media," which allowed individuals to actively participate in the development of their municipality. More specifically, the project aimed to develop mobile and context-aware services that facilitated the communication between citizens and the municipality. This service was intended to stimulate citizens to give suggestions and opinions for how they wanted the municipality to develop, and to identify risks or dangers in their environment.

The development of these services was carried out in an interactive manner in cooperation between citizens, companies, and official authorities. Since the SMART project was an applied research project aimed to develop new technology, our role as researchers became dual, to some extent. From a practical point of view, our role in the project was to elicit citizens' needs related to communication with local authorities and translate these into system requirements. From a research point of view, our role was to reflect on how a theoretical framework could support this process. We focus on the latter process in this paper.

To facilitate a participatory approach, the project was set in a Living Lab context. The foundation of Living Labs is the involvement of four different stakeholders (government, companies, researchers, and end-user representatives) in innovation processes. The aims involve close cooperation among involved stakeholders in facilitating innovation, and developing products and services that users really need and that are designed to fit their life patterns and preferences. During this development process, the products and services are also tested by end users in their real-world environments. Since the Living Lab activities can go on 24/7, users can test a product or service in their private, real usage situations and from the perspectives of the various roles they shift between during a day: citizen, parent, sport fan, patient, student, or employee. Hence the users gain a deeper and more practical understanding of the function of a new product or service and how it fits into their usage contexts (Mulder et al., 2007; Ståhlbröst, 2006). With this approach, the innovation system becomes human-centric, in contrast to technology-centric.

The Living Lab milieu used in our study is an open Living Lab, called Botnia Living Lab, in which anyone with an interest in the development of IT-related products and services can cooperate. The basic idea of Botnia is to engage interested individuals, ultimate end users, and stakeholder organizations in an interactive and iterative process from needs identification and idea generation through concept development and prototype testing to market validation. Because Botnia is open to all kinds of IT stakeholders in the value chain, the process aims to help these stakeholders manage their development processes with a user-centered approach.

Over the years, Botnia has built up a community of end users with whom it easily and frequently communicates. Thus, this community comprises approximately 7,000 accessible and volunteer "test pilots." Generally, the test pilots represent the diversity found among citizens when it comes to age, gender, living conditions, education, carriers, and so on. These volunteers, however, are motivated by and share a common interest in innovation and technology, are curious to try out new technical artifacts, and seek the opportunity to influence them (Ståhlbröst, 2008).

RESEARCH APPROACH

In order to illuminate the issues discussed above, a qualitative approach was deemed most appropriate, since it enables researchers to capture the richness and the detail of the citizens' experiences. While qualitative research methods were used, the epistemological underpinnings were more broadly interpretive (Orlikowski & Barudi, 1991; Walsham, 1995), focusing on social constructions such as language and shared meanings. Interpretive studies generally attempt to understand phenomena through the meanings that people assign to them through human sensemaking activities. Given that the focus of the research project is to increase citizen involvement in local government, it is crucial that we gain an understanding of the potential users and their needs within the context in which they operate. In order to help achieve this, we used scenario-based focus group interviews, since this method helps stimulate the creation of interactive communication among newly constituted groups that share characteristics of interest (Bloor, Frankland, Thomas, & Robson, 2001; Wibeck, 2000). Focus groups can generate a broader scale of ideas and views compared to traditional one-on-one interviews (Wibeck, 2000), and create a situation in which participants can validate and discuss each other's perspectives and experiences. These interviews aimed to gather the opinion and discussion of the group, not to follow individuals' statements.

In this study, we planned and carried out five focus groups with Living Lab community participants. Arguably, this may skew the results since the community consists of self-selected participants who are interested in technology development. However, we were able to approach a cross section of the Living Lab participants, allowing to emerge a heterogeneous group covering a diversity of citizens, based on residential location, gender, age, and occupation (see Table 2). Initially, around 200 people were approached to participate in the project and—as is typical—approximately 10% responded positively, resulting in 23 participants (13 male and 10 female respondents, with an age varying between 18 and 50 years). The participants were divided into five focus groups, each consisting of between four and six respondents. To account for geographic variation and experiences in different municipalities, three of the surroundings

Table 2. Focus Groups and Their Composition.

Focus group interviews	Composition of the focus groups
(FG1) Carried out with six people from the larger city area	Female; age 38; single; student but works part time in theatre and dance; lives in the city Female; age 30; single; student; lives in the suburb Male; age 52; family; works as an engineer; lives in a rural area Female; age 48;family; works as a nursing assistant; lives in the city Female; age 24; partner; student; lives in the city
(FG2) Carried out with four people from the larger city area	Male; age 35; single; student; unemployed; lives in the suburb Female; age 28; partner; student; lives in the suburb Female; age 22; single; student; lives in the suburb Male; age 36; family; works as a journalist; lives in the suburb
(FG3) Carried out with four people from the larger city area	Female; age 41;family; unemployed; lives in a rural area Male; age 24; single; student; lives in the city Female; age 32; partner; works in an office; lives in the suburb Male; age 26; single, student; lives in the suburb
(FG4) Carried out with four people from one of the smaller city areas	Female; age 41;family; works as a university administrator; lives in the suburb Male; age 38; family; works as a sales engineer; lives in the city Male; age 25; partner; student; lives in the city Male; age 40; family; works as a quality control manager; lives in a rural area
(FG5) Carried out with six people from one of the smaller city areas	Male; age 55; family; works as an IT-consultant; lives in the suburb Male; age 32; partner; works at a newspaper, lives in the city Male; age 37; family; works as a teacher; lives in the suburb Female; age 24; partner; works as a social worker; lives in the suburb Male; age 27; partner; student; lives in the city Male; age 34; family; works as a social worker; lives in the suburb

(about 75,000 inhabitants) and two groups (FG4 and FG5) were carried out with citizens drawn from smaller cities and their surroundings (about 25,000 inhabitants) and rural locations. The focus groups with citizens from the larger city were held on the city's university premises, while the other two took place at the city hall, a central location with easy access, in two of the smaller cities.

The focus groups were formed in order to generate discussion and gain an appreciation of citizens' opinions and experiences in relation to communication with municipalities and governments. An open-ended, structured interview format facilitated the discussions and ensured that the main topics had been covered in each of the focus groups. However, the primary aim was to encourage the citizens to talk as much as possible with each other, rather than respond to the questions posed by the researchers.

In order to encourage this, various scenarios were used (Bødker, 2000; Bødker, Kensing, & Simonsen, 2004; Carroll, 2000) as stimuli. Scenarios can be used in two different ways: to help get the discussion process started or, when asked from users, to elicit concrete experiences or intended points. In the fieldwork, both approaches were adopted and included (a) a typical day in their life, (b) their relationship with the municipality and the process of communication, with illustrations, (c) communication of alarms and potential hazards (e.g., notifying authorities regarding a hole in the road) and their experiences of this contact, (d) suggestions as to how they would like things to operate in the future and how their needs could be accommodated, and (e) how ICTs could be used to support their future needs and aid their communication with the municipality. Once the scenarios had kick-started the group

discussion, we allowed the discussion to flow naturally until that particular area of concern had been exhausted, at which point we introduced a new theme or question. These discussions, which lasted for 1-2 hours, took place in Swedish in order to allow all the respondents to speak their native language. All of the focus groups were audio recorded. We transcribed the interviews verbatim in Swedish to facilitate the analysis process and then translated the selected expressions, reported in this paper, into English.

The method of analysis was based on an ongoing iterative process of reflection and discussion among the research team to help identify concepts, themes, and issues. The "interpretive generalizations" (Walsham, 1995) that emerged from the findings are intended to be insightful and assist scholars and practitioners in deepening their understanding of the use of ICTs by everyday citizens in their encounters with local municipalities.

TRANSLATING USER EXPRESSIONS INTO REQUIREMENTS

We start this section by giving a brief overview of the main analysis and translation processes, followed by an in-depth description of the translation process with illustrative examples. The analysis process consisted of two steps. The first step was an open analysis aimed at identifying patterns and then structuring the data into overarching themes. The second step focused on clustering user expressions into subcategories within the overarching themes. After the analysis, the translation process, consisting of three steps, formally began. The first step translated user expression into general needs and motivators. The second step translated the general needs and motivators into design-oriented and operational needs, that is, needs that aim to guide the design process. By merging user expressions with general needs, we could construct user needs related to the particular system to be designed. The third step translated the design-oriented needs into system requirements.

The first step of the analysis process began with an open analysis of the transcribed focus group interviews, which were read as a whole with a focus on identifying patterns and overarching trends in the users' expressions. In this process, 15 overarching themes were elicited based on what the users had expressed. These themes were Gaining Access, Information, Feedback, Efficiency, Competence, Freedom, Mobility, Respect, Reward, Influence, Security, Interaction, Functionality, Availability, and General Issues. To guide the process of gathering users' expressions into themes, we decided that each theme should have at least three user expressions related to it to count as a theme. Expressions that were not related to any of these themes were gathered into the theme called General Issues. This criterion of a minimum number of expressions was set to make sure that a theme represented the perspective of more than one respondent. We gathered into each theme all of the expressions related to it. For instance, related to the theme Feedback were user expressions such as, "*I do not want any auto-generated feedback. If they give feedback, it needs to be personal*" (FG4, male, 25), and "*I want to see some kind of feedback from the authorities such as, 'We have received your suggestion and this is what we have done with it'; something like that is all I need*" (FG3, female, 32).

The second step of the analysis was to cluster the users' expressions into subcategories within these themes. For example, in the Gaining Access theme, we extracted three

subcategories: people, information, and contact information. Here, all the expressions related to people gaining access to information were gathered into one category, and so forth.

The first step of our translation process aimed to get a thorough and scientifically based understanding of the underpinning needs and motivators for the users' expressions related to the subcategories, hence Reiss' (2004) theoretical framework was applied. By means of this framework, the users' expressions were analyzed with the intention to understand and clarify their expressions, but also to generate new and innovative ways of interpreting them. By applying the theoretical framework, it became possible to discern if an expression was more strongly related to another theme, if new themes needed to be added, or if a theme could be taken away. In this process two new themes were identified (Alarm and Technical Solutions/Characteristics); two themes, Respect and Competence, were combined into one; and, four themes (Security, Availability, Freedom, and Mobility) were integrated into other existing themes. Due to this, the initial 15 overarching themes were modified and reduced to 12 themes: Gaining Access, Information, Feedback, Efficiency, Respect/Competence, Reward, Influence, Alarm, Interaction, Functionality, and Technical Solutions/Characteristics, and General Issues.

In the second step of the translation, we focused on constructing design-oriented user needs related to the SMART system based on the expressions and general needs (Ståhlbröst & Bergvall-Kåreborn, 2008). During this process we also discerned two different types of design-oriented needs. The first type focused on what motivated people to use the SMART service and therefore related strongly to the main purpose of the service. The most frequently occurring needs here were: Idealism, Power, Status, Acceptance, Curiosity, and Tranquility. The second type of needs described or explained how these motivational needs could be operationalized, and what motivated people once they were using the system. As such these needs were related to the specific form and function of the system and in the SMART project they focused on Saving, Order, and Independence.

In the third step of the translation, the design-oriented and operational needs were translated into systems requirements. In this work, we strived to develop a number of alternative system requirements based on each design-oriented need in order to highlight the fact that the requirements are means for fulfilling the needs and, as such, there are many possible way this can be done.

As in any interpretation and translation process, individuals will see, select, and value the same expressions more or less differently, depending on their worldview (Checkland, 1999). This is true as well for the process described in this paper. However, in order to help validate the findings, we members of the research team carried out analyses individually before comparing and discussing our results. This revealed quite similar interpretations, even though some minor differences existed. Some of these differences are illustrated in the three expressions below, where we indicated more than one possible interpretation. These differences were not seen as problems, or as a deficiency in the process; rather, we viewed them as a strength of the processes, since they enriched the expressions and elevated our understanding of them. In relation to the transformation processes, we also want to clarify that the use of a theoretical framework in order to clarify user expressions in no way hinders a feedback loop between the researchers and the users. However, before such a feedback loop can take place, the data need to be analyzed and clustered in order to identify the points that need clarification.

In the following subsections, we will illustrate our process by means of three separate user expressions that were translated into requirements. These expressions were selected

because they illustrate some of the diversity of interpretation found in the users' stories, varying among general needs, design-oriented needs, processes, activities, problems, situations, and solutions. What is important to keep in mind here, however, is the specific aim of this paper: We do not intend to present representative and general users' needs related to the government derived from the data, but rather we focus solely on the translation process from identified user expressions into system requirements. The first expression illustrates a general need, the second a condition for use, and the third a preferred workflow.

Expression 1

An emerging discussion theme among the citizens in the focus groups concerned their desire to feel involved and have influence and control over their lives; the users' expressions related to these discussions rendered the theme Influence. The illustration of this theme stems from a discussion within FG2 that centered on how citizens viewed the opportunity for active involvement by communicating suggestions and alarms to authorities. One citizen commented, "*You need to feel that you are* involved *and have the* power *of your own life. That is important*" (FG2, male, 36).

This comment represents a clearly defined need statement expressing possible motives or underlying rationales that would stimulate and enthuse this citizen to use the service. Focusing on the key word *power* in the sentence, we at first related this to the need for power. However, in analyzing the text against Reiss' (2004) framework, we quickly realized that since the citizen talked about power related to the possibility of having influence over his own life, rather than having influence over other people and their lives, the end motive was independence, not power. This made us shift our classification of the text from the need of power to the need for independence, which included the desire to be autonomous and the intrinsic feeling of freedom.

Besides these basic human needs, the users also expressed needs that were more design-oriented and operational. These expressions took the forms of freedom to select communication channels, media, and structures when expressing a suggestion or alarm; freedom to judge the importance of the suggestion or alarm; and freedom to communicate a suggestion or alarm independent of place and time. However, on a more fundamental level, these needs do not address the basic need of independence and autonomy related to the citizens' own lives. Rather, they are limited to creating a feeling of freedom when using the system.

These design-oriented needs where then translated into requirements. Some of these requirements state that (a) the system should be able to receive data from the phone, mobile technologies, and the Internet; (b) the service should allow the user the opportunity to indicate the importance of her/his input; and (c) the response should be a mobile solution.

The meaning of the user's statement "you need to feel that you are involved" was harder to interpret since we easily could relate it to many different needs. Some of these were status and the desire for social standing and attention, social contact and a desire for peer companionship, and idealism and the desire to improve society.

When we analyzed this first part of the sentence from the perspective of these different needs, we realized that different interpretations lead to quite different design-oriented needs and requirements and sometimes even different functions within the service. Interpreting the expression as a need for status resulted in design-oriented needs, such as making people and

their contribution visible to the community, translated into the requirement of displaying the name of the person giving the suggestion.

Interpreting it as a need for idealism brought forth the design-oriented need of displaying information about whether or not citizen suggestions resulted in actual implementation of the suggestions or, if not, the reasons for this. This translated into the requirement of a service where citizens could follow their suggestions through the decision process of the governmental organization; from initial suggestion to final decision and/or implementation of the suggestion. Table 3 summarizes the process of translating the user expression "You need to feel that you are involved and have the power of your own life. That is important" (FG2, male, 36) into system requirements.

Table 3. Clarification of the Translation Process from General Needs to Requirements.

User Expression	General User Need/Motivator	Design-oriented needs	System Requirement
"You need to feel that you are *involved* and have the *power* of your own life. That is important."	Independence	Create a feeling of freedom when using the system	Support mobile solutions Handle data from multiple communication channels Visualize the importance of the users' input
	Status	Make people visible to the community	Display contributors name Visualize the importance of the users' input
	Idealism	Provide information regarding the suggestion	Indicate and display the process for the suggestion

Expression 2

Related to the theme Efficiency, the focus group interviews revealed clear differences among the citizens in their expressed levels of required effort. While some of the users were keen to participate in and contribute to municipality issues, others were less inclined to do so; for this latter group, participation rested on minimal effort. To illustrate one perspective on efficiency, we have selected a participant's expression from a discussion about reporting suggestions for improvement. This citizen said,

> *If I am to pick up the phone and give a suggestion for improvement in society, I cannot be hindered by whether I have money on my cash card or not. If there are no hindrances, then I would make the call.* (FG5, Male, 37)

This comment does not express a direct user need but rather a condition that the service should fulfill for this person to report suggestions for improvement to the municipality. Further, even though the condition posed by the citizen on the service is quite clear (it cannot require money on cash cards), it was hard for us to relate the expression to a need. The need that we started to elaborate with was *saving,* with its value of frugality, since one possible interpretation of the underlying rational behind the expression "money on cash card" could be an unwillingness to pay for the call.

We also saw a link between the need for saving and the expression, "*If there are no hindrances, then I would make the call*." However, the expression here was not related to monetary aspects, but rather to saving time through a smooth and easy process. In this instance, we found a lack in Reiss' (2004) framework, since the need for a smooth process—or, more generally expressed, the feeling of efficiency—is not present within the motivator saving.

In our discussions around the above expression, by shifting our focus between the context of the expression, the expression as a whole, and separate clauses and words in the expression, it became clear that just because users expressed design-oriented needs related to a service, this does not mean that they have an actual need of that service. On one hand, this indicates an efficiency demand that the service needs to fulfill for the citizen to use it. On the other hand, "no hindrances" also indicates that the citizen sees little use for the service, since he is only willing to make the efforts to communicate a suggestion when it is effortless to do so.

Based on this expression and our interpretation of its underlying rational, we suggested the requirement that the service should be free of charge, since such a requirement would increase the likelihood of a smooth and easy process as well as address the economic issue. In the end, however, the project team agreed on a Web-based interface for the service that was free of charge since designing a mobile-based service with no cost lay outside the scope of the project. In Table 4, the translation of Expression 2 into a system requirement is summarized.

