
An Implicit Test of UX: Individuals Differ in What They Associate with Computers

Martin Schmettow

University of Twente
Drienerlolaan 5
7522 NB Enschede, Netherlands
m.schmettow@utwente.nl

Matthijs L. Noordzij

University of Twente
Drienerlolaan 5
7522 NB Enschede, Netherlands
m.l.noordzij@utwente.nl

Matthias Mundt

University of Twente
Drienerlolaan 5
7522 NB Enschede, Netherlands
m.mundt@student.utwente.nl

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2013 Extended Abstracts, April 27–May 2, 2013, Paris, France.

Copyright © 2013 ACM 978-1-4503-1952-2/13/04...\$15.00.

Abstract

User experience research has made considerable progress in understanding subjective experience with interactive technology. Nevertheless, we argue, some blind spots have remained: individual differences are frequently ignored, the prevalent measures of self-report rarely undergo verification, and overly focus is on utilitarian and hedonic dimensions of experience.

A Stroop priming experiment was constructed to assess what people implicitly associate with a picture of a computing device. Three categories of target words were presented: hedonic, utilitarian and “geek” words. Longer response times were interpreted as stronger associations. Need-for-cognition and subject of undergraduate study (computer science vs. psychology) were taken as predictors for a hypothetical geek personality. The results suggest that persons with a geek predisposition tend to think of computers as objects of intellectual challenge and play, rather than tools or extensions of the self.

Author Keywords

geekism; hedonism; usability; user experience; Stroop task; priming; multi-method; implicit method; mixed-effects models; individual differences; need for cognition

ACM Classification Keywords

H.5.2. User Interfaces: Evaluation/methodology

General Terms

Experimentation, Measurement, Theory

Introduction

Researchers in the field of Human-Computer Interaction have always sought to quantify and compare qualities of interactive products. Many studies in classic usability engineering, and a majority in more recent user experience research, measure qualities in use by self-reported judgments. While self-report instruments, such as Likert scales or semantic differentials have economy and flexibility as their merits, they also have their limitations and potential biases. The *first goal* of the present work is to extend the current (explicit) methodologies with an experimental method, the Stroop priming task, to *implicitly* assess the spontaneous associations with a product.

In classic usability research, quality assessment of products has primarily focused on directly observable utility attributes. This is prominently represented in the ISO standard 9241-11, with its two sub-criteria effectiveness and efficiency, as well as its emphasis on user goals and tasks [23]. An array of behavioral measures have been developed and used by researchers to objectively assess or compare effectiveness and efficiency of user-system interaction [22].

At the same time, the ISO 9241-11 acknowledges that the value of a system or design extends beyond the mere utilitarian qualities. It introduces the subjective value of *user satisfaction* as "*freedom from discomfort, and positive attitude to the use of the product.*" The

advent of user experience (UX) research has largely expanded the view on subjective values in user-system relation. Users' subjective experience with interactive products is now widely recognized as holistic, dynamic, and multidimensional [3]. In addition to the classic definition of satisfaction, users' attitudes are now also defined by experiential aspects such as beauty [21], joy [27] and growth [18]. However, in recent UX research, little attention has been paid to individual differences. As of current, it is unclear whether all users have the same preference towards qualities of usability or UX. The *second goal* of this study is to examine if implicit measure can reveal individual differences in how users associate values with computer products.

We were specifically interested in a hypothetical user attitude that has rarely been regarded in usability or UX research. We coin the term "geekism" to capture a predisposition that we associate with great affinity for exploring and tinkering with technological devices. The *third goal* of this research is to test if signs of geekism can be observed as associations in the Stroop priming task.

Self-report measures

Bargas-Avila and Hornbæk, in a review of user experience research [3], found that a large majority of studies use methods of self-report to assess quality attributes, such as aesthetic appeal, enjoyment or hedonic quality. They also criticize that subjective rating scales are frequently constructed ad hoc, and rarely undergo a rigorous validation. Furthermore, they suspect that some common scales, such as hedonic value and beauty, may suffer from low discriminative validity.



Figure 1 Conditions of the classical Stroop task (ink colors from top to bottom: green, blue, red). Response times in the incongruent condition are slowest.

