Development of Attractiveness Walkthrough

Aliaksei Miniukovich

University of Trento, Trento, Italy
miniukovich@disi.unitn.it

Abstract. In this paper, I describe the development of Attractiveness Walkthrough (AW), a new expert-based technique of evaluating attractiveness of graphical user interface (GUI). I detail an underlying theoretical model, translation of the model implications into an evaluation protocol and planned empirical studies of the model and protocol. GUI designers and engineers may use AW for iterative GUI improvement and as a measure of GUI quality.

Keywords. Attractive user interfaces, expert-based evaluation, user experience.

1 Introduction

Despite more than a decade of intensive attractiveness-related research in HCI, we still do not know how to make an interface attractive. Mechanical summarization of preceding findings in simple laws of attractiveness does not work; thus, the guideline "*colour use should be balanced and low saturation pastel colours should be used for backgrounds*" [8] is simple but obviously contradicts the finding "*10 images receiving the lowest ratings in all three dimensions predominantly had a white background whereas colors of the 10 highest rated were all quite bright and saturated*" [4]. Apparently, a simple solution does not exist: numerous contextual and cultural factors, and ever-evolving user values make the "one solution for everybody" scenario unrealistic. A potential way out is a method that would define how to define what is attractive in which circumstances and for what audience.

My work aims at developing Attractiveness Walkthrough, an expert-based method of GUI evaluation that allows designers (especially, non-professional designers) to find specific attractiveness issues and to propose design solutions cheaper than doing user studies. This work involves three types of activities. First, I combine relevant theoretical findings from HCI, psychology and consumer behavior research in a theoretical model of how people appreciate things. Second, I translate the implications of the model into a set of instructions and questions, which is the protocol of AW. Finally, I test out the model, and then, the protocol, gather and analyze feedback, and improve the protocol.

2 Progress to date

In the beginning of my PhD, I assumed that a structured expert-based technique for UI attractiveness evaluation might satisfy the demand for a cheap, valid and reliable method of validating designers' work. This technique – Attractiveness Walkthrough (AW) – is based on a theoretical model of appreciating things (Figure 1); the protocol of AW resembles the protocol of Cognitive Walkthrough [6]. The scope of AW is limited to GUIs and first-time users, initial impressions of whom AW tries to model.

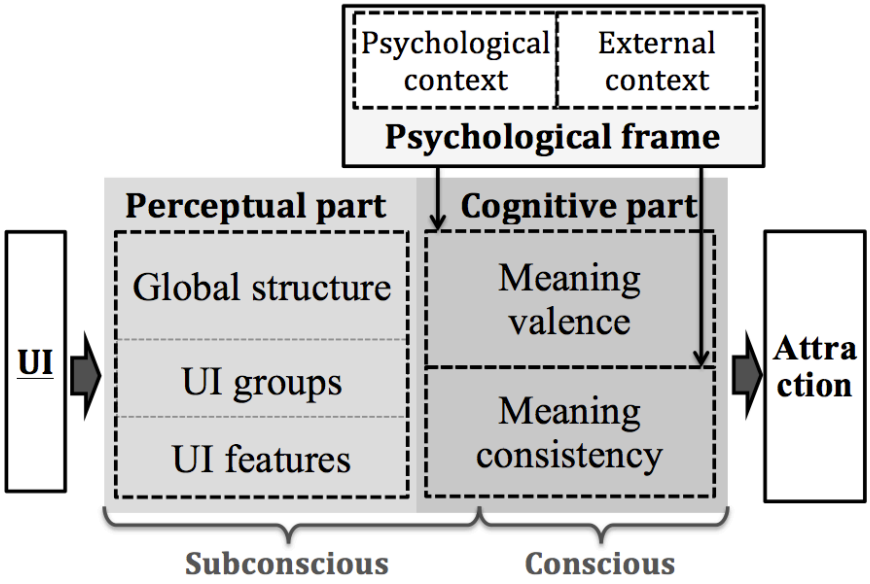

Fig. 1. The theoretical model of AW.

2.1 Theoretical Model

Existing theoretical models do not suit AW. HCI-originated models often concentrate on UX rather than on attractiveness (e.g., [1]), whereas psychology-originated models address art stimuli predominantly (e.g., [3]). I reviewed psychology, HCI and consumer behavior literature and synthetized relevant findings into a theoretical model (Figure 1) that AW is based on.