Table 4. Clarification of Translation Process for a Condition for Use.

User Expression	General User Need/Motivator	Design-oriented needs	System Requirement
"If I am to pick up the phone and give a suggestion for improvement in society, I cannot be hindered by whether I have money on my cash card or not. If there are no hindrances, then I would make the call."		Assure a smooth and easy process Support ease of use	Be free of charge to the user

Expression 3

For many of the citizens in this study, their level of engagement was related to the extent of the ease of use in the communication process; many expressed the desire for interaction that was both straightforward and quick. In relation to the theme Functionality of a future solution for citizen involvement, one of the citizens remarked, "*I just want to pick up my phone, make a short video recording, add a voice message, and then just send it away*" (FG1, female, 30).

This expression also does not directly express a general need, but rather a preferred process or work flow. It also gives a clear indication of desired requirements that the system should fulfill. Based on this expression, we added the requirement that the final system should support the use of different types of data, such as video recordings, voice messages, and pictures.

Further, when we used Reiss' (2004) framework to analyze the expression stated by the citizen, we saw two possible interpretations. Firstly, from a savings perspective, a smooth process is desired since it is related to frugal use of time. Secondly, from the perspective of order, a smooth and clearly defined process is desired since it is linked to a feeling of

stability. However, in analyzing the text from both perspectives, we realized that the word *just* had key importance and was the determining factor for why we choose to set the expression as an expression for saving rather than order. Hence, in this way, one need (order) became the means for achieving another need (savings).

In traditional requirements engineering, an expression that clearly expresses a requirement is usually not elaborated further. Hence, the above expression would result in requirements that support mobile video recording and voice message. Relating the expression to the framework, however, made us reflect further on its meaning and this resulted in design-oriented needs related to time efficiency through ease of use, which in turn were translated to system requirements. For example, the log-on procedure became focused on creating as few steps as possible and, on an overarching level, the focus on saving made us prioritize ease of use and speed over aesthetics in the interface. As in the two previous expressions, a summary is given of the translation from user expression to system requirement, see Table 5.

Table 5. Clarification of the Translations Process for a Preferred Workflow.

User Expression	General User Need/Motivator	Design-oriented needs	System Requirement
"I just want to pick up my phone, make a short video recording, add a voice message and then just send it away".	Saving (ends)	Support ease of use	Have easy log-on procedure
	Order (means)	Assure a smooth and well defined process	Handle different input data automatically

DISCUSSION: DIFFERENT CLASSIFICATIONS OF USER NEEDS IN RELATION TO DESIGN

When we translated users' expressions into needs and then into requirements, we noticed that user needs could be associated with two distinct levels of abstraction. The first level described or explained what might motivate a user to buy and/or use a product or service, while the second level described or explained how these motivational needs could be operationalized. The motivational needs for a service were often expressed quite generally and could be related almost directly to one of Reiss' general motivators (see Expression 1). The operational needs, on the other hand, were seldom expressed as needs but rather as goals, conditions, activities, and processes. Here, the importance and influence of the interpretation and translation process became clear. To be of value for the development process and the programmers, these expressions (see Expression 2 for illustration) needed to be translated into design-oriented needs. For example, the need for a communication channel between the citizen and the receiving organization must be easily accessed and require as few preconditions from the citizens as possible.

While Reiss (2004) helped us avoid a translation and categorization process driven solely or primarily by our preconceptions, his framework did not help us to understand the two different levels found within our data. For this we turned to authors within design and

development disciplines who have translated the basic and general human needs from psychology into more prescriptive definitions.

Patnaik (2004), for example, presented four categories of individual and collective needs that originate from several design and strategy investigations, interviews, and observations: qualifier needs, activity needs, context needs, and common needs. Qualifier needs result from problems with existing systems and solutions and exist for all of those who use a particular solution in a similar way. Activity needs are the results of specific activities that persons perform or want to perform. These needs are the same for all people who want to perform the same activity. Context needs reflect the situations where people live, work, or operate: Similar needs will exist for people operating within the same industry, profession, region, culture, and so on. The common needs are the most fundamental and universal needs of all and relate to needs such as socialization, feeling loved, being comfortable, and so on.

According to Patnaik (2004), developers usually end up with a list of mostly qualifier needs, useful in that such a list leads to incremental improvements of current solutions, but not radical innovations. Context needs and common needs are often left unexplored, leading to a loss of opportunity for companies to create valuable, profitable, and strategically powerful solutions for their customers. For each category, Patnaik also suggested various solution types, for example, a qualifier need is related to new features of an existing solution while context needs are related to new families of offerings.

The categorization by Patnaik (2004) is more prescriptive than those of Maslow (1954) and Reiss (2004), and provided a mental model for understanding needs in design situations. More precisely, it helps designers understand the origin of needs and design wisely, as well as to relate needs to different aspects of a design. However, the framework also has some weaknesses. The most significant are the lack of theoretical underpinning of this work and the lack of methodological description of the research process. Also, Patnaik's categories are somewhat inconsistently described and need to be further developed.

A second example of categorizations of needs, developed from a design perspective, is Kankainen's (2003) two types of needs: motivation-level needs and action-level needs. Motivational-level needs answer the question of *why* a person is doing what s/he is doing. The action-level needs are more cognitive and relate to behavior, since they are related to a mental model of *how* to conduct an action. Hence, the motivational-level needs explain the underlying rational for carrying out an activity while the action-level needs explain the nature and process of the activity. In our view this categorization misses an important category, namely the category of *what* (i.e., what a person is doing).

Oulasvirta (2005) extended the categories presented by Kankainen (2003) by dividing the motivational needs into two subcategories: basic needs and quasi needs. While basic needs relate to general physiological, psychological, social needs, and so on, quasi needs are more ephemeral and situationally induced needs. They are not full-blown needs in the same sense as basic needs but they still affect how we think, feel and act. Oulasvirta illustrates quasi needs as the need or desire for an umbrella when standing in the rain, or the need for money when shopping in a store.

We see many similarities between Oulasvirta's (2005) basic motivation-level needs and Patnaik's (2004) common needs. As such, motivational needs provide a promising starting point for discovering design opportunities. The quasi need, on the other hand, is described as a solution rather than a need. For example, an umbrella does not represent a need, as is

argued by Oulasvirta's (2005). When it is raining, we feel the need to keep dry. This can be managed with different solutions, where an umbrella is one solution and a raincoat is another. Moreover, Oulasvirta's (2005) does not clarify how these various needs influenced or directed the different studies described in his paper.

To fully harvest the potential of these more design-oriented classifications of needs, these needs should be related to the design process as such. It is the design-oriented needs that offer the best potential for helping designers understand what needs to focus on in relation to different phases of development and to reflect on how needs and expressions of needs change form as the design process moves on.

Our classification of needs into two hierarchical levels, motivational needs and design-oriented needs, is one such example of how this can be accomplished. It helped us to identify and shift focus between needs that will motivated a user to buy and/or use a product or service and needs that will motivate the user to keep using the product and service. To further clarify the difference between the two levels of needs, as well as their place and function in the design process, we name the motivational needs as needs *of* the service, and the design-oriented needs as needs *in* the service.

FINAL REMARKS AND REFLECTIONS ON THE FINDINGS

Based on the trends and weaknesses found in the present literature related to user needs, this paper aimed to contribute to the field by presenting an approach for translating user expressions into needs and, later, into requirements. This process was illustrated using a case study focused on increasing citizens' involvement in municipality matters.

One of the most important benefits in using a framework for analyzing user expressions and translating them into requirements was the discussion and reflection it generated. Through these discussions and reflections, the importance of separating needs into two hierarchical levels or categories crystallized. The first is related to needs *of* the service, meaning what motivates a user to buy and/or use a product or service. Based on our interpretation, user needs of the service gave an indication of what the citizens considered important in their lives and what motivated them to interact with public authorities. The second is related to needs *in* the service, that is, when using a service, what needs are important for the users. In our study, typical needs of the service were Idealism, Power, Status, Acceptance, Curiosity and Tranquility, while typical needs in the service, those influencing the design of the implemented system most, were Saving, Order, and Independence.

The translation process from user expressions into requirements generated an interesting debate on what constitutes a need and what the meaning of an expression might be. It also illustrated that users do not always clearly express stated needs. Rather, their contributions include a mix of needs, suggestions, conditions, and problems. Due to this, the analysis and interpretation of user expressions becomes very important. Here, the framework helps to avoid a translation and categorization process driven only by the preconceptions of the researcher by providing scientifically sound theories on user needs. Documenting the translation process also provided a clear pattern of traceability between expressions, needs, and requirements.

Further, if we are to harvest the potential benefits of a needfinding approach, it is crucial that we are able to identify the needs and translate them into relevant requirements and solutions. In this process, the applied framework made it possible for us to see needs hidden in general expressions and to reformulate these accordingly. However, the weakness in this framework applies to Reiss' framework and generally to other frameworks as well: That is, in using a framework to support the analysis of needs, a risk always exists that the researcher may force a need into a predetermined box. Since this might hinder the development of new types of needs, the analyst needs to be attentive and open to this.

The analysis also confirmed that expressions of motivators are situated; that is, they are unique and arise within the situated reality in which the individual takes part—a certain context at a certain point of time. Hence, from a design perspective, the situated needs and motives give the direction or design implication, not the motivators as such. Further, the analyses demonstrated that interpreting the expressions from different motivators generated different requirements and, as such, resulted in different services. Finally, the analysis has generated interesting ideas for new possible requirements or functions of the developed product or service.

REFERENCES

Bergvall-Kåreborn, B., Holst, M., & Ståhlbröst, A. (2008). Creating a new leverage point for information systems development. In M. Avital, R. Boland, & D. Cooperrider (Eds.), *Advances in appreciative inquiry: Designing information and organizations with a positive lens* **(**pp. 75–95). New York: Elsevier.

Bloor, M., Frankland, J., Thomas, M., & Robson, K. (2001). *Focus groups in social research*. London: Sage Publications.

Bødker, K., Kensing, F., & Simonsen, J. (2004). *Participatory IT design: Designing for business and workplace Realities*. Cambridge, MA, USA: The MIT Press.

Bødker, S. (2000). Scenarios in user-centered design: Setting the stage for reflection and action. *Interacting with Computers*, *13*, 61–75.

Carroll, J. M. (2000). Introduction to the special issue on scenario-based systems development. *Interacting with Computers*, 13, 41–42.

Checkland, P. B. (1999). *Systems thinking, Systems practice: Includes a 30-year retrospective*. Chichester, UK: John Wiley & Sons.

Ericson, Å., & Ståhlbröst, A. (2005). In search of innovation: Grasping the concept of needs. *International Journal of Technology, Knowledge and Society*, *2*, 35–49.

Esbjörnsson, M., Juhlin, O., .& Östergren, M. (2004). Traffic encounters and Hocman: Associating motorcycle ethnography with design. *Personal and Ubiquitous Computing*, *8*, 92–99.

Ha, T. S., Jung, J. H., & Oh, S. Y. (2006). Method to analyze user behavior in home environment. *Personal and Ubiquitous Computing*, *10*, 110–121.

Herzberg, F., Mausner, B. & Snyderman, B. (1959). *The motivation to work*. Chichester, UK: Wiley.

Holtzblatt, K. (2005). Customer-centered design for mobile applications. *Personal and Ubiquitous Computing*, *9*, 227–237.

Kaasinen, E. (2003). User needs for location-aware mobile services. *Personal Ubiquitous Computing*, 7, 70–79.

Kankainen, A. (2003, June). UCPCD: User-centered product concept design. Paper presented at the Conference on Designing for User Experience (DUX'03), San Francisco, California, USA.

Kankainen, A., & Oulasvirta, A. (2003). Design ideas for everyday mobile and ubiquitous computing based on qualitative user data. In N. Carbonell & C. Stephanidis (Eds.), *User interfaces for all* (pp. 458–464). Berlin, Germany: Springer-Verlag.

Kankainen, A., Tiitta, S., & Rantanen, M. (2003, December). Exploring everyday needs of teenagers related to context-aware mobile services. Paper presented at the Human Factors in Telecommunications conference. Berlin, Germany.

Madsen, K. B. (1970). *Motivation: Drivkrafterna bakom våra handlingar* [Motivation: The momentum behind our actions]. Stockholm, Sweden: Wahlström & Widstrand.

Maslow, A. H. (1954). *Motivation and personality*. New York: Harper and Row.

Mulder, I., Fahy, C., Hribernik, K. A., Velthausz, D., Feurstein, K., Garcia, M., Schaffers, H., Mirijamdotter, A. & Ståhlbröst, A. (2007, October). Towards harmonized methods and tools for Living Labs. Paper presented at the eChallenge Conference, Hauge, The Netherlands.

Olphert, W., & Damodaran, L. (2007). Citizen participation and engagement in the design of e-Government services: The missing link in effective ICT design and delivery. *Journal of the Association of Information Systems, 8*(9), 491–507.

Orlikowski, W. J., & Barudi, J. (1991). Studying information technology in organizations: Research approaches and assumptions. *Information Systems Research*, *2*, 1–28.

Oulasvirta, A. (2004). Finding meaningful uses for context-aware technologies: The humanistic research strategy. In E. Dykstra-Erickson & M. Tscheligi (Eds.), *Proceedings of ACM CHI 2004 Conference on Human Factors in Computing Systems* (pp. 247–254). New York: ACM Press.

Oulasvirta, A. (2005). Grounding the innovation of future technologies. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*, *1*, 58–75.

Patnaik, D. (2004). System logics: Organizing your offerings to solve people's big needs. *Design Management Review*, *15*(3), 50–57.

Patnaik, D., & Becker, R. (1999). Needfinding: The why and how of uncovering people's needs. *Design Management Journal*, *10*(2), 37–43.

Reiss, S. (2000). A mindful approach to mental retardation. *Journal of Social Issues, 56,* 65–80.

Reiss, S. (2001). Secrets of happiness. *Psychology Today*, *34,* 50–56.

Reiss, S. (2004). Multifaceted nature of intrinsic motivation: The theory of 16 basic desires. *Review of General Psychology, 8*, 179–193.

Reiss, S. (2005). Human individuality and the gap between science and religion. *Zygon, 40,* 131–142.

Reiss, S., & Havercamp, S. (1996). The sensitivity theory of motivation: Implications for psychopathology. *Behavioural Research and Theory, 34,* 621–632.

Reiss, S., & Havercamp, S. (1998). Toward a comprehensive assessment of fundamental motivation: Factor structure of the Reiss profiles. *Psychological Assessment, 10,* 97–106.

Schein, E. H. (1970). *Organisational psychology*. London: Prentice Hall.

Sharp, H., Rogers, Y., & Preece, J. (2007). *Interaction design: Beyond human-computer interaction* (2nd ed.). Chichester, UK: John Wiley & Sons Ltd.

Ståhlbröst, A. (2006). Human-centric evaluation of innovation. Unpublished licentiate thesis, Department of Business Administration and Social Sciences, Luleå University of Technology, Sweden.

Ståhlbröst, A. (2008). Forming future IT: The Living Lab way of user involvement. Unpublished doctoral thesis, Department of Business Administration and Social Sciences, Luleå University of Technology, Sweden.

Ståhlbröst, A., & Bergvall-Kåreborn, B. (2008). FormIT: An approach to user involvement. In J. Schumacher & V. P. Niitamo (Eds.), *European Living Labs: A new approach for human centric regional innovation* **(**63–76). Berlin, Germany: Wissenschaftlicher Verlag.

Tiitta, S. (2003, March). Identifying elderly people's needs for communication and mobility. Paper presented at the Include Conference. London, UK.

Walsham, J. (1995). The emergence of interpretivism in IS research. *Information Systems Research, 6*, 376–394.

Wibeck, V. (2000). *Fokusgrupper: Om fokuserade gruppintervjuer som undersökningsmetod* [Focus groups: About focused group interviews as research method]. Lund, Sweden: Studentlitteratur.

Authors' Note

This work was supported by EU's Objective 1; The Swedish Agency for Innovation Systems (Vinnova); and Swedish Foundation for Strategic Research.

All correspondence should be addressed to
Birgitta Bergvall-Kåreborn
Social Informatics
Luleå University of Technology
SE-971 87 LULEÅ
SWEDEN
Birgitta.Bergvall-Kareborn@ltu.se

Human Technology: An Interdisciplinary Journal on Humans in ICT Environments
ISSN 1795-6889
www.humantechnology.jyu.fi

HUMAN TECHNOLOGY

An Interdisciplinary Journal on Humans in ICT Environments ISSN: 1795-6889

www.humantechnology.jyu.fi Volume 6 (2), November 2010, 230–249

TEMPTING TO TAG: AN EXPERIMENTAL COMPARISON OF FOUR TAGGING INPUT MECHANISMS

Mark Melenhorst
Novay
Enschede, the Netherlands

Lex van Velsen
University of Twente
Department of Technical and Professional Communication
Enschede, the Netherlands

Abstract: *Tagging helps achieve improved indexing and recommendation of resources (e.g., videos or pictures) in large data collections. In order to reap the benefits of tagging, people must be persuaded to label the resources they consume. This paper reports on a study in which four different tagging input mechanisms and their effect on users' motivation to tag were compared. The mechanisms consisted of a standard tag input box, a chatbot-like environment, a bookmarking mechanism, and a "tag and vote" game. The results of our experiment show that the use of the nonstandard tagging input mechanisms does not affect users' motivation to tag. In some instances tagging mechanisms were found to distract users from their primary task: consuming resources. Persuading people to tag might be accomplished more effectively by using other motivating tagging mechanisms (e.g., tagging games), or motivation could be created by explaining the usefulness of tagging.*

Keywords: *information retrieval, tagging, motivation, interaction design, experimental research.*

INTRODUCTION

One of the core innovations in Web 2.0 applications has been the possibility for users to share content, such as Web references, photos, videos, or presentations, with other users. When users upload their content, they can supply it with self-devised keywords. Depending on your privacy preferences, these applications allow other users to retrieve your content by means of these keywords. Asking readers to describe their content in their own words seems to be a sensible thing to do, since the inclusion of self-devised keywords, also known as *tags*, are beneficial for the indexing of the resource. Tags describe a resource (e.g., a scientific paper, photo, or video) in layman's terms or add information to the content (Berendt & Hanser, 2007).