Even when carefully designed, measures of self-report are not without issues. A sequence of complex cognitive operations mediates between reading a questionnaire item and setting a mark on a Likert scale [25]. Biases and spurious results are likely to happen on every stage of the judgment process. A comprehensive review of self-report measures is beyond the scope of this paper, instead the reader is referred to other treatments of the topic, such as the work of Schwarz and colleagues (e.g., [32,33]) and to Lucas and Baird [25]. The latter authors conclude that “*errors that result from respondents’ inability to remember past behaviors or their unwillingness to accurately report their feelings are unlikely to be shared across different measurement techniques.*” Hence, they call for multi-method assessment, where experimental techniques or implicit measures complement or validate self-reports.

The Stroop priming task

Facing the limitations of self-report scales, the fields of personality psychology and consumer psychology are currently adopting *implicit methods* to assess an individual’s predisposition [30]. As an implicit experimental method, the *Stroop priming task* assesses which associations an individual has with a priming stimulus, for example, the picture of a smart phone. Longer response times in the subsequent color-naming task indicate strong associations between prime and target.

In the classical Stroop task, the subjects see color words, but have to name the ink color, not the word itself. Word and ink color are either congruent or incongruent, or the word is neutral (see Figure 1). The typical result is, that subjects’ response is delayed in the incongruent condition. This is called the *Stroop effect*. The Stroop effect is commonly interpreted as an inter-

ference of the target word’s meaning with the color naming task. The Stroop effect has been replicated in dozens of studies over several decades and has been found robust in many variants [26].

A more general interpretation of the Stroop effect is, that the stronger a target word captures the attention of a person, the more delayed is the response. People who have been disposed to think about a certain topic typically show slowed reaction times in naming the ink color, when this topic is semantically associated with the target word. As an example, imagine, a person sees a picture of a bank and makes an association with “money”. This person is likely to show delayed reaction time when seeing the target word “dollar”, compared to, let’s say, “flower”.

This effect is called *priming* and can be explained with the theory of *spreading activation*. Concepts are thought to be represented in semantic memory as nodes, and learned relations between them as associative pathways [2]. When a node in the network gets activated, this activation will spread along the associative pathways to connected areas in memory. This spread of activation increases availability of related concept for further cognitive processing, which is referred to as *priming* [36]. At the same time, a forming association consumes attention and distracts the person from the primary task of naming the color. In the original Stroop experiment, the task itself is the distracting prime: the meaning of the color word is literally related to the task itself, which is naming a color.

In a recent applied example, Sparrow, Liu and Wegner [34] found that persons who had been given a knowledge question as a prime, showed delayed re-

sponse times on target words related to computers and the World Wide Web. They take this as a sign of how modern technology has changed the way we approach knowledge tasks, in that we rapidly think of computers.

In our study, we apply the Stroop priming task to measure associations that are activated when seeing an interactive computing device, for example a smart phone. To assess the direction of the association, we used three categories of target words: hedonic, utilitarian, and "geek". We predict individual differences in the direction of the associations.

Individual differences

In 1996, Dillon and Watson regretted: "*The study of individual differences is as old as psychology itself, and one may wonder how it has remained so marginal to mainstream HCI which is usually receptive to psychological theory.*" [10:620] In the meantime, quite a few HCI studies examined individual differences. Most seem to examine the impact of cognitive, perceptual or motor capabilities on interaction performance (e.g., [12,14]). A seemingly smaller number of studies address the role of personality and traits in user-system interaction (e.g., [1,6]).

In contrast, virtually all recent UX studies seem to resort to what Dillon and Watson call an experimentalist, as opposed to a differentialist, perspective, assuming "*relative homogeneity among subjects of whatever ability is required to perform a task, often relegating inter-subject differences into the category of error variance.*" [10:621] For example, studies on the interplay between usability and hedonic properties typically focus on co-variation of different subjective ratings on products [19], but little is known on the subjective rele-

vance of usability compared to hedonic qualities. Some other studies address situational factors, such as the impact of different instructions in an interaction task. For example, Hassenzahl and Ullrich showed how induced instrumental goals change retrospective evaluation of a usage episode [20]. But, it has to our knowledge not been examined how individuals differ in their predisposition to form instrumental, or other, goals when using a computer.

