Applying the Implicit Association Test for the Evaluation of Persuasive Technology - Implicitly Measuring Attitudes and Attitude Changes

Susen Döbelt[1], Johann Schrammel[1], Özge Subasi[1] and Manfred Tscheligi[1, 2]

[1] CURE – Center for Usability Research & Engineering,
Modecenterstraße 17 / Building 2, 1110 Vienna, Austria
[2] ICT&S Center, University of Salzburg, S.-Haffner-Gasse 18, 5020 Salzburg, Austria

Abstract. The effects of persuasive technologies are typically evaluated by means of behavioural indicators; the similarly important effects on the users' attitudes often remain unevaluated. In this paper we propose the use of the Implicit Association Test (IAT) as a method to reliably evaluate the effects of persuasive technologies on attitudes/attitude changes. Results showed that two tested distinctive user groups (usability experts and software developers) differed as expected concerning their IAT score. Based on results of reliability and validity tests we suggest the application of IAT in a persuasive technology context and as a promising method to evaluate the short- and long-term impact of persuasive technology tools.

Keywords: persuasive technology evaluation, Implicit Association Test, attitude measurement

1 Related Work

Incorporating persuasive strategies into technology may help achieve desired changes in human behavior and attitude. Due to the fact that attitude and attitude change is difficult to study there is an over-representation of studies dealing with effects on behavior [9]. The question of how to evaluate persuasive technology on the level of attitude change arises.

Exemplary studies dealing with ecological behaviour-enhancing persuasive technology (e.g. [1], [2]) reported desirable changes of behavioral indicators, but their use of questionnaires led to an inability to demonstrate similar changes in attitudes. The authors themselves also mentioned problems concerning explicit awareness of attitudes. These limitations became apparent during our own work on usability-enhancing persuasive tools as well. The Implicit Association Test (IAT) was developed to address the problems for which commonly-used questionnaires and rating scales have been criticized [8]. The IAT is typically applied in fields of research related to prejudices (e.g. [5], [6]). It provides information about individual differences in the automatic activation of evaluative and semantic associations. In a first IAT study [3] dealing with the association between locative pictures and trust on websites, the authors highlighted the need for development and adaptation of appropriate experimental methods and suggest methods from social sciences to provide deeper insights. In our study we address the measurement of attitude towards

usability, which, as with the subject of trust, is difficult to assess. We suggest an implicit measurement of attitude as an appropriate method and constructed an IAT aiming at the evaluation of usability-enhancing persuasive tools.

2 Constructing an IAT for measuring attitude towards usability

The IAT assesses subjects' comparative attitudes (i.e. implicit preferences) toward two target categories by comparing reaction times to different category-attribute pairings. Subjects have to allocate target items presented in the middle of a screen as fast as possible by pressing a key ("e" or "i"), depending on whether the item belongs to the permanently presented category (attitude object), attribute (good or bad) or category-attribute pairing on the left or on the right. This resulted in final IAT score range from 2.0 to -2.0 indicating an attitude towards either the first or second concept-attitude pairing. The IAT score considers reaction times of *Concept-attribute pairing I*, in which one concept-attribute pair is presented at each side in the top of the screen, and *Concept-attitude pairing II*, in which the attributes switch sides while the concepts maintain their positions. Calculation of reaction times included reaction times of 120 trials (60 per block). Response times faster than 300ms and slower than 10000ms and subjects with >10% of error trials were excluded [8].

Using a licensed but individually configurable software product (Inquisit by www.millisecond.com), we created an IAT to measure the association level of concepts *usability* and *programming* that are relevant to our case. Experts were asked to allocate common usability and programming terms into the two concept categories; the items that were allocated unanimously were then used in our IAT. The concept of *usability* thus included the terms: accessibility, user friendliness, human computer interaction, user centered design, user ergonomics, usability engineering and usability; the concept of *programming* included the terms: software architecture, object oriented programming, efficient algorithms, data structures, class hierarchies, compiler and programming environment. IAT standard attribute items for *good* (fabulous, excellent, nice, joyful, wonderful, great and terrific) and *bad* (tragic, cruel, annoying, horrible, awful, embarrassing and ugly) were used.

3 Validation of the IAT

Altogether 17 study participants took part in the validation. Two different groups were assembled according to the participants' professions: six usability experts, working for a usability consultancy company and eleven software developers, working in various machine building companies. First, to enable analysis of test-retest reliability, participants were asked to complete a full set of IAT trials twice immediately after each other; the short retest interval was chosen to avoid situational influence on the activation of evaluative and semantic associations. Second, participants had to provide information about their concern during a typical project by locating themselves on a continuum between usability and programming pole and subjects' marks were converted into numerical data. Third, the participants completed

a specifically constructed questionnaire aimed at measuring the subjects' personal preference with regard to usability and programming. The subjects could rate their agreement to seven statements indicating either a usability- or programming-oriented attitude on a five stage agreement scale (from “I totally disagree” to “I totally agree”). Cronbach’s alpha was used to analyze internal consistency of the two questionnaire scales. Nondistinctive items were excluded. The remaining item set of the questionnaire showed a Cronbach’s α of .784 for the usability and .679 for the programming scale.