This model aims at simulating user thinking that may result in approach behavior (e.g., higher liking ratings and intention to buy), which is a manifestation of being attracted. Thus, the model utilizes product attractiveness as inclination towards a product due to pleasing effects of forming a positive mental image of the product. Immediacy, subjectivity, impermanence and automatism are, thus, the characteristics of attractiveness.

Perceptual Part. Perceptual mental processes are subconscious and very fast. They occur when human brain extracts distinct features from the scene (e.g., areas and lines colored differently), groups these features (e.g., same-color dots that belong to the same object) and perceives the scene as a whole (e.g., detects symmetry, which is the characteristic of the entire scene). Facilitating any of these processes leads to faster, effort-free and error-free perception of stimulus, which also leads to mild pleasure and higher liking ratings [7].

Cognitive Part. Given the focus of AW on initial impressions, I have included in the model only fast and automatic cognitive processes relevant to appreciation. The first type of processes – experiencing meaning consistency – leverages the notions of fitting input information within existing knowledge, i.e. obtaining "cognitive mastery" over the stimulus [3] and "feeling of control" of the stimulus [5]. While obtaining cognitive mastery, users might consciously experience salient, affective GUI messages, and, if the messages are consistent with the context and among themselves, they will grasp them easier and favor the GUI. Contrary to cognitive mastery, feeling of control is primarily unconscious and based on non-salient, non-affective GUI messages (e.g., "this colored area is a button"), which still should match with users' existing knowledge, so the user understands the GUI and favors it.

The second type of relevant cognitive processes – experiencing meaning valence – deals with re-feeling past painful or pleasing experiences, which a GUI triggers. The positive/negative affect of these experiences is often associated with the GUI as a whole (this is known in psychology as the halo effect). Consequently, featuring positive messages, which trigger positive associations, contributes to GUI liking.

Psychological Frame. Psychological frame can be seen as the projection of context on the user mind; it represents all semantic concepts that have been framed (i.e., preactivated) prior to interaction. Thus, all contextual factors (i.e., the task at hand or place of interaction) influence attractiveness, but indirectly, through the psychological frame. The components of psychological frame may resemble the components of context (cf. [2]): psychological context (past experiences), external context (interaction environment) and incidental context (incidental events). The model (Figure 1) includes only psychological and external contexts since incidental context is impossible to predict, and therefore, is of little practical value.

Psychological frame strongly influences only cognitive mental processes: user values and expectations may influence what is seen as positive, whereas the (mis)match between framed concepts and GUI-featured messages influence the feel of consistency. The influence of psychological frame on perceptual mental processes is dismissively small.

2.2 Evaluation Protocol

Any theoretical model is difficult to apply unless it is translated into a set of practical instructions. This section describes such a translation based on the protocol of Cognitive Walkthrough [6].

Prerequisites. Sufficient complexity and element diversity are the main prerequisite: overly monotonous or overly simple GUIs have a very few aspects to evaluate, the impact of design on their appreciation will be largely influenced by incidental factors and indistinguishable from the statistical error.

Actors. Conducting AW requires a pool of evaluators (3 to 5 people, cf. [6]) familiar with the project and designing: HCI, user factors, usability and ergonomics specialists, software engineers and, of course, designers are qualified. Using a pool of evaluators instead of one evaluator might increase the reliability and validity of evaluation results.

Procedure. Similar to Cognitive Walkthrough, AW consists of two main phases: preparation and evaluation.

Preparation. Prior to evaluation, one specialist does one of two preparation activities: describes the psychological frame of user, or partitions a GUI into basic and group elements. In the beginning of each evaluation session, the rest of specialists familiarize themselves with collected artifacts and verify their correctness.

For describing the psychological frame, a specialist familiar with project goals – desirably an ethnographer or human factor specialist – leverages his/her experience and external data sources (e.g., surveys done before the project started) to create plain-text descriptions of the target user and typical usage situation. Then, the plain-text descriptions are compressed into two lists of concise points. While compressing, the specialist may leverage the questions "What characteristics differentiate an average *target* user from an average user", "What aspects of interaction time and location are highly specific for the developed tool" and "What high-level functional intentions the target user has". Lastly, based on the lists, the specialist infers potential user expectations regarding interaction and GUI.