Tagged resources are easier to find in large collections (Melenhorst & van Setten, 2007). Furthermore, the addition of tags makes it easier for other users of a system to understand the

URN: NBN:fi:jyu-201011173093

contents of the resource quickly, since tags provided by peers have been found to be more descriptive of the content than keywords provided by external professionals (Matusiak, 2006). And finally, tags reflect the interests of a user, upon which the personalization of output can be based. In this context, tags are seen as indicators of users' characteristics and contexts. An overview of how tags can be employed for user models is provided in Wartena (2010).

Of course, tagging also has its limitations (see Mathes, 2004). Tags can be ambiguous, which can make searching for a resource difficult (e.g., searching for pictures of *spoon* can result in both tableware and the rock band Spoon), and a lack of synonym control can result in a set of almost identical tags describing a resource (e.g., a picture of the Alps can be tagged as "mountain" and "mountains"). The many successful on-line services that provide the option to tag resources have made it clear, however, that the limitations do not have to stand in the way of the advantages.

A look at the popular Websites on which one can tag shows that, more often than not, platforms that facilitate the sharing of nonprofessional content do not force the users to tag, despite the advantages we just described. In the cases where the use of tags is supported, the choice of whether or not to tag is the user's. This does not have to prevent people from tagging. Several on-line services, like del.icio.us and Flickr, acquire many tags even though the choice to tag or not is left to the user. By not obligating users to tag the resources they are consuming or uploading, services such as these avoid the risk of irritating users by forcing them to devise tags, at the cost of less user-generated metadata. However, new services, or services with a small number of users, do not have such a well-developed folksonomy (a large set of tags that describes resources) to their avail, which hinders advanced indexing. In cases like these, the service will want to create a folksonomy with a small number of people or in a relatively short amount of time. One way to achieve this goal is to tempt people to tag.

Several studies have delved into the issue of how to motivate people to voluntarily tag content so as to increase the amount and diversity of metadata. The study described in this paper contributes to the discussion on how to tempt users to tag on-line content by comparing four different tagging input mechanisms. The results of this comparison can inform designers regarding the usefulness of creating and implementing tagging input mechanisms that are different from the standard input field currently used in most applications that facilitate the use of tags.

The remainder of this paper is organized as follows. In the next section, we will discuss the relevant literature. This is followed by the presentation of the experimental setup and the four tagging input mechanisms that were compared. We then discuss the results of the comparison, and finalize this paper with our conclusions.

THEORETICAL BACKGROUND

Tagging is a way to add metadata to a resource. These resources can be very diverse, ranging from videos (e.g., www.youtube.com), to songs (e.g., www.last.fm), to books (e.g., www.librarything.com). The various definitions of social tags emphasize the freedom of users in assigning the keywords they deem fit. Wu, Zubair, and Maly (2006) define tags as freestyle descriptors of a resource, while Sen et al. (2006) define them as short free-form labels used to describe items in a domain. The literature concerning tagging shows two main

streams: evolution and effect studies, and studies delving into the matter of how to motivate people to tag on a voluntary basis.

Evolution, Effect, and Motivation

Evolution studies try to map the development of a collection of tags, associated with a large information collection (e.g., Golder & Huberman, 2006). With this knowledge, it is possible to predict how large collections of tags come about and evolve. As a result, one can determine at what point in their evolution tag collections can be used effectively to improve other functionalities, such as searching. Effect studies on tagging, meanwhile, have shown that the inclusion of tags in content searched via a search engine increases the quality of search results (Melenhorst, Grootveld, & van Setten, 2008; Morrison, 2008). However, it is not correct to say that an increase in the number of tags provided will result in an increase of quality of search results. According to Chi and Mytkowicz (2008), this quality is dependent on the diversity of the tags the users provide. Therefore, an application that is aimed at tempting users to tag should focus not only on quantity, but also on diversity.

Because tagging is an activity performed most often by a small minority of a user population (Marlow, Naaman, Boyd, & Davis, 2006), another strand of research is focused on motivating all users to tag. Heckner, Heilemann, and Wolff (2009) divide the motivations for tagging into two main categories: personal information management and resource sharing. While the former is intended to efficiently and effectively store and retrieve content, the latter class of motivations is more social: the key point is to distribute valuable information with other users. A generic but more fine-grained overview of user motivations to tag can be found in Marlow et al. (2006).

Other researchers have focused on system-specific motivations to tag, assuming that generic motivations may not comply with a system-specific context. Several studies have focused, for example, on users' motivations to tag photos (Ames & Naaman, 2007; Nov, Naaman, & Ye, 2008). Interestingly, the large majority of user motivations are instrumental. In other words, people do not tag because they think it is fun, but they hope to achieve a different goal by means of tagging. For example, on MovieLens,[1] users primarily tag movies to create an overview of the kind of movies they have seen (Sen et al., 2006).

Early motivation research and literature has focused on general human behavior. The seminal work by Maslow (1943), for example, describes how human behavior is the result of needs. These needs range from the physiological (e.g., breathing or food), to love (e.g., friendship) to self-actualization (e.g., creativity). However, human motivation cannot be explained simply by a set of motivational factors, but rather is context dependent (Ryan & Deci, 2000). Situational motivation can be explained by means of the notions of intrinsic and extrinsic motivation. One is intrinsically motivated when someone does something for fun or out of interest, while one is extrinsically motivated when the executed behavior is instrumental to another goal (Deci & Ryan, 1985).

Deci & Ryan (1985) argue that intrinsic and extrinsic motivations are interrelated when rewards are used to stimulate external motivation. When rewards affect the perceived self-determination and perceived competence, they may also increase intrinsic motivation. If not, rewards potentially threaten intrinsic motivation. There is considerable evidence for the negative effect of tangible rewards (e.g., money) on intrinsic motivation (Deci, Koestner & Ryan, 1999).

Various on-line activities are initiated because of the different intrinsic or extrinsic motivations. Using the Internet in the context of work is primarily the result of extrinsic motivation: It should be useful for achieving a certain goal (Teo, Lim, & Lai, 1999). On-line shopping, on the other hand, is an activity that is fun for many of people, and hence, is intrinsically motivated for them (Shang, Chen, & Shen, 2005). Finally, the use of medical Websites has been found to be both intrinsically and extrinsically motivated (Logan et al., 2000).

Tagging Applications: Designing for Motivation

Several publications discuss research that explores different techniques for increasing users' motivation to tag various kinds of context. Drenner, Shen and Terveen (2008) conducted a study with users of MovieLens in four conditions. In one condition, users were not forced to tag. These users were in turn assigned to one of two groups: those who did and those who did not receive a screen in which they could tag on a voluntary basis. In the other two conditions, users had to tag either 5 or 25 movies. More people did not complete the more intensive condition (tag 25 movies) than the less intensive conditions. However, the new users who were forced to tag more movies were more fanatic taggers during their subsequent use of MovieLens, even if forcing the new users to tag did not influence the quality of tags. This study shows that it is possible to shape future tagging behavior of users by means of a specific interaction design.

Other researchers have focused on positively influencing tagging behavior by focusing system design upon the needs and wishes of the target user group. The mobile photo application Zonetag (Ames & Naaman, 2007), for example, gives users the possibility to skip the option to tag their photos. It can also provide users with tag recommendations. The recommendations are based upon the tags of other photos that have been taken at the same location and tags that the user has submitted in the previous 24 hours. These recommendations were found to increase the number and diversity of tags. Similar promising findings for tag recommenders have been found in the context of recommending tags for blog posts (Sood, Owsley, Hammond, & Birnbaum, 2007) and in social bookmarking systems (Jäschke, Marinho, Hotho, Schmidt-Thieme, & Stumme, 2008).

An interesting application that approaches tagging as a game is the ESP game (Von Ahn & Dabbish, 2004). In this game, two persons are shown the same photo that they tag simultaneously without seeing the other person's tags. When an identical tag is submitted by both persons, they are awarded points. Since its launch, the ESP game has proven to be a success, and collected more than 10 million tags in the first few months (Von Ahn, 2006). The success of this game has shown that presenting tagging input mechanisms as a form of entertainment has the potential of greatly improving the number of tags provided by users.

Social Communities: Interface and Interaction Design

In the literature, several implications for the design of motivating interface and interaction design for social communities can be found. Tagging is often done in the context of a social community. Therefore, tagging can be regarded as a voluntary contribution to such a community. YouTube is, for example, more than a huge database of movies, but also a place for people to gather and socialize (Lange, 2008).

A study concerning how to motivate social community members to rate movies identified that members are more willing to rate a movie when the interface shows that the community will benefit from this rating. This effect was even larger when the interface indicated that members themselves would benefit from their actions (Rashid et al., 2006). Another study showed that movie community members are more willing to invest time and effort when they are given specific goals (e.g., rate 16 movies in the next week; Beenen et al., 2004). In the case of a social peer-to-peer downloading service, it has been found that rewarding active members with an upgrade in their membership status (e.g., silver or gold members) and an improvement in the service, such as increased privileges, motivates users to be more active in the community (Cheng & Vassileva, 2005).

The works mentioned in this section have generated some useful guidelines for the design of social communities that can also be used in the context of applications that facilitate the possibility to tag (see, e.g., Preece, 2000). However, most guidelines concern general interaction design. In order to improve the design toolkit for tagging applications, we investigated how different tagging input mechanisms affect users' motivation to tag.

INPUT MECHANISMS SELECTION AND EXPERIMENT SETUP

This study is the next in a line of work that focuses on user motivation to tag videos. In this section we will first briefly discuss our previous studies. Then we will present the four different tagging input mechanisms that we compared. We will conclude this section with describing the experimental procedure we applied.

Previous Work: Video Tagging and Motivation

In the first stage of our work into user motivations to tag video content, we focused on putting together a list with users' possible motivations to tag a video on the Internet. Based on focus groups, we compiled a list with possible motivations related to indexing, socializing, and communicating (Van Velsen & Melenhorst, 2008).

Next, a large group of intensive Internet users ranked these possible motivations for two cases: uploading a video onto an on-line news Website and watching a video on an on-line music community (Van Velsen & Melenhorst, 2009). In both cases the motivations related to indexing were the main motives to tag an uploaded or watched movie. The motivation "tagging as a means to make others able to find a movie" was in both cases the most important motivation of all. Interestingly, affinity with the subject at hand did not lead to a higher motivation to tag: People tag certain video content to achieve another goal (e.g., improved indexing of a movie) not because they think a video is funny or interesting. Based on these findings, one can say that video tagging by means of a traditional tag entry box is extrinsically, rather than intrinsically, motivated.

The next step in our research was to take these insights, translate them into tagging input mechanisms, and to put these to the test. These tagging input mechanisms were the result of several brainstorming sessions.

Brainstorming Sessions

As a first step in developing the different tagging mechanisms to be compared in our experiment, two brainstorming sessions were held. The first session was held with a class of 25 third-year college students majoring in digital communication at the Hogeschool Utrecht (a university of applied science in the Netherlands). First, it was explained what tagging entailed. Next, groups of five to six students were assigned to discuss and come up with ideas for motivational tagging systems. To promote the elaboration of the ideas, the ideas from one group were passed to another group after which all of the ideas were further discussed and new ideas were generated in the group as a whole.

The second brainstorming session was held with 11 people: six experts from the fields of digital communication, cross media studies, and usability, who were teachers at the aforementioned school, and five student researchers in the field of digital communication. The process was the same as the process that was followed in the first brainstorming session.

The results of these brainstorm sessions was a long list of ideas. These ideas, listed in Table 1, represent potential means to motivate users to contribute tags.

Several ideas that served a purpose other than motivating users to contribute tags were left out. From the extensive list of ideas, three ideas were selected and further elaborated into working prototypes, hereafter referred to as tagging mechanisms. *Bookmarking* was selected because of our earlier research: We found that personal indexing was the most important motivation for users to tag (Van Velsen & Melenhorst, 2009). Therefore, it could be considered the most promising mechanism.

The *chatbot/chatbox* was selected because of its attempt to transform tagging into chatting. This is an activity in which many Internet users engage because of a social, intrinsic motivation (Stafford, Stafford, & Schkade, 2004).

The *tagging game* was selected because it appeals to users' motivation for competition and play (Marlow et al., 2006). In addition, voting for tags may improve their quality.

Descriptions of Tagging Mechanisms

The selected mechanisms were integrated into a Web environment specifically designed for this study. In this environment, the outline of the study, the experimental environment, and the concept of tagging were explained to the user. After this introduction, the user interacted with the interfaces one by one. For each mechanism, the user was asked to watch two videos, presented in a YouTube-like style. For each mechanism, help information was made available and, if necessary, the researcher could assist the participants. In Appendix A, a screen dump is displayed for each of the mechanisms.

Condition 1: Tag Box

Rationale. This mechanism does not have a specific motivational quality. It represents the way tagging is implemented in most Websites today. As such, it is the baseline against which the other mechanisms in this study are compared.

Table 1. Long List of Tagging Input Mechanisms.

Idea	Description	Intended Motivation
Bookmarking	*Tags could serve as input for a bookmarking system. By tagging certain content, they would be able to find it again more easily. The system would automatically order and display the "tagged" favorites by type of content, such as videos about pets or videos containing spoken language. As such, this idea resembles Del.icio.us.*	Future retrieval, contribution and sharing
Personal homepage	Introducing new videos on personal home pages of social network sites.	Self-presentation
Involving the social network	On the Website, one could give users the ability to create a personal friends list, or allow users to put themselves on an uploader's friends list. When the uploader shares a new video, an e-mail would be sent to his or her friends containing a link to the video in question and the request to create some tags for it, or to comment on the resource. Such a subscription method is already being used on YouTube.	Attract attention
Reward system	When users assign tags, they could be rewarded with more (related) content.	*
Commercial tagging	Following some review Websites, financial rewards could be given, for instance, based on a share of advertisement revenues	*
Chatbot and chatbox	*Users could be invited to chat about a video. When no other users are watching a video at the same time, a chatbot invites users to talk about the video. Tags can be derived from the chatlogs.*	(see section on Condition 2)
Tagging game 1	*Users could tag and subsequently vote for tags that they think are good. Votes are counted and prominently displayed. As such, users are encouraged to compete with each other to generate many high-quality tags.*	Competition and play self-presentation
Tagging game 2	Two players could simultaneously see the same image and try to come up with the same tags. If they do, they would be awarded points. The high scores of individual players would be displayed on the site prominently. Such a game is already present in the form of the ESP game.	Competition and play self-presentation
Tagging game 3	"Where is Waldo" is a game in which a little figure is hidden in the to-be-tagged resource. Multiple quick frames of Waldo could be hidden in a video. After the video is complete, the user can indicate at which frames Waldo appeared or, rather, what happened when Waldo appeared. If more people give the same answer (e.g., tags), they receive points, and their description of the scene becomes a tag.	Competition and play self-presentation
Tagging game 4	When key frames are extracted from the videos, they can be compared against Flickr photos. A game could ask users to identify the differences between the Flickr photo and the YouTube clip, from which tags can be extracted.	Competition and play self-presentation
Tagging game 5	After a video has ended, the system could present the user with a small quiz. For example, when there is an image of a cat walking across the street, a quiz question could be: "What did the cat pass on his way to the other side?" From these answers, tags could be derived.	Competition and play Self-presentation

Note. Selected ideas in italic. Motivations indicated by * give rewards in cash or in content, which are not covered by taxonomies for tagging incentives. Intended motivations are based on Chi & Mytkowicz (2008) and Marlow et al. (2006).

Functionality. This mechanism consists of the usual text box with a Tag button. No specific attempts were made to encourage users to tag here.

Condition 2: Chatbot

Rationale. Apart from the suggested mechanisms, the brainstorm sessions led to the conclusion that the propensity to tag could increase when tagging as an uninspiring activity is avoided. Earlier research (e.g., Sen et al., 2006; Van Velsen & Melenhorst, 2009) has shown that there is no intrinsic motivation for tagging, but that it is only done to achieve a certain objective. The chatbot idea does not encompass a classical tagging activity but replaces it with something that could appeal to an intrinsic motivation: to get involved with other people and friends. When chat functionality is offered next to a movie clip, it can be assumed that conversations revolve around this movie clip. Tags can be derived from the chat protocols by extracting the most salient and often-used words. Statistical techniques can be used to filter out off-topic conversations (e.g., Wartena & Brussee, 2008) and to distinguish topic related words from other salient terms (Wartena, 2010).

Functionality. Users can chat about the video in a chat window that is presented next to the video clip. When no other users are on-line, users can chat with a chatbot (an artificial-intelligence-based computer that can communicate with users more or less like a human being) that invites the users to tell him what the video clip is about. However, this was presented as an invitation in order to avoid pressuring the users to use the chatbot.

Condition 3: Bookmarking

Rationale. In a previous study (Van Velsen & Melenhorst, 2009), we found that personal indexing or indexing for others are the most important motivations for users to engage in video tagging. This prototype draws on this motivation. To a certain extent, it resembles Del.icio.us.

Functionality. Users can organize their bookmarks into folders and tag them. Subsequently, they can retrieve their bookmarks via these tags. Thus, in addition to a basic tagging mechanism, it allows users to organize their content by means of tags.