Hedonism, utility, and geekism

As we have outlined above, the classic view on product quality is instrumental, focusing on utility. Hassenzahl and Ullrich showed that having an instrumental motivation to use a product, changes behavior and attitude of a user [20]. If focusing on utility is a motivational state, we find it compelling to ask: is this just situational or may there be an underlying *motif*, where individuals differ in how utilitarian they are? The same question may be asked for an individual's tendency to prefer hedonic qualities, (e.g. visual appeal, social identity) over utility, or as Diefenbach and Hassenzahl put it: be-goals over do-goals [9]. A purely utilitarian individual, we presume, thinks of technology as a tool to complete tasks and reach goals. For utilitarian users, functionality and usability are preferred qualities. In contrast, a hedonic user would appreciate the surface features of a product, such as brand and visual appeal.

We presume that a third motivational predisposition may play a role: technology enthusiasm, or *geekism* as we call it. Here, with *geekism* we denote an individual's strong urge and endurance to understand the inner workings of a computer system. The stereotypical geek user prefers a Linux box over computers equipped with Windows, spends more time on customizing a smart

Target word examples

Hedonism: attractive, popular, stylish, impression, exciting, elegant, pride

Utility: useful, potent, performance, tool, govern, serving, multifunctional

Geekism: understand, improve, master, configure, play

Prime picture examples

phone than using it, and is enthusiastic about the idea that all electronic devices in his household communicate to each other via network. Likely, he or she also helps friends, parents and grandparents in maintaining and upgrading their software and devices.

The word “geek” originally denoted actors in freak shows, but has since become a synonym for persons with devotion for computers, at the expense of normal social life. We, in contrast, have no intention to carry on any stereotypical or stigmatizing ideas about the social life of persons. Neither do we deal with any other form of enthusiasm, for example music or cars. The concept of geekism in this paper draws solely upon the assumption that persons differ in enthusiasm for computers, especially the tendency to see computers as interesting objects in themselves.

In our study, we examine whether geekism can be distinguished from the two prevalent dimensions of attitude towards computers: hedonism and utilitarianism. The Stroop priming task serves to elicit associations with computers, as an indicator for the prevalent attitude of a person. Accordingly, three categories of target words were created for the Stroop task: hedonism, utility, and geekism. The strength of association with one of the word categories is measured by response latency in the color naming task. It is expected that individuals with a stronger geek predisposition show stronger associations with geekism words, after being primed by a picture of a computing device.

Assessing the geek predisposition

What could be a prevailing motif of geek users? We imagine geek users as having a strong urge to understand the inner workings of systems. They appreciate

the intellectual effort to master a technical system. So, the degree to which an individual enjoys intellectually demanding tasks, may be a good predictor.

In order to approximate individual differences related to the geekism concept we have opted for the Need for Cognition Scale (NCS) [8]. Individuals with high scores on the NCS scale tend to be flexible in their choice of learning strategies. In addition, they are usually highly motivated for challenging tasks, not strongly influenced by surface features (e.g. aesthetic aspects of a system), and they have excellent control over their attentional resources [7,31]. In contrast, individuals low in need for cognition show little affection for complex thought and are considered to rely more on others to find meaning in outside events [13]. Taken together, we expect a relation between the need for cognition scores (as an approximation for geekism traits) and the strength of associations with geekism target words.

As a second predictor for geekism, we chose the subject of undergraduate study of the participant. It is assumed that computer science students have a stronger predisposition for geekism, as compared to psychology students. While the need-for-cognition score captures appreciation of intellectual challenges, subject-of-study is believed to capture the preference for computer-related topics.