A specialist familiar with creating GUIs does the second preparation activity – partitioning a GUI into basic elements and group elements. Basic elements of GUI are separate semantic or functional entities that, oftentimes, are well-known and conventional (e.g., buttons, labels and icons). Group elements of GUI consist of several spatially-proximal basic elements that share one or several characteristics (e.g., color, shape or font size), e.g. a menu with several buttons. After basic and group elements are all extracted, the specialist finds similar looking elements and leaves only one of them. The one-by-one analysis of the rest may well not result in more attractiveness issues found: similar elements have similar issues.

Evaluation. Each evaluation session consists of four steps. First, evaluators check how fluently the user can perceive an interface: the easier, more fluently an individual is capable to make sense out of raw sensorial data, the more pleasing is this experience [7]. Accordingly, people often prefer simple and prototypical stimuli. Evaluators check simplicity and prototypicality by answering a set of questions (Table 1) for basic and group GUI elements, and entire layout. This evaluation step corresponds to the entire perceptual part of the AW model (which consists of experiencing GUI features, GUI groups and entire layout).

Second, evaluators check how easily the user can understand a GUI. GUIs often consist of many functional elements that do not trigger strong associations in users, but still carry meanings and are recognized. Thus, understanding a GUI is a cognitive process, still primarily subconscious, which corresponds to Norman's "feeling of control" [5]. Evaluators check this parameter by answering questions from Table 2. This evaluation step corresponds to the Meaning Consistency part of the AW model.

Table 1. Questions for evaluating perceptual complexity of UI

UI aspect	Question
Basic elements	How easy to distinguish from background and other elements is the element? How prototypical is the element? How easy to read is the text within the element? (if applicable)
Group elements	How salient is the pattern or symmetry of the group element? How many characteristics do the sub-elements of group element share?
Entire UI layout	How salient is the pattern or symmetry of the layout? How prototypical is the layout?

Table 2. Questions for evaluating easiness-to-understand of UI

UI aspect	Question
Basic elements and group elements	How obvious to the user is the type of the element? How obvious to the user is what functionality the element has?
Group elements only	How many sub elements does the group element have? (*3 to 9 is the optimum amount*) How confusing to the user is the mismatch between the actual group element and what is seen as a group?

Third, evaluators check how consistent with the psychological frame and among themselves GUI-featured messages are. GUIs often feature several messages, which might relate to prevailing colors, non-prototypical GUI elements, images, GUI general stylistic impression, company brand and main headlines. Evaluators outline these, most-salient messages, relate them to the concepts of psychological frame (outlined

prior to evaluation), and single out the messages that are not or weakly connected. These messages are inconsistent with the psychological frame; they provoke confusion in users, and thus, diminish GUI appreciation. This evaluation step corresponds to the Meaning Consistency part of the AW model.

Finally, evaluators check how emotionally positive or negative GUI-featured messages are. GUIs often exhibit emotionally affective messages, which, due to the halo effect, affect general GUI appreciation. Consequently, exhibiting positive messages is what a GUI should do. Evaluators look through the messages they outlined and split ambiguous messages into two, e.g. the "Facebook-like" becomes "Facebook" (positive) and "plagiarism" (negative). Then, they rate the valence of each message based on common sense and user expectations (which are a part of psychological frame outlined prior to interaction), and thus, single out negative messages that detriment GUI attractiveness. This final evaluation step corresponds to the Meaning Valence part of the AW model.

Output. Besides setting scores (quantitative data), evaluators may note down prospective design solutions and ideas, which they came across during evaluation (qualitative data). Based on this, evaluators can produce reports where they substantiate which GUI aspects should be changed and why, and how this can be done.

3 User Studies

I plan on conducting three user studies with computer science students as participants and website screenshots as stimuli. Computer science students, after they are introduced to AW, will have necessary background for doing AW tryouts, and so they will be used. Website screenshots will be used since websites are easily accessible and are diverse enough (e.g., on their visual complexity and meaningfulness). Although, other HCI-related stimuli with rich GUI, such as mp3 players or mobile apps, could also be used.