Condition 4: Tag & Vote

Rationale. This mechanism was created on the assumption that people like it when they can display their competence by being named in a high score list.

Functionality. Users can tag video clips and rate other users' tags by voting for what they think is the best tag. Tags receiving more than three votes are visible to other users. Users are able to see how many votes their tags received and what their position in a high score list is.

Experimental Set-Up

We constructed an experiment in which we evaluated the motivational effect and the appreciation of the interfaces with the implemented mechanisms that were described in the previous section.

Participants

Forty participants were informally recruited. They were, on average, 23.4 years old (*SD* = 5.0): 29 were male and 11 were female. They were all college students. However, students attending programs in digital communication, information science, and related disciplines were not allowed to participate in the study, since their prior knowledge about the topic may have interfered with the objectives of the study.

All but one of the participants use the Internet on a daily basis. Typical Web 2.0 applications are not used regularly, apart from YouTube and Hyves (a Dutch Facebook-like community). Twenty-five of the 40 participants used YouTube once a week or more, while Hyves is used once a week or more by 26 out of 40 participants. No one used Del.icio.us, one participant used Flickr once a week or more, and only six used Last.fm once a week or more.

With respect to their on-line activities, the results show that only three participants tag more than once a week, while 29 participants never tag. Sixteen out of 40 participants contribute to a forum once a week or more, while instant messaging is most popular: 28 of 40 use IM messaging more than once a week.

In sum, for this group of participants, popular social tagging applications are used only to a small extent, indicating that tagging is not so widespread among the group of participants, who may be considered as frontrunners with regard to the use of Web 2.0 applications. This result is consistent with our earlier work (Van Velsen & Melenhorst, 2009), in which we found that only 20% of the information elite knew what tagging was about.

Procedure and Tasks

The experimental procedure was completed one person at a time and consisted of the steps listed below. The entire procedure was presented within the electronic environment. The language used in the interfaces was Dutch, even though some of the movie clips were in English. Even though this environment guided the participant through the experiment, a researcher was available for questions and technical assistance.

1. *Introduction*. The experiment's steps were explained to the participant. Two things were assessed here: the participant's study subject and his/her familiarity with tagging.
2. *Reading an introduction to tagging*. Next, the core concepts and principles for tagging were explained. Each participant had to read this introduction, even if the user was already familiar with tagging: The purpose was to create a common understanding of tagging.
3. *Experimenting with the mechanisms and watching the video clips*. The participant went through all four prototypes. The order in which the prototypes were presented was randomized. For each prototype, two video clips were shown. After each video clip, a short survey was administered with questions concerning the participant's appreciation of the video and his/her propensity to tag the video clip. Following the second video in each condition, the participant was questioned additionally about the appreciation of the tagging mechanism in question and about the added value of tagging when presented this way.

4. *Survey*. The final part of the experiment consisted of a survey with questions regarding demographics and use of Web 2.0 applications.

Materials

Eight short YouTube video clips were selected and paired for each condition: Four clips were meant to entertain users and four clips were of an informative nature. They all lasted about three minutes. The titles and URL's of the videos can be found in Appendix B. The clips were presented by means of YouTube's embedded player within the ePaxperimental environment.

Data Collection

Using surveys, we collected the following data by means of short surveys:

1. Appreciation for the content, using 5-point scales and a holistic mark on a scale from one to ten (after each clip)
2. Propensity to tag (after each clip and after each mechanism): the participant's inclination to tag using the mechanism provided
3. Perceived usefulness and usability of the tagging input mechanisms (after each mechanism)
4. Background characteristics (at the end of the study)

The surveys are displayed in Appendix C.

RESULTS

In this section we discuss the results of our study. First, we address the results regarding the tagging input mechanisms. Then, we address the role of the content and its influence on the propensity to tag.

Appreciation of Tagging Input Mechanisms

After the users watched the two video clips per tagging mechanism, they were asked to provide a generic evaluation of the mechanism. We first tested whether the appreciation for the different mechanisms differed. The results are shown in Table 2. The bottom row represents the test-value of the within-subjects effect resulting from a repeated measures analysis with "tagging mechanism" as within-subjects factor.

As Table 2 shows, the scores regarding usefulness items received moderate scores. The usability items were more positively scored with means around four. Contrary to our expectations, the propensity to tag is above the neutral point of 3. We think this is somewhat surprising since the literature suggests that a small percentage of Internet users engage in tagging. Hence, we expected values to be lower than the neutral point. The added value of tagging the movie clips is considered relatively low, with a score slightly below the neutral point of 3.

Table 2. Effect of Bookmarking Mechanism on Perceived Usefulness and Usability.

Condition	Usefulness			Usability			
	Tag propensity	***Added value***	***Use in real life***	***Ease of use***	***Learna-bility***	***Compre-hension***	***Fun to use***
Control condition [C]	3.1 (1.1)	2.5 (.9)	3.0 (1.1)	4.1 (.7) ^{B}	4.2 (.6) ^{B}	4.1 (.8) $^{Ch, T, B}$	3.0 (1.2) ^{T}
Chatting [Ch]	3.5 (1.1)	2.9 (1.1)	3.2 (1.1)	4.1 (.8) ^{B}	4.2 (.8) ^{B}	3.7 (1.1)	3.5 (1.0)
Tagging & Voting [T]	3.2 (1.1)	2.6 (1.1)	3.3 (1.1)	3.8 (.9)	4.0 (.8)	3.5 (1.0)	3.6 (.9)
Bookmarking [B]	3.3 (1.1)	2.7 (1.1)	3.3 (1.1)	3.5 (1.0) ^{C,Ch}	3.6 (1.1) $^{C, Ch}$	3.2 (1.1)	3.4 (1.0)
F-Value[a]	1.34	1.98	.81	5.18 **	5.35 **	7.27 ***	3.28 *

Note. Values for the prototype-evaluations could range from 1 to 5. Standard deviations between parentheses. Significant differences between one mechanism and another are indicated by a superscript that refers to the first character(s) of the other mechanism. The significance level is .05.
[a] statistical significance: * = . at .05 level; ** at .01 level; *** at. 001 level

With regard to the perceived added-value of tagging, no statistically significant differences were found between the tagging mechanisms. The control condition (with a basic tag box) did not result in a lower perceived added value in comparison to the other tagging mechanisms.

Table 2 does show some differences in the perceived usability of the input mechanisms. The bookmarking mechanism was less easy to use and had a more troublesome learnability than the control condition and the chatbox. Not surprisingly, the control condition was the easiest to understand. The "fun to use" criterion did yield somewhat ambiguous results. Significant differences between the control condition and the tag and vote condition were found, but not between the control condition and the other conditions. This is somewhat surprising since we expected all mechanisms to be more fun to use than the control condition. In the case of the chatbot, this effect may have been caused by the absence of other users to chat with: Chatting with other users will probably be more appreciated than chatting with an automatic chatbot.

To get a better understanding of the relationship between propensity to tag and usability, we computed correlations between ease of use, learnability, instant comprehension, and the propensity to tag. In the bookmarking condition, each of the usability criteria was positively correlated with the propensity to tag ($.39 < r < .57$; $p < .05$). For the voting condition, learnability was positively correlated with the propensity to tag ($r = .37$; $p < .05$). For the chatbox condition and the control condition no correlations were found. These results suggest that usability can affect users' intention to tag.

Appreciation of Movie Content

The tagging input mechanisms cannot be considered in isolation from the content they are presented with since the content may influence users' appreciation of the mechanisms. Therefore, we investigated the relations between the content and the input mechanisms. After each video clip, the appreciation of the video clip was assessed by means of six items, derived

from Norris & Colman (1994). Participants had to award up to 5 points on each of the 6 appreciation items. Cronbach's alpha for the scale was .93. Table 3 displays the scale means.

We performed a MANOVA analysis with the tagging input mechanism as an independent variable, and average content appreciation and propensity to tag as dependent variables. Familiarity with the movie clip, presentation order of the tagging mechanisms, the type of movie clip, and the position of the subject in the experiment were introduced into the model as covariates. The model proved to be statistically significant, $F(2, 75) = 191.99$, $p < .001$.

Further inspection of the between-subjects results showed that the tagging mechanism had a statistically significant effect on the appreciation of the content, $F(3, 75) = 5.64$, $p < .01$. However, as Table 3 shows, advanced tagging mechanisms do not lead to a higher appreciation for the content than the simple tag box: The differences between the control condition and the other mechanisms were not significant.

Furthermore, the video clips were appreciated less in the bookmarking condition (Bonferroni post-hoc test; $p < .01$) and the voting condition (Bonferroni post-hoc test; $p<.01$), compared to the chatting condition, but not in comparison with the control condition. The lower appreciation for tagging & voting and bookmarking could be the result of distraction, since the items assessing usability pointed out that the participants found the bookmarking and the voting mechanism more difficult to understand than the mechanism in the control and the chatbot condition. This could have interrupted their attention to the video clips, possibly affecting their appreciation for the content. In contrast to the ratings, the propensity to tag was not affected by the tagging mechanism, $F(3, 75) = 2.50$, n.s. In other words, each of the mechanisms resulted in the same propensity to tag.

To further explore the relationship between the propensity to tag and the appreciation for the content, we computed correlations between both variables. There proved to be a significant correlation between the appreciation for the content and the propensity to tag ($r = .32$, $p < .001$). The next step was to construct a regression model with propensity to tag as the dependent variable, and appreciation for the content as the independent variable. The model proved to be significant with a R^2 of .10, $F(1, 318) = 36.04$, $p < .001$) with a highly significant B ($B = .38$; $t = 6.00$, $p < .001$). Introducing tagging mechanism as a second independent variable led to a nonsignificant Beta ($B = -.-6$, $t = -1.10$, n.s.).

Table 3. Content Appreciation and Propensity to Tag by Individual Clip Type and Condition.

Tagging mechanism	**Content appreciation**			**Propensity to tag**		
	Clip Type			***Clip Type***		
	I*	E	Avg.	I	E	Avg.
Control [C]	3.4 (1.0)	3.6 (.9)	3.5 (.9)	2.7 (1.3)	2.5 (1.3)	2.6 (1.3)
Chatting [Ch]	3.8 (.9)	3.6 (.9)	3.7 (.9)	2.7 (1.1)	2,7 (1.2)	2.7 (1.2)
Bookmarking [B]	3.3 (1.1)	3.0 (1.2)	$3.1(1.1)^{Ch}$	2.3 (1.1)	2.0 (1.2)	2.2 (1.2)
Tagging and Voting [T]	3.3 (1.0)	2.9 (.9)	$3.1(1.0)^{Ch}$	2.3 (1.3)	2.3 (1.2)	2.3 (1.2)
Average	3.4 (1.0)	3.2 (1.0)	3.3 (1.0)	2.5 (1.2)	2.4 (1.2)	2.4 (1.2)

Note. Significant differences between one mechanism and another are indicated by a superscript that refers to the first character(s) of the other mechanism.
The significance level is .05. * I stands for Informational video; E stands for entertaining video.

CONCLUSIONS

This paper presented an experiment in which we compared four different tagging input mechanisms and investigated how each mechanism affected users' motivation to tag. The four mechanisms were the result of two brainstorming sessions with students of digital communication and Web 2.0 professionals. The mechanisms consisted of a control condition (a standard tagging text input box with a tag button), a chat window in which one can chat with other users or a chatbot and from which tags are derived automatically, a del.icio.us-like bookmark mechanism and, finally, a mechanism by which one could tag and then vote for "good" tags that were kept on a high score list. The experimental results show that the different input mechanisms tested in the experiment do not lead to different perceptions of the added value of tagging, nor do they affect the users' propensity to tag. The appreciation of the content to tag was affected by the tagging input mechanism. This might have been the result of presenting a relatively complicated tagging input mechanism, which might have distracted the participants from the video content.

Our results indicate that implementing "fancy" tagging input mechanisms that utilized a chatbot, a voting mechanism for the best tag, or a bookmarking feature do not lead to a higher motivation to tag. Of course, this finding does not rule out more advanced tagging input mechanisms in general. It is possible that a different tagging design could have been perceived as prettier or more interesting by the participants which, on its turn, might have influenced usability or motivation scores. In the Evolution, Effect and Motivation section, we discussed several other promising tagging input mechanisms (like the ESP game). The implementation of these tagging input mechanisms might well be more fruitful than the mechanisms tested in this study and might motivate users to tag more than by use of a standard tag entry box. However, our opinion is that the prototypes used in this study are a proper and realistic embodiment of the ideas behind them. Therefore, although a different prototype might have marginally influenced results, we think that the general trend that can be observed in our results holds. Of course, a definitive verdict on this issue can only be realized by means of a replication study using our prototypes and new prototypes that represent the same idea.

It is possible that the tagging input mechanisms we tested did not affect user motivation because users need to be convinced of the added value of tagging in a different way. The point at which it is explained to users what tagging is and what purposes it can serve could be a crucial moment. After reading or watching this explanation, users will have to decide for themselves whether they find tagging worth the effort. Only after making this decision, then, should the users be confronted with the tagging input mechanism. The various mechanisms are used to achieve the benefits of tagging and are instruments for that, but they are not motivators. In other words, taggers do not make the decision to tag or not on the basis of the tagging input mechanism presented to them. Because tagging is done with a higher goal in mind, it is not as strongly affected by the interface and interaction design of the tagging input mechanism as we presumed it to be.

The best moment in time at which users can be persuaded to tag and in which form depends heavily on the users' contexts and goals. These contexts and goals can be identified by applying a user-centered design approach. In such an approach, (potential) users are consulted as early as possible in the design process, after which their characteristics, wishes,

and contexts lead the design (Gould & Lewis, 1985). The functional design and the interface and interaction design that are the fruit of this design approach may well be highly motivating. Therefore, case studies on user-centered design of motivating tagging interfaces will be a very welcome addition to the tagging literature.

The data provide us with some evidence that the mechanisms that were considered more difficult to understand might have disturbed the participants' appreciation for the movie clips. However, the setup of the study does not allow for a thorough analysis of the relation between these factors. Because we think that the interplay between usability and motivation to tag is an important determinant of tagging motivation, future research should delve deeper into this relationship.

However, we can safely conclude that when designing tagging mechanisms, software developers have to be careful to pay attention to usability criteria as well as their motivational quality. In any case, tagging input mechanisms should not be too intrusive, as also noted previously by Sen et al. (2006).

Interestingly, the results of the study suggest that appreciation of the content was positively co-related to a users' propensity to tag. This result is contradictory to the results we found previously (Van Velsen & Melenhorst, 2009), where higher affinity with the content to be tagged did not lead to a higher propensity to tag. We can only conclude that the relation between these two factors is unclear at the moment. Further research will have to shed more light on this relation.

IMPLICATIONS FOR PRACTITIONERS

In this study, several tagging input mechanisms were compared to test which one encouraged users to supply most tags in the context of video tagging. These involved a chatbot mechanism (chatting with a chatbot so tags can be derived from the logs), a bookmarking condition (where users can organize their bookmarks by using tags), and a tag & vote condition (where users can tag videos and rate other users' tags), as well as a traditional tagging method (tag input box) as a comparison condition. The results show that the advanced tagging input mechanisms do not improve users' motivations to tag. Therefore, designers of tagging applications have two options. The first is to use a standard tag input box where users can type tags. This is an easy and cheap solution and yielded the same results as the more advanced tagging input mechanisms we tested. The second option is to motivate users to tag by means of implementing other tagging input mechanisms. For instance, tag recommenders or tagging as a game as incorporated in an ESP game[2] have been shown to be more promising than the mechanisms we tested in this study.

Taking into account the results from our earlier studies, reported in Van Velsen and Melenhorst (2009), we can state that motivating users' to tag video content requires the careful selection of the right focus and instruments. The primary motivation to tag a video is to make a video (or other medium) easier to find for others or yourself. This activity can be simplified by using the use of tag recommenders (Melenhorst et al., 2008). An alternative approach may be to make tagging fun by means of a game. Which approach works best will be dependent on the system, the context, and the user. Even though the interplay between these dependencies is a research topic in itself, system developers need to determine the best

approach by exploring the system's context of use and identifying its target group. Based on this knowledge the most appropriate approach can be selected.

A second finding from this study is that usability is related to users' motivation to tag. Therefore, it is important to ensure that a tagging input mechanism is usable. In order to achieve this goal, we recommend applying a user-centered design perspective while creating tagging input mechanisms and testing the interface and interaction design of a mechanism before launch.

Finally, in the case of video tagging, it is very important that tagging input mechanisms do not distract the user too much from watching the video. A novel or relatively complex tagging input mechanism might prove too distracting and has implications for the tagging process and effectiveness. It is paramount to present the users' primary goal—that is, watching a video—as the main activity in the interface and interaction design of a video application.

ENDNOTES

1. See www.movielens.org
2. See www.espgame.org

REFERENCES

Ames, M., & Naaman, M. (2007, April-May). *Why we tag: Motivations for annotation in mobile and online media*. Paper presented at the Computer-Human Interaction conference (CHI 2007), San Jose, CA, USA.

Beenen, G., Ling, K., Wang, X., Chang, K., Frankowski, D., Resnick, P., & Kraut, R. E. (2004, November). *Using social psychology to motivate contributions to online communities*. Paper presented at the Computer Supported Cooperative Work conference (CSCW '04), Chicago, IL, USA.

Berendt, B., & Hanser, C. (2007, March). *Tags are not metadata, but "just more content" to some people*. Paper presented at the International Conference on Weblogs and Social Media, Boulder, CO, USA.

Cheng, R., & Vassileva, J. (2005, January). *User motivation and persuasion strategy for peer-to-peer communities*. Paper presented at the the 38th annual Hawaii International Conference on System Sciences, Big Island, HI, USA.