Method**Sample**

Forty-one Dutch University students participated in the study, 16 were enrolled in a Computer Science (CS) program and 25 were enrolled in a Psychology program (PSY). CS students were rewarded with six Euro and the chance to win one of two coupons with a value of



Figure 2 Sequence of a Stroop priming trial

30 Euro. The PSY students participated as part of their course fulfillment.

Materials & Apparatus

One author generated and classified 90 Dutch target words, 32 for hedonism, 28 for utility and 30 for geekism (for examples see preceding page). Another author independently classified the words and Cohen's Kappa was calculated to 0.835, giving support for an acceptable inter-rater reliability. We attempted to exclude overly technical terms in the geekism category, as this may introduce a confounding with word familiarity between the two groups of students. Note however, that word familiarity effects often are negligible in the Stroop experiment, or, if they happen, are in the direction that familiar words are processed quicker [11]. Remember, that response is slower in the presence of strong associations. This means, that any confound with word familiarity would obscure rather than produce the sought effects.

Black-and-white pictures of five smartphones, five tablets and five notebooks were used for priming. The pictures were rather neutral, neither showing the product in operation, nor being very stylish. Visible brand marks were erased from the pictures to exclude a possible bias. The Stroop task experiment was designed with E-prime [29] and administered on a Windows PC, optimized for reliable response time measures.

Procedure

Individual sessions started with the experiment. First, subjects were briefed to use the Z, X, N, and M keys to respond to the color of the shown targets (red, blue, green and yellow) [4]. They got instructions to watch the priming picture and respond as accurate as possible

to the color naming task. Two training blocks with neutral primes (grayscale pictures of fruits) and targets ('XXXXXX') made the subjects familiar with the task.

The actual experiment consisted of six blocks with 15 trials each, and five short breaks to prevent fatigue. A set of 15 prime pictures were used, each appearing once per block. Ninety different target words of the three categories (hedonism, utility, geekism) were randomly assigned to the same number of trials. Colors were randomly assigned to the trials. The within trial sequence starts with the priming image, which is shown for 5 seconds, giving subjects sufficient time to form associations. Following a fixation cross (1s) the Stroop color naming task is given (Figure 2) and response time is measured.

After the subjects completed the Stroop task, they were asked to fill out the need-for-cognition scale, consisting of 18 five-point Likert scale items, such as: "The notion of thinking abstractly is appealing to me." Finally, the subjects were debriefed and thanked for their participation.

Data analysis

In the experiment the subjects' undergraduate study and NCS score are between-subject factors. The word class (hedonism, utility, and geekism, in the following denoted as HUG) is a within-subject factor. Every subject encountered each of the 90 HUG words once.

The relationship between the predictors Study and NCS and the response time was estimated by a mixed-effects model. Mixed-effects models have several advantages compared to classic repeated measures ANOVA, such as handling unbalanced and incomplete

Table 1 Estimated fixed effects coefficients, with 95% confidence limits and alpha error. Reference groups for treatment contrasts are HUG=Geekism and Study=PSY.

Variable	Coef	CI95	P
Intercpt	1126.6	1043.6 1213.8	.00***
Order	-1.6	-2.5 -0.6	.00***
CS	-52.3	-173.2 66.8	.39
Hedo	-3.6	-58.7 54.9	.90
Util	-51.9	-112.1 7.9	.10 ⁺
NCS	32.1	-29.8 89.8	.28
CS:Hedo	-65.4	-122.1 -1.7	.03*
CS:Util	-15.6	-77.8 47.1	.63
NCS:Hedo	-18.1	-45.7 12.2	.22
NCS:Util	-32.4	-62.6 -1.5	.04*

experimental designs, offering much greater flexibility in choosing covariates and better statistical power [16].

Assessing statistical significance in mixed effects models by asymptotic statistical tests (such as the F-test) is not without problems [5]. For this reason, it was opted for a Bayesian estimation via Markov-Chain Monte-Carlo sampling, using the *MCMCglmm* program from the correspondent package [17] as supplied in the R system for scientific computing [28]. Weakly informative priors were used.

Confidence limits of coefficients were obtained from the highest posterior density and used for hypothesis testing. Note that although we estimate several effects at once, we did not correct the alpha level for multiple hypothesis testing. According to Gelman, correcting the family-wise error level usually is unnecessary in Bayesian mixed-effects models: due to the mechanism of partial pooling, estimates are pulled closer towards the grand mean, attenuating any effect. This is also called "shrinkage".[15].