3.1 Model validation

The first user study will test out the underlying model of AW. Based on the model, I have inferred four dimensions of GUI (which also corresponds to four steps of AW protocol), which might contribute to perceived GUI attractiveness and which will be tested. These dimensions are perceived visual complexity, perceived understandability, perceived meaningfulness (context compliance) and perceived emotional valence of GUI. Please note the usage of perceived parameters (as opposite to objective parameters): attractiveness, as it is defined in this paper, is a subjective quality of GUI and is based on perceived (subjective) parameters.

In the study, participants will read short stories (descriptions of the user and usage situation) and imagine the stories happen to them, which is needed for forming similar psychological frames across participants. Then, the participants will be looking at each story-related website screenshot for four seconds and will rate it on four afore-

mentioned dimensions and attractiveness. I expect the ratings on four dimensions to strongly correlate with the ratings of attractiveness, which will support the theoretical model. These user ratings will also be used in the third study.

3.2 Protocol refinement and validation

The second user study will refine the AW protocol; it will explore which most-salient aspects of GUI users attend to. The third step of AW protocol – assessing the consistency of GUI-featured messages with psychological frame – requires knowing what GUI aspects can trigger an association in user, i.e. what aspects convey salient-enough messages. Currently, the list of these GUI aspects (see the 3rd paragraph of *Evaluation* section) is solely based on a literature review and should be supported empirically. In this study, I plan on asking participants to read short stories (the same as in the first study), imagine the stories happen to them, and then, to look at website screenshots for 60 seconds and note down their immediate most-salient one-word associations with the website. The analysis of these associations will let me extract classes of attention-attracting GUI features and augment my current literature-based list of such features.

The third user study will test out AW. I plan on presenting AW to several groups of evaluators (students taking an HCI course). After this, I will ask them to conduct AWs on several website screenshots, produce reports where they outline the most serious attractiveness issues found and to fill out post-evaluation satisfaction questionnaires. Then, I will compare the amount and quality of evaluator-found attractiveness issues with the user-set ratings of website attractiveness (obtained while pre-selecting screenshots), and thus, will measure the validity of AW. Further, I will compare attractiveness issues found by different teams of evaluators, and thus, will measure the reliability of AW. Then, I will correlate the number of attractiveness issues by each step of AW (the AW protocol has four steps) with screenshot ratings on related dimensions (visual complexity, understandability, context compliance and emotional valence – obtained in the first study). Weak correlations will point out the parts of AW that should be improved. Finally, the analysis of post-evaluation questionnaires will let me improve the wording of AW questions and instructions.

4 Conclusion

The development of AW is not finished. I have combined relevant findings from literature into the AW underlying model and proposed the AW evaluation protocol. However, empirical studies should validate the model and protocol, and, currently, I am developing such studies. Feedback from the CHItaly community could let me improve AW cheaper than doing an additional test-improve round, and therefore, I will appreciate any feedback regarding the theoretical model and protocol of AW, and design of user studies.

Acknowledgements. I would like to thank Antonella de Angeli and Sylvie Noel for their invaluable help in refining the ideas described in this paper. I am also grateful to anonymous reviewers for their helpful comments on an earlier draft of this paper.

References

1. Hassenzahl, M. The thing and I: understanding the relationship between user and product. *Funology* (2005), 31-42.
2. Kleine III, R. E., & Kernan, J. B. Contextual influences on the meanings ascribed to ordinary consumption objects. *Journal of Consumer Research* (1991), 311-324.
3. Leder, H., Belke, B., Oeberst, A., & Augustin, D. A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, *95*, 4 (2004), 489-508.
4. Lindgaard, G., Dudek, C., Sen, D., Sumegi, L., & Noonan, P. An exploration of relations between visual appeal, trustworthiness and perceived usability of homepages. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *18*, 1 (2011).
5. Norman, D. A., & Ortony, A. Designers and users: Two perspectives on emotion and design. In *Proc. of the Symposium on Foundations of Interaction Design at the Interaction Design Institute* (2003).
6. Polson, P. G., Lewis, C., Rieman, J., & Wharton, C. Cognitive walkthroughs: a method for theory-based evaluation of user interfaces. *International Journal of man-machine studies*, *36*, 5 (1992), 741-773.
7. Reber, R., Schwarz, N., & Winkielman, P. Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, *8*, 4 (2004), 364-382.
8. Sutcliffe, A. Assessing the reliability of heuristic evaluation for Web site attractiveness and usability. *In proc. of the 35th Annual Hawaii International Conference on System Sciences* (2002), 1838-1847.