Chi, E. H., & Mytkowicz, T. (2008, June). *Understanding the efficiency of social tagging systems using information theory*. Paper presented at the 19th ACM Conference on Hypertext and Hypermedia, Pittsburgh, PA, USA.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*, 627–668.

Deci, E. L., & Ryan, R. M. (1985). *Motivation and self-determination in human behavior*. Heidelberg, Germany: Springer.

Drenner, S., Shen, S., & Terveen, L. (2008, October). *Crafting the initial user experience to achieve community goals*. Paper presented at the the 2008 ACM Conference on Recommender Systems, Lausanne, Switzerland.

Golder, S. A., & Huberman, B. A. (2006). Usage patterns of collaborative tagging systems. *Journal of Information Science, 32*, 198–208.

Gould, J. D., & Lewis, C. (1985). Designing for usability: Key principles and what designers think. *Communications of the ACM, 28*, 300–311.

Heckner, M., Heilemann, M., & Wolff, C. (2009, May). *Personal information management vs. resource sharing: Towards a model of information behaviour in social tagging systems.* Paper presented at the International AAAI Conference on Weblogs and Social Media (ICWSM), San Jose, CA, USA.

Jäschke, R., Marinho, L., Hotho, A., Schmidt-Thieme, L., & Stumme, G. (2008). Tag recommendations in social bookmarking systems. *AI Communications, 21*, 231–247.

Lange, P. G. (2008). Publicly private and privately public: Social networking on YouTube. *Journal of Computer-Mediated Communication, 13*, 361–380.

Logan, R. A., Nuttall, R. J., Hazelwood, S. E., Parker, J. C., Johnson, J. C., Hewett, J. E., & Reid, J.C. (2000). Audience motivations to use an arthritis website. *Arthritis Care and Research, 13*, 320–329.

Marlow, C., Naaman, M., Boyd, D., & Davis, M. (2006, August). *HT06, tagging paper, taxonomy, Flickr, academic article, to read*. Paper presented at the Seventeenth Conference on Hypertext and Hypermedia, Odense, Denmark.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review, 50*, 370–396.

Mathes, A. (2004). *Folksonomies: Cooperative classification and communication through shared metadata*. Retrieved September, 10, 2010, from http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html

Matusiak, K. K. (2006). Towards user-centered indexing in digital image collections. *OCLC Systems & Services: International Digital Library Perspectives, 22*, 283–298.

Melenhorst, M., Grootveld, M., & van Setten, M. (2008). Tag-based information retrieval of educational videos. *EBU Technical Review, 314*, 1–18.

Melenhorst, M., & van Setten, M. (2007, October). *Usefulness of tags in providing access to large information systems*. Paper presented at the IEEE Professional Communication Conference, Seattle, WA, USA.

Morrison, P. J. (2008). Tagging and searching: Search retrieval effectiveness of folksonomies on the World Wide Web. *Information Processing and Management, 44*, 1562–1579.

Norris, C. E., & Colman, A. M. (1994). Effects of entertainment and enjoyment of television programs on perception and memory of advertisements. *Social Behavior and Personality, 22*, 365–376.

Nov, O., Naaman, M., & Ye, C. (2008, April). *What drives content tagging: The case of photos on Flickr*. Paper presented at the the 26th annual SIGCHI Conference on Human Factors in Computing Systems, Florence, Italy.

Preece, J. (2000). *Online communities. Designing usability, supporting sociability*. New York: John Wiley & Sons, LTD.

Rashid, A. M., Ling, K., Tassone, R. D., Resnick, P., Kraut, R., & Riedl, J. (2006, April). *Motivating participation by displaying the value of contribution*. Paper presented at the CHI 2006, Montréal, Canada.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*, 54–67.

Sen, S., Lam, S. K., Rashid, A. M., Cosley, D., Frankowski, D., Osterhouse, J., Harper, F.M., & Riedl, J. (2006, November). *Tagging, communities, vocabulary, evolution*. Paper presented at the Conference on Computer Supported Cooperative Work, Banff, Canada.

Shang, R., Chen, Y., & Shen, L. (2005). Extrinsic versus intrinsic motivations for consumers to shop on-line. *Information & Management, 42*, 401–413.

Sood, S. C., Owsley, S. H., Hammond, K. J., & Birnbaum, L. (2007, March). *TagAssist: Automatic tag suggestion for blog posts*. Paper presented at the International conference on weblogs and social media, Boulder, CO, USA.

Stafford, T. F., Stafford, M. R., & Schkade, L. L. (2004). Determining uses and gratifications for the Internet. *Decision Sciences, 35*, 259–288.

Teo, T. S. H., Lim, V. K. G., & Lai, R. Y. C. (1999). Intrinsic and extrinsic motivation in internet usage. *Omega, 27*, 25–37.

Van Velsen, L., & Melenhorst, M. (2008, July). *User motives for tagging video content*. Paper presented at the Adaptation for the Social Web workshop, Hannover, Germany.

Van Velsen, L., & Melenhorst, M. (2009). Incorporating user motivations to design for video tagging. *Interacting with Computers, 21*, 221–232.

Von Ahn, L. (2006). Games with a purpose. *Computer, 39*, 92–94.

Von Ahn, L., & Dabbish, L. (2004, April). *Labeling images with a computer game*. Paper presented at the Conference on Human Factors in Computing Systems, Vienna, Austria.

Wartena, C. (2010). Automatic classification of social tags. In M Lalmas, J. Jose, & I. Frommholz (Eds.), *Lecture notes in computer science: Research and advanced technology for digital libraries* (Vol. 6273; pp. 176–183). Berlin, Germany: Springer-Verlag.

Wartena, C., & Brussee, R. (2008, September). *Topic detection by clustering keywords*. Paper presented at the the 19th International Conference on Database and Expert Systems Application, Turin, Italy.

Wu, H., Zubair, M., & Maly, K. (2006, August). *Harvesting social knowledge from folksonomies*. Paper presented at the the Seventeenth Conference on Hypertext and Hypermedia, Odense, Denmark.

Authors' Note

We thank the students from the Utrecht University of Applied Science who participated in the software development and data collection for this study. The research was financed through the European Commission FP7 Project MyMedia (www.mymediaproject.org) under the grant agreement No. 215006.

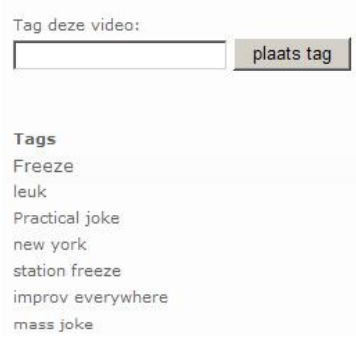
All correspondence should be addressed to
Mark Melenhorst
Novay
PO Box 589
7500 AN Enschede, the Netherlands
Mark.Melenhorst@novay.nl

Human Technology: An Interdisciplinary Journal on Humans in ICT Environments
ISSN 1795-6889
www.humantechnology.jyu.fi

APPENDIX A

Chatbot condition

Control condition



This is the control condition. In the text box, tags can be entered that appear in the tag list after pressing the *Plaats tag* button

Tag & Vote condition

Below *Uw gegevens* the user's statistics and position in the *higschore* is displayed. The green and red arrows represent positive and negative votes for tags. Below *Tag deze* video, new tags can be entered.

Bookmarking Condition

The user's collection of tags is displayed below *Mijn tags*. Folders can be found below *Mijn mappen*. In *Mijn favorieten* new video clips and tags can be entered.

A collage of screendumps of the tagging mechanisms.

APPENDIX B

Titles and URLs of YouTube Videos.

Title	URL
Frozen Grand Central	http://www.youtube.com/watch?v=jwMj3PJDxuo
Japanese way of folding t-shirts	http://www.youtube.com/watch?v=b5AWQ5aBjgE
Dove evolution	http://www.youtube.com/watch?v=iYhCn0jf46U
OFFICIAL - Terry Tate Office Linebacker "Superbowl Spot"	http://www.youtube.com/watch?v=Kg5cdZ-Fnpc
Learn Popular Magic Illusions : The Penetrating Pinky Illusion Magic Trick Explained	http://www.youtube.com/watch?v=RDGCR4W7Yn8
Fonejacker: Latest Episode: Bank Robber Vs Locksmith	http://www.youtube.com/watch?v=rr2d7YYUHEI
Bud Light Swear Jar	http://www.youtube.com/watch?v=JI3Y1auTFpU
Big Band Explained With Mince Pies	http://www.youtube.com/watch?v=WdCqtnS_cOA

APPENDIX C

After each movie clip the following questions were asked:

What do you think of this clip? Please indicate what you think of this clip on a scale from 1 to 5 on the following points.

Entertaining	1 – 2 – 3 – 4 – 5
Exciting	1 – 2 – 3 – 4 – 5
Humorous	1 – 2 – 3 – 4 – 5
Amusing	1 – 2 – 3 – 4 – 5
Nice	1 – 2 – 3 – 4 – 5
Funny	1 – 2 – 3 – 4 – 5

What mark between 1 and 10 would you assign to this clip?

Suppose you would see this clip on a video site. Please indicate to what extent you would like to provide this clip with tags.

- I would definitely not tag this clip
- I would probably not tag this clip
- I am not sure if I would tag this clip
- I would probably tag this clip
- I would definitely tag this clip

Apart from questions about the movie clip, after the second movie of each pair associated with a tagging condition, the following questions were asked about the tagging input mechanism:

Please provide a mark between 1 and 10 for this interface:

Please indicate for the statements below to which extent you agree with them:

This interface is easy to use	Strongly disagree – Disagree – Neutral – Agree – Strongly agree
It is easy to learn how to use this interface	Strongly disagree – Disagree – Neutral – Agree – Strongly agree
When I saw the interface, I could see what I could do immediately	Strongly disagree – Disagree – with it Neutral – Agree – Strongly agree
The interface is fun to use	Strongly disagree – Disagree – Neutral – Agree – Strongly agree
If this website would exist in real life, I would definitely use it	Strongly disagree – Disagree – Neutral – Agree – Strongly agree
The interface encourages to tag	Strongly disagree – Disagree – Neutral – Agree – Strongly agree
Tagging has added value for me	Strongly disagree – Disagree – Neutral – Agree – Strongly agree

HUMAN TECHNOLOGY

An Interdisciplinary Journal on Humans in ICT Environments ISSN: 1795-6889

www.humantechnology.jyu.fi Volume 6 (2), November 2010, 250–268

CAPTURING USER EXPERIENCES OF MOBILE INFORMATION TECHNOLOGY WITH THE REPERTORY GRID TECHNIQUE

Daniel Fallman
Interactive Institute Umeå and
Department of Informatics
Umeå University, School of Architecture
Sweden

John Waterworth
Department of Informatics
Umeå University
Sweden

Abstract: *We describe the application of the repertory grid technique (RGT) as a tool for capturing the user experience of technological artifacts. In noting the artificiality of assessing the emotional impact of interactive artifacts in isolation from cognitive judgments, we argue that HCI techniques must provide practical solutions regarding how to assess the holistic meaning of users' interactive experiences. RGT is a candidate for this role. This paper takes the reader step by step through setting up, conducting, and analyzing a RGT study. RGT is a technique on the border between qualitative and quantitative research, unique in that it respects the wholeness of cognition and does not separate the intellectual from the emotional aspects of the user experience. Compared to existing methods in HCI, RGT has the advantage of treating experiences holistically, while also providing a degree of quantitative precision and generalizability in their capture.*

Keywords: *user experiences, mobile HCI, repertory grid, design.*

INTRODUCTION

Adopted from the cognitive psychology of the 1970s and in force until relatively recently, the main theoretical approach to understanding human–computer interaction (HCI) was to view a person interacting with a computer generally as a disembodied information processor. Similarly, the standard methodological practice was to perform various lab-based quantitative experiments to gain empirical insight into the usability of a particular interactive device or environment, typically understood in terms of the specific qualities of the information processing involved. The nature of the users' *experiences* during interaction, that is, how he or she felt about it, was not considered or addressed.

In the last two decades, many of the limitations of this approach have been well documented within HCI by, for instance, Suchman (1987), Winograd and Flores (1986), Landauer (1991), and others. To a large segment of the HCI community, it has been clear for

URN: NBN:fi:jyu-201011173094

many years that there is more to the interaction between human users and interactive artifacts than information processing, and that methods other than tightly controlled experiments are needed if more experiential aspects of interaction are to be captured. Thus, since the early 1990s, HCI researchers have increasingly explored broader issues to gain an understanding of the relationship between the user and the artifact in terms of, for instance, affective qualities, fun, and playability. In other words, researchers and practitioners are starting to consider the user not just as a processor of information and an experimental subject, but rather as an individual with hopes, desires, expectations, and emotions.

During the period when this change of perspective in HCI was gradually taking place, psychological approaches to cognition had already moved on. After a long period in the psychological wilderness, emotion became recognized within mainstream cognitive science as a fundamental component of cognition, of our making sense of the world. As neuroscientists such as Antonio Damasio (1994, 1999) pointed out, not only are our experiences limited without emotion, but we cannot make decisions. Affect is seen as an essential component of reasoning about the world, not an opposing force. Although we may loosely speak of emotion versus reason, both too much and too little emotion will have a negative impact on cognition, with the latter being the more pathological. Understanding the nature and varieties of conscious experience is also a central topic for contemporary cognitive science. For example, huge advances have been made in identifying the neural correlates of a range of subjective states, and relating these to verbal phenomenological reports and behaviors. Experiences and behaviors are viewed as two integrated effects of the same neural events, not as separate things.

In attempts to deal with and speak about these new issues in HCI, which are far more complex than the simple human processing and associated usability views they have come to replace, the concept of *user experience* has become a key concept in recent HCI research. While there is no unified theory about the role and implication of experience to design (Forlizzi & Battarbee, 2004), a number of efforts have been made recently within HCI to establish a better understanding of the role of user experience in interactive systems design (see, e.g., Fallman, 2003, 2006; Forlizzi & Battarbee, 2004; Forlizzi & Ford, 2000; Hassenzahl & Tractinsky, 2006; Ketola & Roto, 2008; Law, Roto, Hassenzahl, Vermeeren, & Kort, 2009; McCarthy & Wright, 2004; Waterworth & Fallman, 2007).

A central issue in current user experience research is methodological: Exactly how do we best capture the experiences users have while being exposed to various designs? Purely quantitative measures, such as success rate and reaction time, do not seem to relate directly to users' experiences even though they may be useful in predicting some aspects of user performance under certain conditions. On the other hand, qualitative approaches, such as interviews and questionnaires, often lack any external validation and are limited in terms of generalizability and reliability. What is needed is a hybrid approach that provides a quantifiable and reliable measure, while also capturing subjective aspects of the experiences engendered by specific HCI designs.

In this paper, we provide an example of a candidate technique that we believe can be useful for getting insights into users' experiences of interactive artifacts in a quantitative way. We start from the position that interaction is about finding meaning, and that this involves judgments that result from a highly integrative blending of rational and affective elements, each relying on the other in producing a user's experience of an artifact. *Meaning* here refers to the sense individuals make of artifacts; we take things to mean what they are experienced

to be, reflecting the close coupling of rationality and affect. As observers of our own experiences, we cannot separate the two, except perhaps in extreme cases. In what follows, we describe and illustrate what we consider to be a promising technique for capturing the dimensions of meaning that characterize user experiences of technology in a holistic, yet also quantitative, way: the repertory grid technique (RGT).

THE REPERTORY GRID TECHNIQUE

The repertory grid technique (RGT).is a structured procedure of eliciting a repertoire of conceptual structures and for investigating and exploring them and their interrelations (Bannister & Fransella, 1985; Dalton & Dunnet, 1992; Landfield & Leitner, 1980). It has been found to be a useful technique for eliciting meaning in several different domains, for instance in organizational management, education, clinical psychology, and particularly in the development of knowledge-based systems (Boose & Gaines, 1988; Shaw, 1980; Shaw & Gaines, 1983, 1987).

RGT is a methodological extension of Kelly's (1955) personal construct theory. Kelly argued that we make sense of our world through our own construing of it. That is, we tend to model what we find in the world according to a number of personal constructs that are bipolar in nature and structure our experiences of the world. For instance, according to Kelly, we judge other people through forming personal constructs such as tall–short, light–heavy, handsome–ugly, and so on. A *construct* is essentially a single dimension of meaning for a person allowing two phenomena to be seen as similar and thereby as different from a third (Bannister & Fransella, 1985). Experiences arise from the interaction of multiple personal constructs.

What is a Repertory Grid?

While RGT is a technique for eliciting personal constructs, and a *repertory grid* is the outcome of a successful application of the technique. It is a table, a matrix, whose rows contain constructs and whose columns represent the *elements* of the phenomena under investigation. Repertory grids also typically embody a rating system used to relate each element quantitatively in relation to the qualitative constructs. An individual repertory grid table is constructed for each subject participating in a RGT study. This construction process, which will be described in detail later in this paper, is fairly straightforward. First, an individual participating in an elicitation session produces her (usage intended to be inclusive) own constructs, that is, what bipolar dimensions of meaning the person sees as the most important for talking about the elements (the investigated phenomena). The construct elicitation process is typically facilitated by the use of *triads*, through which the participant becomes exposed to sets of three elements at a time and is asked to describe and put a label on what he or she sees as separating one of the elements in the group from the other two. Second, after having provided her own individual, qualitative constructs, the participant is asked to rate the degree to which each element in the study relates to each bipolar construct according to some scale (typically a binary or Likert-type scale). Hence, in RGT, constructs and elements are the two building blocks of each individual's unique repertory grid table, and which are quantitatively related to each other by the use of some rating system. The constructs represent the qualities the

participants use to describe the elements in their own personal words (Fransella & Bannister, 1977). Constructs thus embody the participant's meaning and experience in relation to the study's elements.