Following [24], two intercept random effects were introduced: one subject-level random effect for the overall reaction time of a subject, and one material-level random effect for the overall tendency of words.

Results

In total, we obtained 3690 response time measures, 90 per subject and 41 per word. Mean response time was 1007ms ($sd=368$). The need-for-cognition score (NCS) ranged from 2.4 to 4.6 ($m=3.7$, $sd=0.54$). For all subsequent analyses, the z-standardized NCS score was used. The data was visually screened, following the protocol in [37]. Residuals of the response time had a

slight skew, and variance seemed to weakly increase by the predicted means. A logarithmic transformation did not reduce the problem. So, we stayed with the original scale, but explicitly modeled for heteroscedasticity between groups. An association between NCS and Study was observed ($F=9.221$, $p<.01$), with CS students showing higher need for cognition on average. While statistically significant, this association is weak ($R^2=0.17$), so there is low risk of colinearity. Although subjects underwent a practice phase, a visible learning effect remained. Therefore, we added the trial order as control variable. Visual inspection suggested that the learning rate varied strongly between individuals. This is accounted for a slope random effect for trial order. No effect of word length was observed.

The primary question, whether subjects' response time to geekism words depends on NCS or Study, is expressed as two interaction terms $NCS \times HUG$ and $Study \times HUG$. Table 1 shows the estimated fixed-effects coefficients. First, the order effect is evidential; on average subjects respond 1.6ms faster with every trial. Computer science students on average are faster, too ($\Delta_{RT}=53ms$), however not reaching statistical significance. While hedonic words have almost the same response time as geek words (the reference group), responses to utility words seem to be faster ($\Delta_{RT}=53ms$, $p \leq .1$). Higher need-for-cognition is associated with longer response times ($\Delta_{RT}=35ms$ per z-standardized step), but not beyond chance level.

The interaction effects between HUG word categories and subject-of-study are illustrated in Figure 3. The lines of geekism and utility are almost parallel, but responses to hedonic words are over-proportionally shorter in the CS group ($\Delta_{RT}=65ms$, $p \leq .05$). Figure 4

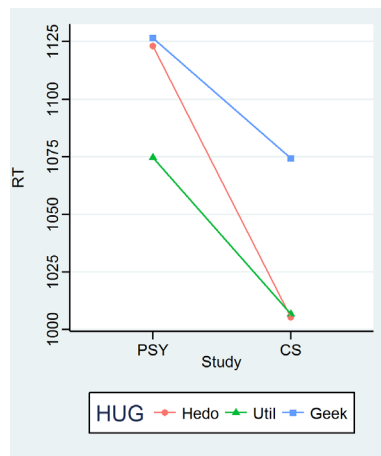


Figure 3 Interaction effect between word category (HUG) and Study on response time

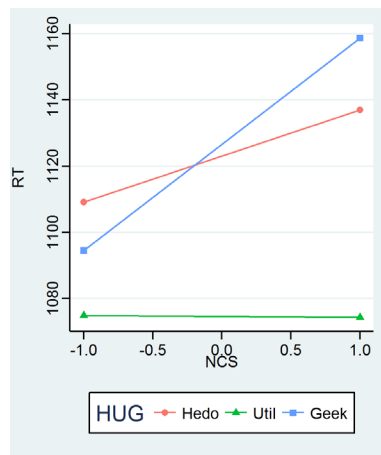


Figure 4 Interaction effects between word category (HUG) and need-for-cognition score (z-standardized) on response time.

shows the interaction effects between HUG categories and need-for-cognition scale (NCS). While response times for geekism words are delayed by 35ms per standard deviation, Hedonism words are delayed by only 16ms. Response times for utility words are virtually unrelated to NCS (2ms). The difference between geekism and utility words in association with NCS reaches statistical significance ($p \leq .05$).

In summary, subjects with a geek predisposition, as assessed by Study and NCS, show stronger associations with geekism words. Computer science students show stronger associations with geekism words compared to hedonic words. Individuals with high need-for-cognition have stronger associations with geekism words, compared to utility words.