The two IAT scores of all subjects resulted in a test-retest correlation of $r = .506$, which can be considered as satisfying. This indicates that the IAT delivers stable results. To investigate convergent validity of the IAT we correlated the mean IAT score of all subjects with the questionnaire data. The questionnaire score showed a highly significant correlation ($r = .740$) and the overall self-assessment scale a significant correlation ($r = .579$) with the IAT score. Overall the self-assessment scale and questionnaire showed a significant correlation ($r = .559$). These results show that the IAT, the questionnaire and the self-assessment scale measure similar concepts.

Comparing the average reaction latencies between the tested user groups, usability experts show faster reactions (M = 823.60ms, SD = 199.70 ms) when *usability* and *good* are paired compared to the *usability* and *bad* pairing (M = 1083.50ms, SD = 200.66ms). This resulted in a mean IAT score of D = -0.6299 (SD = 0.2983) indicating a strong positive attitude towards the concept *usability* compared to *programming*. Software developers showed slight faster reactions for the combination *usability* and *good* (M = 1048.97ms, SD= 184.57ms) compared to the combination *usability* and *bad* (M = 1194.98ms, SD = 255.07ms). This resulted in a mean IAT score of D = -0.1194 (SD = 0.1862) which reflects a slightly positive attitude towards *usability* compared to *programming* for software developers.

We used a t-test for independent samples to calculate differences between usability experts and software developers concerning their attitude towards *usability* vs. *programming.* Welch-corrected for unequal variance (Levene p = .033), we found a significant difference between the mean IAT scores of the two groups (p = .006) on .05 level. Additionally questionnaire data showed highly significant differences (p = .001). Groups did not differ in their overall self-assessment scale.

4 Conclusion

In this paper we suggest an implicit method for measuring attitudes as a new approach to evaluate the impact of persuasive technology tools especially in sensitive contexts, when people are not willing to say what’s on their mind or it is not accessible for them. We presented first results of tests as well as results of validation and reliability analysis of our constructed IAT measuring attitude towards *usability* vs. *programming*.

IAT scores of the two tested groups (usability experts and software developers) are statistically different from each other. This is supported by the results of self-assessment data which showed differences between usability experts and software developers’ attitude as well. Hence, we conclude that the constructed IAT is an

appropriate measurement to reflect and assess attitude towards usability. The test-retest reliability of the constructed IAT found in this study could be considered as satisfying based on results known from other studies with a mean around r = .56 (for an overview see [7]). Furthermore, correlations with explicit self-assessment measurements turned out to be significant, which indicates that our constructed IAT is a valid method for measuring usability-oriented mindsets.

However, there are limitations concerning the convergent validation relating to the fact that the constructed IAT showed correlations with constructed self-assessment measurements. These are commonly-used measurements to evaluate attitude change, but we aim to include behavioral measurements as well in order to validate the constructed IAT. In addition to the evaluation of subjects' current attitudes, we aim to conduct further longitudinal studies to evaluate long term-effects of persuasive tools by means of implicit measurements. For this purpose we think that additional to an explicit attitude measurement, implicit measurements could help to "develop a deeper understanding of how computing systems can be designed to change attitudes and behaviors" [4].

References

1. Froehlich, J., Dillahunt, T., Klasnja, P., Mankoff, J., Consolvo, S., Harrison, B., Landay, J. A.: UbiGreen: investigating a mobile tool for tracking and supporting green transportation habits. In Proc. of the 27th international Conference on Human Factors in Computing Systems, Boston, ,USA, ACM, New York, 1043--1052 (2009)
2. Shiraishi, M., Washio, Y., Takayama, C., Lehdonvirta, V., Kimura, H., Nakajima, T.: Using individual, social and economic persuasion techniques to reduce CO2 emissions in a family setting. In Proc. of the 4th International Conference on Persuasive Technology, ACM New York (2009)
3. Kostakos, V., Oakley, I.: Designing Trustworthy Situatied Services: an Implicit and Explicit Assessment of Locative Images`Effect on Trust. In Proc. of the 27th international Conference on Human Factors in Computing Systems CHI '09, Boston, USA. ACM, New York, 329--332 (2009)
4. Fogg, B.J.: Persuasive Technology. Using Computers to Change What We Think and Do. Morgan Kaufman Publishers (2002)
5. Kühnen, U., Schießl, U., Bauer, N., Paulig, N., Pöhlmann, N., Schmidthals, N.: How robust is the IAT? Measuring and manipulating implicit attitudes of East- and West-Germans. Zeitschrift für Experimentelle Psychologie, 48, 135--144 (2001)
6. Rudman, L. A., Greenwald, A. G., Mellott, D. S., Schwartz, J. L. K.: Measuring the automatic components of prejudice: Flexibility and generality of the Implicit Association Test, Social Cognition, 17, 437--465 (1999)
7. Nosek, B. A., Greewald, G. A., Banaji, M. R.: The implicit Association Test at Age 7: A Methodological and Conceptual Review (2006)
8. Greenwald, A. G., Nosek, B. A., Banaji, M. R.: Understanding and Using the Implicit Association Test I: an improved scoring algorithm. Journal of Social Psychology, 85, 197--216 (2003)
9. Torning, K., Oinas-Kukkonen, H.: Persuasive system design: state of the art and future directions. In Proc. of the 4th international Conference on Persuasive Technology, Persuasive '09, Vol. 350. ACM, New York, NY, 1--8 (2009)