RGT in Human–Computer Interaction

RGT has been found to be a useful technique for eliciting people's experiences and meaning structures in several different domains, including information systems (Tan & Hunter, 2002), education, clinical psychology, and particularly the development of knowledge-based systems (Boose & Gaines, 1988; Shaw, 1980; Shaw & Gaines, 1983, 1987). Despite its popularity in these fields, the interest in RGT from an HCI perspective peaked in the 1980s, with a special issue devoted to the topic in the *International Journal of Man-Machine Studies* in 1980. Since then, the technique's appearance in HCI-related literature has been sparse, while not completely nonexistent (see, e.g., Dillon & McNight, 1990; Grose, Forsythe, & Ratner, 1998; Hassenzahl & Wessler, 2000; Tomico, Karapanos, Levy, Mizutani, & Yamanaka, 2009). This lack of popularity may be due to fairly strong association with artificial intelligence and expert systems development in the 1980s, developments that came to epitomize the cognitivist viewpoint from which many HCI researchers were intent on distancing themselves.

Tan & Hunter (2002) recommend RGT as a means of studying the cognition of professionals and users of information systems in organizational settings, and review four examples of previous work focusing on its use for knowledge modeling. The emphasis of this kind of work is more on identifying experts' cognitive rules than on the nature of subjective experiences with technology. But recently, there has been a modest resurgence of interest in RGT as a means of capturing dimensions of user experiences with technology, as shown in research on loudspeaker array design (Berg, 2002) and subjective aspects of immersive virtual reality (Steed & McDonnell, 2003); and, more recently, to help understand cross-cultural differences in the experience of different designs of writing pen (Tomico et al., 2009).

The intention of the present paper is to further explore the potential of RGT, and to bring it to the attention of the HCI community as a possible integrative approach to understanding user experiences in HCI. This approach assumes that emotion and reason are essential and interrelated parts of making sense of the world, and provides results that are both subjective and quantitative. The following sections take the reader step by step through the setting up and carrying out of an HCI study using RGT in the context of mobile interaction devices.

USING RGT TO CAPTURE THE EXPERIENCE OF USING MOBILE INFORMATION TECHNOLOGY

In the study described below, we were interested in how people experience mobile information technology, as embodied in existing products and newly developed research prototypes. In addition to a general interest in how people relate to this kind of technology, we wanted particularly to gain empirical insight into what kinds of meanings people ascribed to the different styles of interaction these various devices embodied. The study involved existing off-the-shelf devices, as well as a number of research prototypes that represent a range of alternative means of interaction.

Participants

The empirical data collection process was carried out over a period of 3 weeks. In total, 18 participants took part in the study, all of which had previously volunteered by signing up for a scheduled time slot. Of the total number of participants, 14 (78%) were males and 4 (22%) were females. Eight of the participants (44%) were in the age span of 20–29, seven (39%) were 30–39 years of age, two (11%) were 40–49, and one (6%) was 50–59 years. As assessed by a preparatory questionnaire, three participants (16%) rated themselves as 3 on a 5-graded scale of self-estimated computer literacy, 14 (78%) rated themselves 4, while only one (6%) indicated 5. On a similar scale from 1 to 5, when asked to rate their previous exposure to mobile information technology, one participant (6%) responded with a 2, six (33%) rated themselves as 3, nine participants (50%) rated themselves 4, while two (11%) considered themselves to be 5 out of 5. As a sign of appreciation for their participation in the study, participants were provided cinema tickets. Each session lasted from 45 minutes to two hours, averaging slightly more than an hour. All participants took part in the study individually, with only the participant and the experimenter in the room. With the exception of a single native English speaker, the other 17 participants were native Swedish speakers. The study was carried out in each participant's native language and carefully translated for this paper.

Step 1: Element Familiarization

All 18 sessions began with the participant being exposed to seven different mobile information technology devices. Three of them were examples of existing devices; a Compaq iPaq H3660 personal digital assistant (PDA, known in the study as E0), a Canon Digital Ixus 300 digital camera (E1); and a Sony Ericsson T68i mobile phone (E2).

Four research prototypes were also part of the study (see Figure 1, a–d). The ABB Mobile Service Technician (E5, Figure 1a) is a wearable support tool for service technicians in vehicle manufacturing (Fallman, 2002). The Dupliance prototype (E4, Figure 1b) is a physical/virtual communication device for preschool aged children (Fallman, Andersson, & Johansson, 2001). The Slide Scroller (E3, Figure 1c) combines a PDA with an optical mouse to form a novel way of interacting with Web pages on palmtop-size displays (Fallman, Lund, & Wiberg, 2004). Finally, the Reality Helmet (E6, Figure 1d) is a wearable interactive experience that alters its user's perceptual experience (Fallman, Jalkanen, Lörstad, Waterworth, & Westling, 2003; Waterworth & Fallman, 2003).

Each session started with the seven devices being presented, one by one, to the participant. We provided brief (3–5 minutes each) introductions to the different contexts of the four research prototypes and the projects from which they originated. The participant was then able to try out each device for as long as necessary in order to become familiar with it. The session organizer was always available during the session and willing to answer any questions posed by the participants.

Step 2: Construct Elicitation

After the preparatory questionnaire had been completed, the elicitation of a participant's constructs for the seven elements (devices) began. Each participant sat at a table opposite to the

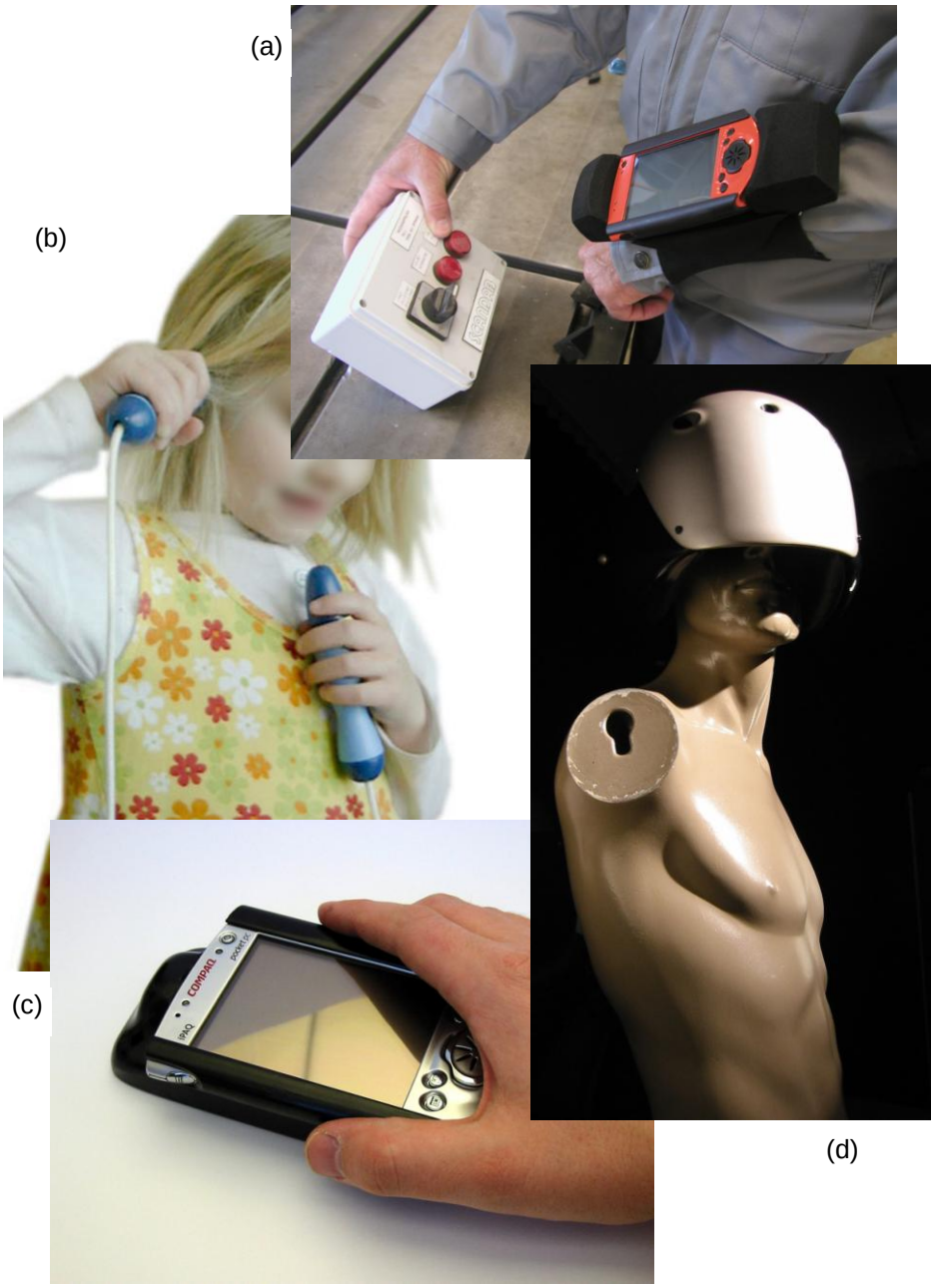


Figure 1. The four research prototypes that, together with three existing devices, were part of this study.

experimenter. On the table, seven palm-sized cards where displayed. Every card contained the following: a photograph of one of the devices; a label on which the name of the device was printed; and the identification number used for organizing the study (i.e., E0 to E6). In each session, the participant was exposed to the seven devices in groups of three; this is known as *triading* in RGT's technical language. Each triad was chosen from a list randomized prior to the study.

On a paper-based form designed especially for this study, the experimenter put down three identification numbers taken from a pre-prepared list, for instance E0, E4, and E5. The experimenter and the participant then together found the corresponding cards on the table and grouped them in front of the participant, while the remaining four cards were put aside. The participant was then asked to think of a property or quality that she considered notable enough to single out one of the three elements (devices) in the triad, and to put a name or label to that property. For instance, among a group of E1, E2, and E3, Participant 10 singled out E1, and labeled her experience as "warm." The participant was then asked to put a name or label on the property or quality that the other two devices in the triad shared in relation to the experience of E1. Participant 10 decided to collectively label E2's and E3's shared quality as "cold."

Some of the participants were fairly quick in finding what they saw as appropriate labels to put on their experiences; others could remain silent for quite some time, thinking carefully to themselves, while a few others discussed loudly and in detail their thoughts and ideas with the experimenter. Although the experimenters tried to answer questions and generally took part in discussions initiated by the participants, we were careful not to generate or imply properties or concepts, in order to avoid putting our words into the participant's mouth. To be able to keep the relation between construct and originator throughout the study, the suffix (S10) was added to each construct elicited from Participant 10. Hence, in this case, the elicited personal construct was Warm (S10)–Cold (S10).

On the form there was also a preprinted table containing the elements, each with its own 7-grade Likert-type scale. After the triading session, the form was handed over to the participant with the instruction to grade each of the seven elements according to the bipolar scale that had just been constructed from the participant's own concepts. That is, for each element of the study as a whole—including those that did not appear in the specific triad from which a particular construct pair was established—the participant was asked to rate or grade that element on a 7-point scale, 1 would represent a *high degree of the property found to be embodied in a singled out device* (e.g., in the case of Participant 10, "warm"), 7 would represent a *high degree of the property embodied by the two other devices* in the specific triad (i.e., "cold").

The Likert scale is the most widely used scale in survey research for measuring attitudes in which respondents are asked to express their strength of agreement, typically using an odd number of response options. For this study, we chose to apply a 7-grade scale, for two primary reasons. First, compared to an even-grade (a so-called forced choice scale), a scale with an odd number of choices does not force people to make choices that might not reflect their true positions). A grade 4 out of 7 thus indicates, statistically, that a construct has no particular meaning for a given element. This is important since the constructs in a repertory grid are constructed from triads in which only three out of seven elements appear. Second, because some people do not like making extreme choices (i.e., 1 or 7 out of 7), the 7-grade scale provides richer data than, for instance, 3- or 5-grade scales.

Thus, for each triad exposed to a participant, two kinds of data were collected. First, a personal construct was elicited (i.e., a one-dimensional semantic space that the participant thought meaningful and important for discussing and differentiating between the elements of a triad). This process provided the study with qualitative data: insight into the participant's own meaning structures, values, and preferences. Second, since each elicited personal bipolar construct was then used as the scale by which the participant rated all seven elements in the study using a 7-point Likert scale, data were also gathered about the degree to which

participants thought their construct had relevance to a specific element. This provided the study with quantitative data used to find out how the different elements compare and relate to each other and to the constructs, described in detail below. This analysis reveals, or at least suggests, whether or not, for example, Participant 10's construct "warm–cold" is purely literal (i.e., referring to the actual temperature of the artifact) or metaphorical (i.e., referring to the emotional effect the artifact has on the participant). The same kind of statistical analysis would not have been possible if we had asked the participants to rank rather than rate the elements.

To keep the length of the sessions roughly equal and in order not to make our participants weary, we decided to limit each session to 10 triads. Thus, from the 18 participants we elicited 180 pairs of personal constructs (i.e., 360 different concepts the participants thought meaningful and relevant) for describing their experiences of mobile information technology. At this point, it should be noted that a specific advantage of the RGT approach is that it is not necessary for the experimenter to share the specific meaning structures a participant holds in relation to an elicited construct at the time of elicitation. These are revealed during analysis by comparing the data connected with elicited constructs to data connected with other groups of elicited constructs.

ANALYSIS OF REPERTORY GRID DATA

While RGT is an open approach that results in a number of highly individual repertory grid tables, some basic structures are shared among the participants. Each table in this study consisted of a number of bipolar constructs; a fixed number of elements (7); and a shared rating system (a scale of 1 to 7). From this setup, there are at least two basic ways in which different people's repertory grid tables may be compared and analyzed interpersonally (i.e., to compare different people's repertory grids in different ways).

First, the finite number of elements and the shared rating system provide the basis for applying statistical methods that search for variations, similarities, and other kinds of patterns in the series of numbers occurring in the numerical data (the ratings). Using relational statistical methods, it becomes possible to compare and divide all constructs from all participants into groups of constructs showing some degree of similarity. This may result in interesting and unexpected correlations between constructs whose relation would most likely have remained unnoticed if one were only looking for semantic similarity. This method may hence be called *semantically blind*, since it is driven primarily by each construct pair's quantitative data in relation to elements.

Second, several seemingly semantically related and overlapping groups of construct pairs appeared across the study's participants. Some similar bipolar scales, for instance, *young–old*, *appliance–multifunctional*, and *work–leisure*, can be spotted among the responses from several of the participants. It would be possible to go through the list of all participants' constructs and gather in groups those that bear semantic resemblance to each other, and analyze these groups (e.g., using discourse analysis). This approach could be regarded as *statistically blind*, since it is driven by an interpretation of the semantic content of the constructs, not taking the numerical ratings into account.

Both of these approaches would result in a number of groups of constructs. In this particular study, we were primarily interested in finding correlations between different

constructs that may or may not seem by semantic resemblance to belong together, but which according to their ratings do. From this, it appeared that a semantically blind statistical approach that compares ratings would be the best choice for exploring the data set.

Step 3: Participant-Level Analysis

The manually collected data from the 18 participants was compiled and put into the WebGrid-III application, a frequently used and feature-rich tool for collecting, storing, analyzing, and visually representing repertory grid data (Gaines & Shaw, 1980, 1993, 1995). Each participant's repertory grid table was used as the basis for three different ways of presenting the data graphically, increasingly driven by and dependent on statistical methods of analysis.

First, a Display Matrix was generated. As the most basic way of presenting a repertory grid, this table simply lays out the numerical results of all constructs for all elements. Second, a FOCUS Graph was constructed for each participant. Here, both elements and constructs are sorted using the FOCUS algorithm (Gaines & Shaw, 1993, 1995; Hassenzahl & Wessler, 2000) so that similar ones are grouped together.

Third, the PRINCOM Map provides principal component analysis of the repertory grid data. The grid is rotated and visualized in a vector space to facilitate maximum separation of elements in two dimensions (Gaines & Shaw, 1980; Slater, 1976). For more detailed information and discussion about these common ways of analyzing and visualizing repertory grid data, see Gaines & Shaw (1993, 1995), Shaw (1980), and Shaw & Gaines (1998).

Step 4: Statistical Analysis of Multiparticipant Data

For our study, we were interesting in seeing if any patterns or other kinds of relationships between different participants' repertory grids could be derived. But how could these highly individual and subjective personal constructs be compared with each other in practice? To be able to perform statistical analysis on multiparticipant data, all 180 bipolar constructs of the participants were put into the same, very large repertory grid. This huge grid then became subject to various kinds of analyses similar to those applied to each individual participant's repertory grid. Hence, a DISPLAY matrix, a FOCUS graph, and a PRINCOM map were constructed from the WebGrid-III application using all the data. These diagrams are immense and unstructured, so the task at this point became to refine and bring order into the data set.

Statistical analysis may be performed on repertory grid data to find similarities and other kinds of patterns among the constructs elicited from different participants. Finding constructs that share a rating pattern indicate that they, mathematically, belong to the same group. This suggests that the coherence in rating also reflects coherence in experience, but one which may have been expressed differently in the semantic terms used. A group whose constructs share a unique topology in ratings thus becomes seen as a specific dimension of meaning in relation to the elements of the study. The part played by the researcher in this process is, through semantic analysis of the constructs that make up such groups, to establish what conceptual similarity they share.

Finding Groups by FOCUS Analysis of Data (1st Round)

To discover groups within the data set, the large repertory grid constructed from all the participants' individual grids was subject to two cycles of FOCUS clustering. The difference between the two rounds was in the manipulation of two rules that were applied to distinguish groups or clusters in the data.