Discussion & Conclusion

We were searching for individual differences and found them in interaction effects between geekism predictors and target word categories. We successfully predicted that computer science students would show strongest associations with geekism target words. When they see the picture of a computer, they tend to rapidly think of concepts such as exploration, play and comprehension.

It surprised us that overall subjects seemed to form stronger associations with geekism words compared to utility words, more so when rating themselves high on need-for-cognition. This seems to contradict the classical perspective in usability engineering: users want interactive products to be simple, task-oriented, and efficient to use, but never “make them think.”

The majority of UX studies only observe main effects and average over a potentially diverse population. In

the Stroop task we measured spontaneous associations with hedonism, utility and geekism words, and found predictable individual differences. We suggest that more UX studies consider individual differences, and think of which traits can make a difference in attitude towards software products and computers. For example, Diefenbach and Hassenzahl argue that self-reported preference of pragmatic over hedonic values is caused by a need for justification, biasing the real preference (which they believe must be hedonic) [9]. Our results partly support this view: on the one hand, utility target words generally had the shortest response latency. Seemingly, people think of utility last. On the other hand, associations with geekism words were strongest for at least a part of the sample. We conclude that fascination for technology *per se* might also play a role in user experience. This hypothetical trait we called geekism, and its role in HCI needs further investigation.

A few recent UX studies raised awareness for lower level cognitive processes that form and potentially bias users’ (or consumers) self-reported judgments. Examples are the above-mentioned justifiability bias [9] and the role of instrumental goals [20]. The only UX study we could find, that uses an implicit measure for a perceived quality, is Tractinsky *et al.* [35]. They found that judgments of beauty were related to response latencies in making this judgment. This relation was not simple, but seemingly curvilinear. In our view, the surfacing complexity of the relationship between judgment and latency hints at the complexity of cognitive processes underlying self-reported measures. Implicit experimental methods, such as the Stroop priming task, may serve to better understand the nature of rating scales in HCI, and give more direct access to users’ spontaneous associations and affects.

References

- [1] Aykin, N. Individual differences in human-computer interaction. *Computers & Industrial Engineering* 20, 3 (1991), 373–379.
- [2] Balota, D. a. and Lorch, R.F. Depth of automatic spreading activation: Mediated priming effects in pronunciation but not in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 12, 3 (1986), 336–345.
- [3] Bargas-Avila, J.A. and Hornbæk, K. Old wine in new bottles or novel challenges. *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*, ACM Press (2011), 2689 – 2698.
- [4] Besner, D., Stolz, J. a, and Boutilier, C. The stroop effect and the myth of automaticity. *Psychonomic bulletin & review* 4, 2 (1997), 221–5.
- [5] Bolker, B.M., Brooks, M.E., Clark, C.J., et al. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in ecology & evolution* 24, 3 (2009), 127–35.
- [6] Borgman, C.L. All users of information retrieval systems are not created equal: An exploration into individual differences. *Information Processing & Management* 25, 3 (1989), 237–251.
- [7] Cacioppo, J.T., Petty, R.E., Feinstein, J.A., and Jarvis, W.B.G. Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin* 119, 2 (1996), 197–253.
- [8] Cacioppo, J.T., Petty, R.E., and Kao, C.F. The efficient assessment of need for cognition. *Journal of personality assessment* 48, 3 (1984), 306–7.
- [9] Diefenbach, S. and Hassenzahl, M. The dilemma of the hedonic – Appreciated, but hard to justify. *Interacting with Computers* 23, 5 (2011), 461–472.
- [10] Dillon, A. and Watson, C. User analysis in HCI — the historical lessons from individual differences research. *International Journal of Human-Computer Studies* 45, 6 (1996), 619–637.
- [11] Effler, M. Interference by Stroop items depending on word frequency training and reaction times of the word components. *Zeitschrift für Experimentelle und Angewandte Psychologie* 18, 1 (1981), 54–79.
- [12] Egan, D. Individual differences in human-computer interaction. In M. Helander, ed., *Handbook of Human Computer interaction*. Elsevier Science Publishers, Amsterdam, The Netherlands, 1988, 543–568.
- [13] Evans, C.J., Kirby, J.R., and Fabrigar, L.R. Approaches to learning, need for cognition, and strategic flexibility among university students. *The British journal of educational psychology* 73, Pt 4 (2003), 507–28.
- [14] Freudenthal, D. Age differences in the performance of information retrieval tasks. *Behaviour & Information Technology* 20, 1 (2001), 9–22.
- [15] Gelman, A., Hill, J., and Yajima, M. Why We (Usually) Don't Have to Worry About Multiple Comparisons. *Journal of Research on Educational Effectiveness* 5, 2 (2012), 189–211.
- [16] Gueorguieva, R. and Krystal, J.H. Move Over ANOVA. *Archives of General Psychiatry* 61, (2004), 310–317.
- [17] Hadfield, J. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software* 33, 2 (2010), 1–22.
- [18] Harbich, S. and Hassenzahl, M. Beyond Task Completion in the Workplace: Execute, Engage, Evolve, Expand. In C. Peter and R. Beale, eds., *Affect and Emotion in Human-Computer Interaction*. Springer, 2008, 154–162.
- [19] Hassenzahl, M. and Monk, A. The inference of perceived usability from beauty. *Human-Computer Interaction* 25, 3 (2010), 235–260.