The first rule was that the threshold level for regarding two constructs as similar was placed at 90%, that is, the constructs needed to share at least a 90% consistency in rating to be grouped together. Naturally, this rule may be discussed and questioned in a number of ways. Most obviously so, why was the 90% mark designated? In reality, this analysis effort most often needs to iterate a few times with different percentages in order to get to know the data set. Settling with 90% as a first rule of the first round was aimed at keeping a balance between (a) the number of clusters that emerge, (b) the size of these clusters, and (c) a reasonable level of internal coherence within each cluster. A higher threshold higher, say at 95%, generates clusters with a stronger degree of internal consistency, but they also become quite few in number. In addition, each cluster becomes fairly limited in terms of the number of contributing constructs. Using an overly high threshold also would leave out many of the constructs from the study and much of the study's semantic "flesh"—the place where the participants' meanings and experiences reside—would be lost. On the other hand, an overly low threshold, set at 60% or 70%, would result in almost all constructs being part of a cluster—thus embracing the lion's share of the meanings with which the participants have charged the elements—but these clusters would be very large in terms of number of constructs, and thus decreasing the clarity or definement of the element they represent. And, since each cluster would consist of a large number of constructs, a low threshold would also result in a small number of clusters in total. Thus, an overly low threshold would associate a particular construct with too many of the other constructs, where meaning would disappear in a few, large, and unmanageable clusters. Through the exploration of different threshold levels during this round, a threshold of 90% was found to be reasonable for a first statistical clustering of the constructs.

As a second rule of the first round, a *cluster* was defined as consisting of three or more constructs. When applying these two rules on the data set, 17 groups emerged consisting of 3 to 12 constructs. Each group was named with the prefix *A* followed by the group's number from top to bottom on the chart generated by the FOCUS algorithm.

Finding Groups by FOCUS Analysis of Data (2nd Round)

While the first round provided a number of statistically coherent groups, a large number of the grid's constructs had not been included. The purpose of the second round was to manipulate the rules for forming clusters so that more of the participants' constructs were included, even at the cost of lower internal coherence. This was done by lowering the threshold level to 85%, so that larger clusters developed around those established in Round 1, as well as a number of completely new clusters. To counterbalance the weaker internal coherence in rating these clusters, the second rule was made more stringent by the additional rule that clusters in this round needed to be made up of four or more constructs. Each of these groups was then named with the prefix *B* and the group's number. Twelve groups were established in this round.

Step 5: Naming Groups by Semantic Analysis

The groups identified so far may be regarded as representing the 29 most pertinent dimensions of the participants' understandings of the elements of the study. The first task of the next step was to create 29 new repertory grids based on the contributing constructs of a group. A Display Matrix, a FOCUS graph, and a PRINCOM map were also generated for each group. The analysis, up to this point, had remained statistical rather than semantic: Each of the 29 groups consisted of a number of constructs whose *ratings* grouped them together. But to be able to address a specific group as a shared bipolar concept, an interpretative analysis became necessary. Each dimension of each construct in each group was thus carefully semantically reviewed and interpreted, and one—or, if needed to better capture the character of the cluster, two or three—of the existing labels (from different participants) was chosen to characterize the group as a whole, and used to form a new bipolar construct representing the group.

At least two issues need to be highlighted in relation to this activity. First, not all constructs in a group fit perfectly well with each other semantically. Some constructs are also odd, unusual, and obviously point at something else than most others in the group. While this is not uncommon when dealing with large amounts of quantitative data, it puts the researcher in the uncomfortable position of having to make judgments about which constructs to include in a group and which to disregard in order to capture the general tendency of the group. In a few cases, no semantic resemblance and no recognized meaning structure could be established from the particular constructs of the group in question, and these groups were excluded at this stage in the procedure. In addition, some of the groups at the B-level are formed around A-level clusters, where the broadening has not always been found to provide any richer semantic information than their corresponding groups at the A-level. Thus, six B-level groups were excluded.

Additionally, even though the interpretative nature of this labeling means that the following analysis is not completely data driven, the potential hazards of experimenter biases and pure misunderstandings are reduced by choosing from existing participants' labels to capture the character of a group, rather than creating new ones. As an example of how this labeling was carried out, the group A16 consists of three contributing constructs, with "Cosmetical (S18)," "Consumer product (S14)," and "Device (S1)" on one end and "Mechanical (S18)," "Professional product (S14)," and "Tool (S1)" on the other. Here, *Device* (A16) was chosen to represent the former and *Professional tool* (A16) to represent the latter end.

Step 6: Calculating Mean and Median Ratings

If these groups, with their labels as representatives, are treated as constructs, it is possible to form a new repertory grid consisting of these 23 groups/constructs and the original elements. But to be able to statistically analyze how they relate to each other and to the elements of the study, a rating for each construct on each element needs to be incorporated into the new repertory grid table. Rather than using the arithmetic mean, these calculations relied on the median value. This was found to provide a result that seems more true to the rating of the participants, one in which the influence of single, extreme values at odds with the majority of the values in the group was de-emphasized. For each value, a standard deviation was calculated, providing clues to which values in a group are the most uncertain. Comparing the

standard deviations for the ratings across the elements of a group, as well as the value for the average absolute deviation from median, tells us something about how certain a specific rating is and provides clues to the lack of agreement by the participants on specific elements.

Step 7: Interpreting and Presenting the Result

When applying an 85% threshold to these 23 clusters and their ratings, the FOCUS algorithm further partitioned them into three groups of four or more constructs, as well as a single clustering between two additional constructs. These clusters may again be treated as groups, and hence, given these four new constructs formed from these clusters together with the remaining six non-clustered constructs, the statistical analysis leaves us with not 23 but rather 10 unique dimensions of the way in which the participants have experienced the devices of mobile information technology that were part of the study.

These 10 dimensions are presented as a FOCUS graph (Figure 2) and as a PRINCOM map (Figure 3), which also shows how the different elements relate to each other. The FOCUS graph sorts the grid for proximity between similar elements and similar constructs while the PRINCOM map makes use of principal component analysis to represent the grid in minimum dimensions (Shaw & Gaines, 1995). These 10 dimensions are thus the most significant ways in which the participants experienced the elements of the study.

The results give us a graphic account of how participants construed the seven devices and, in particular, how their experience of each related to that of the others. We must be cautious of using the construct labels literally, but it is clear that the Reality Helmet, as an example, is semantically distant from the digital camera, as shown by their opposing positions

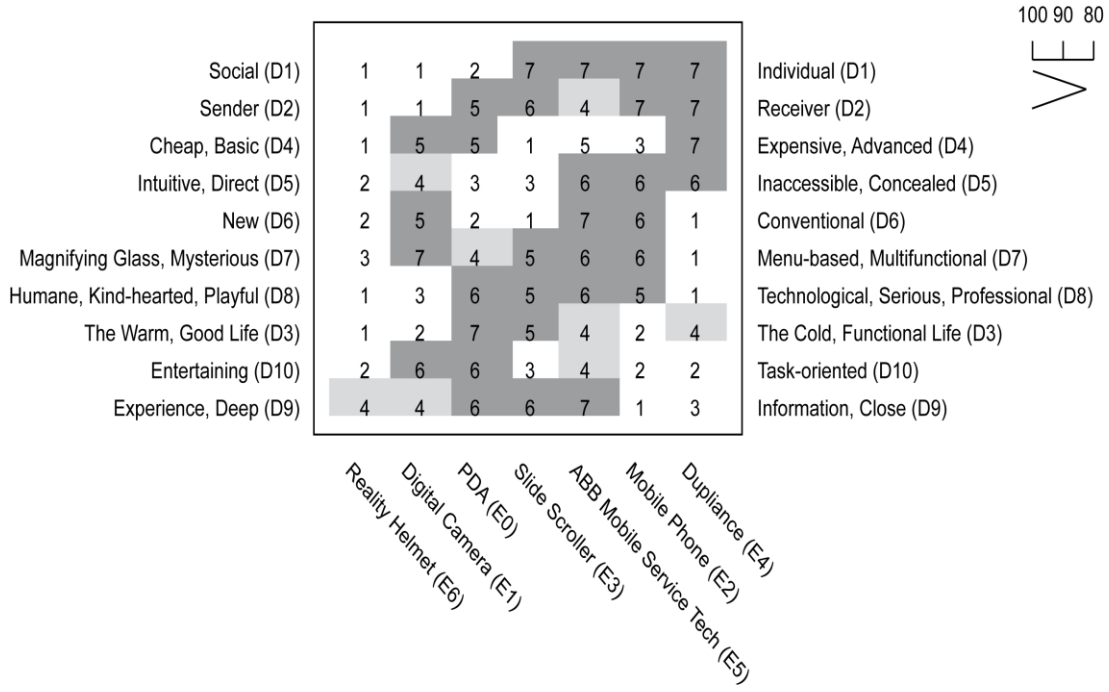


Figure 2. The resulting 10 unique dimensions (D) of the study presented as a FOCUS graph.

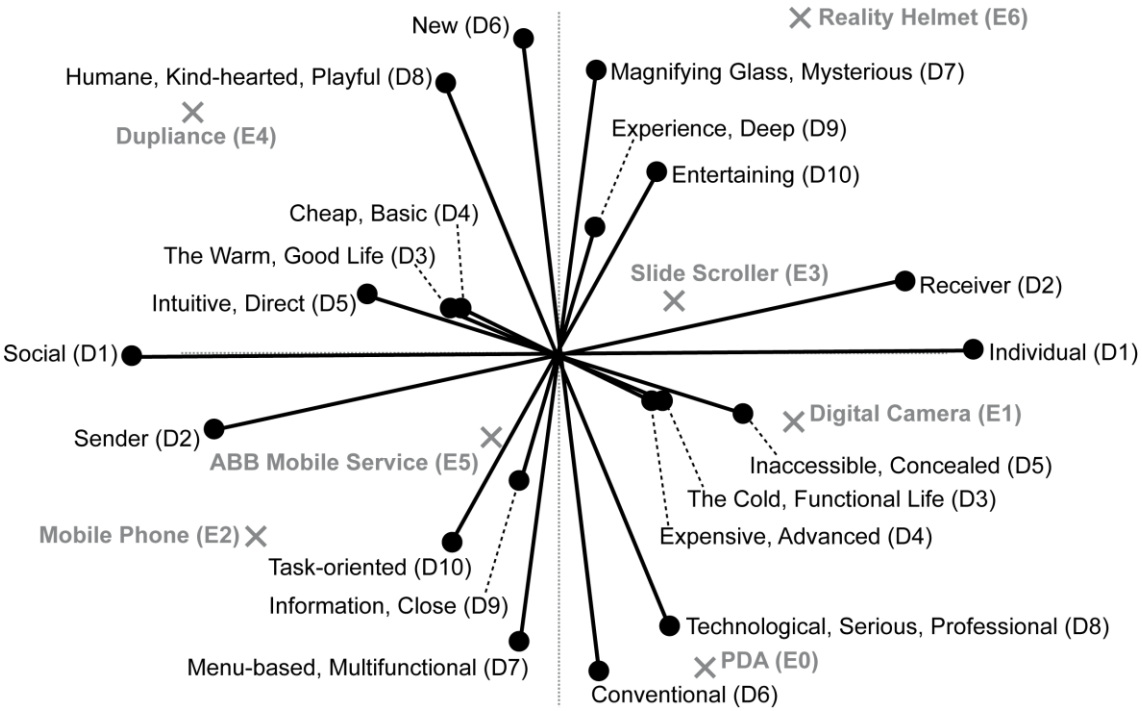


Figure 3. The 10 dimensions presented as a PRINCOM map.

on dimensions such as task-oriented (Digital Camera) versus entertaining (Reality Helmet). Several of the devices were experienced as relatively "social" (Dupliance, Mobile Service Technician, Mobile Phone) as compared to others that were more "individual" (Reality Helmet, Slide Scroller, Digital Camera, PDA). The Dupliance was associated with positive attributes such as "humane," "warm," and "intuitive," whereas the Digital Camera was seen as more "cold" and "concealed." The Mobile Service Technician and the Mobile Phone were quite close to each other, and both were associated with "task-oriented." Taken as a whole, the dimensions provide a wealth of information about how these users experienced the seven artifacts, and how they compared with each other.

DISCUSSION

This paper is primarily concerned with the use of RGT as a methodological tool for getting at people's experiences of using technology, relevant to the current concerns of HCI. We have shown how the procedure may be used to assess the experiences people have of designs, as in the study described above. In the following sections, we reflect further on the use of RGT as an element in research and design efforts, spotlighting ways in which it differs from other approaches in HCI.

Moreover, we point out that the RGT also can be employed during design, when included as a part of an iterative design cycle that aims for the user to have certain experiences. We might want to design, for example, a device that is experienced in a similar way to another existing device. This point is taken up in the concluding section of the paper.

RGT is an Open Approach

There are arguably some potential advantages of using RGT as compared to other candidate techniques for gaining insight into people's meaning structures. While RGT is a theoretically grounded, structured, and empirical approach, it is not restricted or limited to already existing, preprepared, or researcher-generated categories. Alternative approaches showing the same kind of openness as RGT include the semantic differential, discourse analysis, ethnography and similar observational methods, and unstructured interviews.

RGT is both Qualitative and Quantitative

Because a repertory grid consists of not only the personal constructs themselves but also a rating of them in relation to other elements in the study, the researcher not only gains insight into which are the meaningful constructs, but also the degree to which a particular construct applies or does not apply to a particular element. Hence, the RGT technique perhaps may be characterized best as being on the border between qualitative and quantitative research: a hybrid, "quali-quantitative" approach (Tomico et al., 2009).

On the one hand, a repertory grid models the individual perspectives of the participants, where the elicited constructs represent the participants' subjective differentiations. It may be used as such for various kinds of interpretative semantic analysis. On the other hand, since systematic ratings of all elements on all constructs result in a repertory grid consisting not only of elements and constructs but also of quantitative ratings, the resulting repertory grid may be subject to different kinds of quantitative analyses as well. The quantitative aspect of the RGT also provides the necessary means for comparing participants' grids with each other, using contemporary relational statistical methods. While RGT is reliant on statistical methods, semantic interpretation is sometimes needed to carry out specific parts of the analysis. By consistently using codes and markers, it is possible to track these interpretations back to the original data set.

RGT Results are Relational Rather than Absolute

Because RGT relies on comparisons between different elements, all results—such as the 10 unique dimensions of the example study—should be regarded as relative to the group of elements included in the study. The outcome of a study using this technique is not a set of absolute values. Rather, studies using RGT produce insights into people's experiences of particular things and the relationships between them. This potential disadvantage of the method was addressed in our example study by including already existing mobile information technology devices in the study to which the new research prototypes can be related. Doing so provided a result that, while still not absolute, nevertheless has become *situated*. In this respect, use of RGT is similar to the application of psychophysical rating scales to capture observers' perceptual judgments, which are always relative to the range of stimuli presented (e.g., Helson, 1964, Poulton, 1989; Schifferstein, 1995). Experiences can never be captured with the absolute precision of some physical measurements. Experiences can only ever be judged relative to other experiences, and the RGT approach emphasizes this fact.

RGT Addresses the User's Experience Rather than the Experimenter's

A famous contemporary and contrasting attempt at identifying and quantifying meanings and attitudes comes from the work of Charles Osgood in the 1950s (Osgood, Suci, & Tannenbaum, 1957). His semantic differential technique was developed to let people give responses to pairs of bipolar adjectives in relation to concepts presented to them (Gable & Wolf, 1993). The main adjectives used by Osgood included evaluative factors (e.g., good—bad), potency factors (e.g., strong—weak), and activity factors (e.g., active—passive). Each bipolar pair hence conceptually suggests a one-dimensional semantic space, a scale on which the participant was asked to rate a concept. Given a number of such pairs, the researcher is able to collect a multidimensional geometric space from every participant, much like the RGT approach.

However, researchers have raised a number of objections to and reservations about Osgood's technique. Among the most important is the recognition that the technique seems to assume that the adjectives chosen by the experimenter have the same meaning for everyone participating in the study. Also, since the experimenter provides the participants with the bipolar constructs, the former tends to set the stage, that is, provides the basic semantic space, for what kinds of meanings the participant can express for a particular concept. When participants merely rate construct pairs given to them, they are able to dismiss certain pairs as not appropriate or of no significance for a particular concept, but they have no way of suggesting new adjectives that they feel are more appropriate for describing something.

In contrast, the RGT approach does not impose the experimenter's constructs on participants. Rather, the method aims to elicit the users' own understanding of their experiences. In its first phase, RGT is clearly focused on eliciting constructs that are meaningful to the participant, not to the experimenter. The data in a particular participant's repertory grid is not interpreted in the light of the researcher's own meaning constructs.

Invested Effort

One disadvantage of RGT is that it requires a substantial investment of effort by both the experimenter and the participants at the time of construct elicitation, as compared to most quantitative methods. This has implications for both how many participants it is reasonable to have in a study, as well as for the length of each eliciting session. Although it would be better to expose each subject to as many triads as possible, doing so would not have been practically viable in this study, for the following reasons.

First, from around triad 8, we noticed that most participants' ability to find meaningful construct pairs began to decrease significantly, which was something that many of the participants also stated explicitly. Second, 10 triads also kept the length of each session to slightly more than an hour on average, which seemed to be a reasonable amount of time to expect people to concentrate on this kind of task.

Third, with seven elements, the number of possible unique triads exceeds 40, which is clearly far too many to expose to each participant (at least, if there is only a movie ticket at stake). This means that each participant was only exposed to a subset of all possible combinations of triads. However, because different participants were exposed to different triads, each unique group has been covered in the study as a whole.

On the other hand, RGT is more efficient and less time-consuming than most other fully open approaches, such as unstructured interviews and explorative ethnography. And, because the personal constructs elicited from participants constitute the study's data, it follows that using the RGT significantly reduces the amount of data that needs to be analyzed, compared with transcribing and analyzing unstructured interviews or ethnographic records.

Specific Issues Regarding the Elicitation Process

Two potential problems regard the actual conduct of constructing repertory grids. While these are generally not unique to RGT, they are worth noting. First, for various reasons, participants may feel inclined to provide the experimenter with socially desirable responses. In other words, a participant may experience a sense of social pressure during the elicitation session that makes her try to give the experimenter the "right answer." Second, some participants may, again for various reasons (e.g., that they feel uncomfortable in the situation, do not really have the time for the session, do not want to or cannot concentrate, do not really understand the purpose or doubt the study's usefulness, etc.), come to develop a habit of consistently providing moderate answers, or always either fully agreeing or disagreeing with their own constructs.

CONCLUSIONS

In this paper we have commented on the artificiality of assessing the emotional impact of interactive artifacts in isolation from cognitive judgments. We stressed that both emotion and reason are inherently part of any cognitive appraisal, and underlie the user's experience of an artifact. We suggested that studying the one without the other is – literally – meaningless. What HCI needs are techniques that recognize this and that provide practical solutions to the problem of how to assess the holistic meaning of users' interactive experiences.

In this light, a candidate method, the repertory grid technique (RGT), may partly fill this need, and has been presented, discussed, empirically exemplified, and explored. RGT was found to be an open and dynamic technique for qualitatively eliciting people's experiences and meanings in relation to technological artifacts, while at the same time providing the possibility for data to be subjected to modern methods of statistical analysis. The RGT may as such best be described as a research method on the border between qualitative and quantitative research. An example from the area of mobile HCI was used to take the reader step by step through the setting up, conducting, and analyzing of an RGT study.

How should a designer of interactive experiences think about the 10 dimensions of mobile technologies found in this study? Are they only relevant to this study and these devices, or are they general enough to provide a sound understanding of users' experience mobile information technology? The answer probably lies somewhere between these two possibilities.

Since RGT relies on comparisons between different elements, all results—such as the 10 unique dimensions surfaced in this study—must be regarded as relative to the group of elements that were included in the study. The 10 dimensions speak of something that is specifically about the seven technology designs provided to the participants. In a statistical sense, the resulting dimensions are relational to these seven devices. There is no way of

knowing whether they would change dramatically if an eighth device were to be added, without doing such an extended study.

But this limitation was to some extent addressed in the study by including already existing mobile information technology devices to which the new research prototypes can be related. Doing so provided a result that, while still not absolute, nevertheless has become more situated. It would not do justice to the study and the effort put into it by the participants to argue that the results are only valid within the study itself. On the contrary, we believe that the results from this study and the approach it illustrated could be useful for designers of mobile information technology, not the least as a tool for design.

Given that a team of designers wants to provide form and content to a mobile device that should embody certain characteristics, there are at least two ways in which this study can be used to guide the process. First, they may take the three existing devices as a basis and consider the four prototypes to provide a large number of alternative design dimensions. If they want their design to provide its users with a sense of mysteriousness, for instance, then aspects of the Reality Helmet may be taken as influence. Second, designers may use this study as the basis for designing and conducting their own studies in similar ways. If they want to find out whether their design really is experienced as mysterious, they can set up and conduct their own repertory grid study in a similar fashion, perhaps even using the same existing devices as were used here. Such comparisons can at least provide some hints and traces of meaning that may be very useful for further design work. The design team may also wish to embed small repertory grid studies throughout the production cycle to monitor designs against some sought-after set of qualities of user experience: These grids could become a recurring element in organizing the process of interactive artifact design.

RGT is unique in that it respects the wholeness of cognition: It does not separate the intellectual from the emotional aspects of experiences. At the same time, it acknowledges that each individual creates her own meaning in the way she construes things to be, in the context in which they are experienced. RGT has the advantage of treating experiences holistically, while also providing a degree of quantitative precision and generalizability in their capture.

REFERENCES

Bannister, D., & Fransella, F. (1985). *Inquiring man* (3rd ed.). London: Routledge.

Berg, J. (2002). Systematic evaluation of perceived spatial quality in surround sound systems. Doctoral Thesis. Luleå University of Technology (2002: 17), Sweden.

Boose, J. H., & Gaines, B. R. (Eds.). (1988). *Knowledge acquisition tools for expert systems*. London: Academic Press.

Dalton, P., & Dunnet, G. (1992). *A psychology for living: Personal construct psychology for professionals and clients*. London: Wiley.

Damasio, A. (1994). *Decartes' error: Emotion, reason and the human brain.* New York: Penguin Putnam.

Damasio, A. (1999). *The feeling of what happens: Body, emotion and the making of consciousness*. San Diego, CA, USA: Harcourt Brace and Co., Inc.

Dillon, A., & McKnight, C. (1990). Towards a classification of text types: A repertory grid approach. *International Journal of Man–Machine Studies, 33,* 623–636.

Fallman, D. (2002). Wear, point, and tilt: Designing support for mobile service and maintenance in industrial settings. In *Proceedings of DIS 2002: Designing Interactive Systems* (pp. 293–302). New York: ACM Press.

Fallman, D. (2003). Design-oriented human-computer interaction. In *Proceedings of Conference on Human Factors in Computing Systems* (CHI 2003; pp. 225–232). New York: ACM Press.

Fallman, D., Jalkanen, K., Lörstad, H., Waterworth, J., & Westling, J. (2003). The Reality Helmet: A wearable interactive experience. In *Proceedings of SIGGRAPH 2003: International Conference on Computer Graphics and Interactive Techniques, Sketches & Applications* (p. 1). New York: ACM Press.

Fallman, D., Lund, A., & Wiberg, M. (2004). ScrollPad: Tangible scrolling with mobile devices. In *Proceedings of the 37th Annual Hawaii International Conference on System Sciences* (HICSS ’04; on CD) Washington, DC: IEEE Computer Society

Fallman, D. (2006, Nov). *Catching the interactive experience: Using the repertory grid technique for qualitative and quantitative insight into user experience*. Paper presented at Engage: Interaction, Art, and Audience Experience, Sydney, Australia.

Fallman, D., Andersson, N., & Johansson, L. (2001, June). *Come together, right now, over me: Conceptual and tangible design of pleasurable dupliances for children*. Paper presented at the 1st International Conference on Affective Human Factors Design, Singapore.

Forlizzi, J., & Ford, S. (2000). The building blocks of experience: An early framework for interaction designers. In *Proceedings of DIS 2000: Designing Interactive Systems* (pp. 419–423). New York: ACM Press.

Forlizzi, J., & Battarbee, K. (2004). Understanding experience in interactive systems. In *Proceedings of DIS 2004: Designing Interactive Systems* (pp. 261–268). New York: ACM Press.

Fransella, F., & Bannister, D. (1977). *A manual for repertory grid technique.* London*:* Academic Press.

Gable, R. K., & Wolf, M. B. (1993). *Instrument development in the affective domain* (2nd ed.). Boston: Kluwer Academic Publishers.

Gaines, B. R., & Shaw, M. L. G. (1980). New directions in the analysis and interactive elicitation of personal construct systems. *International Journal Man–Machine Studies, 13*, 81–116.

Gaines, B. R., & Shaw, M. L. G. (1993). Eliciting knowledge and transferring it effectively to a knowledge-based system. *IEEE Transactions on Knowledge and Data Engineering, 5*(1), 4–14.

Gaines, B. R., & Shaw, M. L. G. (1995). WebMap: Concept mapping on the web. *World Wide Web Journal*, *1*, 171–183.

Grose, M., Forsythe, C., & Ratner, J. (Eds.). (1998). *Human factors and web development*. Mahwah, NJ, USA: Lawrence Erlbaum Associates.

Hassenzahl, M., & Wessler, R. (2000). Capturing design space from a user perspective: The repertory grid technique revisited. *International Journal of Human–Computer Interaction, 12*, 441–459.

Hassenzahl, M., & Tractinsky, N. (2006). User experience: A research agenda [Editorial]. *Behavior & Information Technology, 25*, 91–97.

Helson, H. H. (1964). *Adaptation-level theory*. New York: Harper & Row.

Kelly, G. (1955). *The psychology of personal constructs* (Vols. 1 & 2). London: Routledge.

Ketola, P., & Roto, V. (2008, June). *Exploring user experience measurement needs*. Paper presented at the 5th COST294-MAUSE Open Workshop on Valid Useful User Experience Measurement (VUUM), Reykjavik, Iceland.

Landfield, A. W., & Leitner, L. (Eds.). (1980). *Personal construct psychology: Personality and psychotherapy*. New York: Wiley.

Landauer, T. (1991). Let’s get real: A position paper on the role of cognitive psychology in the design of humanly useful and usable systems. In J. M. Carroll (Ed.), *Designing interaction: Psychology at the human-computer interface* (pp. 60–73). New York: Cambridge University Press.

Law, E., Roto, V., Hassenzahl, M., Vermeeren, A., & Kort, J. (2009). Understanding, scoping and defining user experience: A survey approach. In *Proceedings of Conference on Human Factors in Computing Systems* (CHI 2009; pp. 719–728). New York: ACM Press.

McCarthy, J., & Wright, P. (2004). *Technology as experience.* New York: The MIT Press.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning. Chicago:* University of Illinois Press.

Poulton, E. C. (1989). *Bias in quantifying judgments*. Hillsdale, NJ, USA: Lawrence Erlbaum.

Schifferstein, H. J. N**.** (1995). Contextual shifts in hedonic judgments. *Journal of Sensory Studies, 10,* 381–392.

Shaw, M. L. G. (1980). *On becoming a personal scientist: Interactive computer elicitation of personal models of the world*. London: Academic Press.

Shaw, M. L. G., & Gaines, B. R. (1983). A computer aid to knowledge engineering. In *Proceedings of British Computer Society Conference on Expert Systems* (pp. 263–271) Cambridge, UK: British Computer Society.

Shaw, M. L. G., & Gaines, B. R. (1987). KITTEN: Knowledge initiation & transfer tools for experts & novices. *International Journal of Man–Machine Studies, 27,* 251–280.

Shaw, M. L. G., & Gaines, B. R. (1995). Comparing constructions through the web. In *Proceedings of CSCL95: Computer Support for Collaborative Learning* (pp. 300–307). Hillsdale, NJ, USA: Lawrence Erlbaum.

Shaw, M. L. G., & Gaines, B. R. (1998, April). *WebGrid-II: Developing hierarchical knowledge structures from flat grids.* Paper presented at KAW '98: The Eleventh Workshop on Knowledge Acquisition, Modeling and Management, Banff, Alberta, Canada.

Slater, P. (Ed.). (1976). *Dimensions of intrapersonal space* (Vol. 1). London: John Wiley.

Steed, A., & McDonnell, J. (2003, Oct). *Experiences with repertory grid analysis for investigating effectiveness of virtual environments*. Paper presented at Presence 2004, Aalborg, Denmark.

Suchman, L. (1987). *Plans and situated actions: The problem of human–machine communication.* New York: Cambridge University Press.

Tan, F. B., & Hunter, M. G. (2002). The repertory grid technique: A method for the study of cognition in information systems. *MIS Quarterly*, *26*, 39–57.

Tomico, O., Karapanos, E., Levy, P., Mizutani, N., & Yamanaka, T. (2009). The repertory grid technique as a method for the study of cultural differences. *International Journal of Design*, *3*(3), 55–63.

Waterworth, J. A., & Fallman, D. (2003). The Reality Helmet: Transforming the experience of being-in-the-world. In *Proceedings of HCI 2003: Designing for Society* (Vol. 2, pp. 1–4), Bristol, UK: Research Press International.

Waterworth, J. A., & Fallman, D. (2007, March). *Capturing users' experiences of interactive mobile technology*. Poster presentation at the British Psychological Society Annual Conference, York, UK.

Winograd, T., & Flores, F. (1986). *Understanding computers and cognition: A new foundation for design*. Norwood, NJ, USA: Ablex Publishing Corporation.

Authors' Note

All correspondence should be addressed to
Daniel Fallman
Interactive Institute Umeå
c/o Umeå University, School of Architecture
SE-90187 Umeå, Sweden
daniel.fallman@tii.se

Human Technology: An Interdisciplinary Journal on Humans in ICT Environments
ISSN 1795-6889
www.humantechnology.jyu.fi

HUMAN TECHNOLOGY

An Interdisciplinary Journal on Humans in ICT Environments ISSN: 1795-6889

www.humantechnology.jyu.fi Volume 6 (2), November 2010, 269–271

BOOK REVIEW

Hautamäki, A. (2010). *Sustainable Innovation: A New Age of Innovation and Finland's Innovation Policy* [Sitra Reports 87]. Helsinki, Finland: Sitra and Edita Prima, Ltd. 144 pages.

Reviewed by Ignacio Del Arco Herrera
Head of Scientific Department and Coordinator
i2BC (Institute of Innovation for Human Wellbeing)
Technology Park of Andalucía
Málaga, Spain

Innovation requires the provision of new capabilities to available resources in order to generate value. For many years, the value generated by innovative activities has been chrematistic, based mainly on the economic incentive for companies producing innovations and/or exploiting them in the markets. However, various driving forces are changing this biased view of the power in the outcomes of innovation.

Permanent change is part of the nature of innovation. But also is knowledge, evolving continuously through research, experimentation, and mastery in order to develop new solutions. Innovators, both individually and collectively, must become and remain fully aware of the actual forces and trends that are shaping the world in the new age of innovation, and use the cumulative base of knowledge and available technologies to anticipate the best possible solutions for today's and tomorrow's global challenges. This forms the basis of the book *Sustainable Innovation.*

The author has included in the book various elements regarding the power of the innovation process to solve many of the problems for present-day research–development–innovation systems and also addressed future challenges beginning to surface on a global level. The book is not about a new innovation model (although it presents the sustainable innovation model), nor is it another book on innovation policy (although it contains very interesting tips and recommendations for leaders and decision makers). Furthermore, the book is not about the basis for the knowledge-based economy (although it reviews the most important implications of innovation with economy). Rather, the book provides a very well founded reflection as a call to action. In a very smart approach, the author shifts from the concept of sustainable development towards sustainable innovation, where innovative activities are based on ethically, socially, economically, and environmentally sustainable principles.

URN: NBN:fi:jyu-201011173095

Sustainable innovation is "the interactive process in which different forms of capital (industrial or physical, human, natural, social) are used in a balanced way to promote long-term human development and the good of people, in Finland and throughout the world" (p. 22). The equation of innovation–productivity–growth no longer works nowadays. It has been replaced by the new formula of innovation-reproduction-wellbeing, by which the economic systems (markets), natural systems (resources) and social systems (citizens and institutions) will find a balanced point to address the "wicked" problems (democracy, poverty, energy, water, forest, etc.).

With Finland being one of the most innovative countries in the world, and having proven the strengths of the links between education, technology and innovation with economy and welfare, some could be puzzled about the need to reformulate the models that have helped develop "the world's best country" according to the *Newsweek Magazine* international study in summer of 2010. As William Pollard (1996, p. 114) said, "Learning and innovation go hand in hand. The arrogance of success is to think that what you did yesterday, will be sufficient for tomorrow." The book also analyzes the Finnish knowledge and innovation systems, providing insightful information about the strategies for success, but focuses also on its weaknesses, providing points for improvement that will prevent the Finnish system from becoming "arrogant" and eventually dying from it own success. Professor Hautamäki moves back and forward in his analysis from a global landscape to the local innovation ecosystems, carefully distilling the most interesting arguments of updated innovation currents and cases worldwide as a learning source for the new concept of sustainable innovation.

Education, creativity, decentralization, networking, and leadership (all of them addressed in the book) are the main characteristics of the modern innovation pathways that are already changing the business activities, regional and national economies, and, hopefully in the near future, the life of more and more citizens of the new world we are starting to live in.

Of these characteristics, I would say that education is one of the most important. Education empowers people through knowledge, so that people can turn their own creative potential into real innovations to improve their living conditions, environment, and quality of life. A central role in the book is given to universities for creating the basis for the innovation activity and even as the core for innovation economy. Quality teaching and research developed in universities may of course exert a direct effect on the economy, although universities in most countries should undergo a deep transformation to facilitate innovation and avoid becoming stuck in their rigid (and in many cases old-fashioned) structures and ways of working. In this transformation, the third mission of the universities must be given the importance it deserves, leveraging the social responsibility of universities as a fundamental piece of knowledge-based modern societies. This vision is valid as well for elementary and secondary education systems, which play a very important role in transmitting the new values, concepts, and resources that will permit a broad base of active participation by citizens in the new society.

Sustainable innovation is based upon five principles: sustainable development; participative, continuous and global innovation; and innovative management. These five factors constitute a strong base to reinforce the performance of innovators and open the eyes of other stakeholders who can contribute to helping sustainable innovative practices achieve their true potential. This perspective is highlighted by two quotes drawn from the conclusions of the book: "The innovation activity of nations and companies must be directed towards

solving the most dramatic and threatening issues facing humankind" (p. 129) and "Sustainable innovation starts with positive thinking, a belief that creativity and innovation can overcome the pending challenges linked with environment and society" (p. 130).

REFERENCES

Newsweek [online]. (2010). The world's best countries: A Newsweek study of health, education, economy, and politics ranks the globe's true national champions. Retrieved November 15, 2010, from http://www.newsweek.com/feature/2010/the-world-s-best-countries.html

Pollard, C. W. (1996). *The soul of the firm*. New York: Harper Collins Publishers.

Human Technology:
An Interdisciplinary Journal on Humans in ICT Environments

www.humantechnology.jyu.fi