- [20] Hassenzahl, M. and Ullrich, D. To do or not to do: Differences in user experience and retrospective judgments depending on the presence or absence of instrumental goals. *Interacting with Computers* 19, 4 (2007), 429–437.
- [21] Hassenzahl, M. The interplay of beauty, goodness, and usability in interactive products. *Human-Computer Interaction* 19, (2004), 319–349.
- [22] Hornbæk, K. Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies* 64, 2 (2006), 79–102.
- [23] International Organisation for Standardisation. *ISO 9241-11:1998 Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability*. 1998.
- [24] Kliegl, R., Masson, M.E.J., and Richter, E.M. A linear mixed model analysis of masked repetition priming. *Visual Cognition* 18, 5 (2010), 655–681.
- [25] Lucas, R.E. and Baird, B.M. Global Self-Assessment. In M. Eid and E. Diener, eds., *Handbook of multimethod measurement in psychology*. Washington, DC, US, 2005, 29–42.
- [26] Macleod, C.M. Research on the Stroop Effect : An Integrative Review. *Psychological Bulletin* 109, 2 (1991), 163–203.
- [27] Porat, T. and Tractinsky, N. It's a Pleasure Buying Here : The Effects of Web-Store Design on Consumers ' Emotions and Attitudes. *Human Computer Interaction*, September (2012), 37–41.
- [28] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2011.
- [29] Richard, L. and Charbonneau, D. An introduction to E-Prime. *Tutorials in Quantitative Methods for Psychology* 5, 2 (2009), 68–76.
- [30] Robinson, M.D. and Neighbors, C. Catching the mind in action: Implicit methods in personality research and assessment. In M. Eid and E. Diener, eds., *Handbook of multimethod measurement in psychology*. APA American Psychological Association, Washington, DC, US, 2005, 115–125.
- [31] Ruiter, R.A.C., Verplanken, B., De Cremer, D., and Kok, G. Danger and Fear Control in Response to Fear Appeals: The Role of Need for Cognition. *Basic and Applied Social Psychology* 26, 1 (2004), 13–24.
- [32] Schwarz, N. Self-reports: How the questions shape the answers. *American Psychologist* 54, 2 (1999), 93–105.
- [33] Schwarz, N. Metacognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology* 14, 4 (2004), 332–348.
- [34] Sparrow, B., Liu, J., and Wegner, D.M. Google effects on memory: cognitive consequences of having information at our fingertips. *Science (New York, N.Y.)* 333, 6043 (2011), 776–8.
- [35] Tractinsky, N., Cokhavi, a, Kirschenbaum, M., and Sharfi, T. Evaluating the consistency of immediate aesthetic perceptions of web pages. *International Journal of Human-Computer Studies* 64, 11 (2006), 1071–1083.
- [36] Tulving, E. and Schacter, D.L. Priming and human memory systems. *Science* 247, 4940 (1990), 301–6.
- [37] Zuur, A.F., Ieno, E.N., and Elphick, C.S. A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution* 1, 1 (2010), 3–14.