

To know or not to know – The Impact of Sharing the Purpose of a Prototyping Session with the Participants

Christian Köppe
Hogeschool Utrecht
christian.koppe@hu.nl

Davy Meers
University of Amsterdam
contact@davymeers.com

Arie van der Veeck
University of Amsterdam
arie@arievanderveek.com

Jouke Stoel
Ordina N.V.
jouke.stoel@ordina.nl

ABSTRACT

Prototyping is a well known technique in software engineering. According to the literature, the successful application of prototyping depends on sharing the purpose of the session with the participants upfront. However, no empirical data is offered to support this claim. In our experiment we measured if the amount of faults found per participant depended on sharing the purpose of the prototyping session upfront. The results show that in our experimental context the claim is not supported.

Keywords

Prototyping session preparation, usability testing

Introduction

It is commonly accepted that prototyping is a successful technique in software engineering which can be used for different purposes related mainly to requirements elicitation, requirements validation, software design, and testing [1, 4, 5]. Much research is done on prototyping techniques. While the discussed techniques, their applicability and their forms differ broadly, it can be often found that authors claim that it is critical to the success of a prototyping session to prepare the audience in advance [1, 4, 16]. This includes the communication of the specific purposes of a prototyping session, especially the question of what is and what is not addressed by the prototype [1, 2, 4].

While it seems intuitively right to claim this, it is remarkable that none of the authors offer reliable studies for proving this claim to be correct. It is also not described if this claim applies to all kinds of prototyping sessions, regardless of other factors which probably influence the results of such a session.

Our work will be part of filling that gap, which helps in answering the main research question:

How should a prototyping session be set up in order to improve its success?

This paper is organized as follows: we first look at the theory regarding the preparation of a prototyping workshop and some principles which influence its success. Then we formulate two hypotheses and describe how our experiment is set up. After the hypotheses and setup an analysis of the experiment results is given. Finally our conclusions are presented and discussed.

Theory

Preparation of a Session

"Understand your audience and intent" is one of the guiding principles for preparing a workshop [16]. The audience of a prototyping session has to be chosen carefully. Real (potential) users should be preferred to representative users [1].

An important part of pre-work is establishing a shared purpose (or intent) for the workshop [2]. All participants have to know the purpose of the workshop before it begins in order to be effective [4]. However, this *shared purpose* actually consists of two parts (although not mentioned explicitly in other work): the defined purpose itself and the sharing of it with the participants.

Knowing the intent of a workshop helps in determining the other important parts of the workshop: which scenarios are used, which research questions should be answered and what kind of prototype is used for the workshop [16].

Facilitation of a Session

The facilitator should be sitting besides the participant and work together with the participant to uncover usability problems, this approach is also termed *co-discovery* [3]. The facilitator helps the participant to produce the appropriate and desired data. It is really important here that the facilitator does not lead the participant so much that he or she interferes with the purpose of the session.

Priming

Priming is a psychological technique used to set expectations of an audience and therefore guiding their

attention and focus [16]. People give certain responses when they are prepared, primed, with certain information. The implicit memory is responsible for these responses. A property of priming is that items are remembered best in the form in which they were originally encountered [17]. So according to this theory sharing a purpose of a prototyping session in advance should lead to better results.

Experiment Design

The Purpose of the Experiment

The purpose of our experiment is to measure the impact of knowing the purpose of a prototyping session on the number of faults that are found. Of particular interest is the number of found faults of the specific type of fault given in the purpose of the session. We think that knowing the purpose of a prototyping session in advance works as priming.

In our experiment we use following pairs of variables:

- independent variable: *sharing or not sharing the purpose in advance.*
- dependent variable: *more identified faults or less identified faults.*

and

- independent variable: *sharing or not sharing the purpose, which causes the participants to focus on a specific type of fault or not.*
- dependent variable: *more identified faults of the same type or less identified faults of a different type.*

The Prototype

A prototype of a job site was used in the experiment. The used prototype was a computer-based low-fidelity prototype that only had a fraction of the complete functionality and did not have the look-and-feel of a completed system.

The used prototype only contained functionality to add a résumé. Data is not persisted and it can be used repetitively.

Seeded Faults

Fault seeding is a technique mostly used in the field of software testing [8]. It is the process of intentionally adding known faults to those already in a computer program for the purpose of monitoring the rate of detection and removal, and estimating the number of faults remaining in the program [8].

A total of 18 faults were inserted into the prototype, ranging from obvious faults like tab sequence to less obvious faults like omission of field validation. To be able to distinguish between different types of faults a classification of faults was used. This classification was developed by looking at the different types of faults in

prototypes described in literature: usability faults [9,15], structure faults [1], and requirement faults [13].

The content validity was increased through introducing faults that violate best practices in UI-design, as described in [6, 14].

Hypotheses

Based on the previous sections we define our hypotheses as follows:

Sharing the purpose of a prototyping session upfront leads to more identified faults concerning the purpose.

and

Sharing the purpose will focus the participants on a specific type of fault, which leads to fewer identified faults of a different type.

Setup of the Experiment

We chose a formative evaluation for this experiment, as this is applicable for the purpose of our experiment. We are interested in the amount of seeded faults that the participants will find in the prototype. These data can be classified as subjective and qualitative data. The seeded faults are of all three types discussed in the *Seeded faults* section.

All participants were asked to perform tasks. These tasks were described in the form of a scenario. The scenario only specifies *what* the participant should do and not *how* this should be done, as suggested by [3]. All instructions given to the participants are in written form, including the scenario. This is to make sure that every participant got the instructions in exactly the same form.

We used two groups for this experiment. The selection of the participants for one of the groups was done through randomized selection, which prevents a potential *selection bias* [10].

Both groups received introductory instructional remarks including the procedure to be followed. The instructions for group one also included the purpose of the session, which was finding *usability faults* in the prototype. They were also given a list of examples of usability faults in a different kind of system. The instructions for the second group did not include the purpose.

The participants did not know that there were different groups. To make sure that the facilitator and observer also did not know a-priori which participant belonged to which group, the assignment of the participants to one of the groups was done outside of the workshop room by a third researcher. This helped preventing the so called *experimenter expectancy effect* [10] by making the facilitator and the observer *blind experimenters*. All participants were registered anonymously with a number.

One possible threat to internal validity is the plausible rival hypothesis that one of the facilitators influenced the group more than the other [10]. Therefore the facilitators

switched frequently so they facilitated a nearly equal amount of participants from both groups.

A single session lasted a maximum of fifteen minutes. At the end the participants had to fill in a short questionnaire to give some additional context. The questions on the questionnaire were divided into two categories. One set of questions was about the participants experience with job sites. The other set of questions was about the perceived reliability of the prototype.

During the experiment the technique of *structured observation* was used. Structured observation is a technique where explicitly formulated rules are employed for the observation and the recording of behavior. Rules were defined as to what the observer should look for, namely usability faults as well as structure and requirements faults. The recording was performed with a checklist, which included the seeded faults and room for note-taking regarding non-seeded faults. The observer also monitored if the facilitator did not lead the participants too much.

Participants

The participants in this experiment were 14 undergraduate students ranging aged between 18 and 20. They had experience with writing résumé's, had applied for part time jobs and had limited experience with job sites. They had no extensive experience with prototyping. Having the correct skills is important [3, 9]. Finding a job is an activity that almost all people do at least once in their lifetime, but it is not done as frequently as ordering a book from the internet. Therefore almost all people should have the correct skills to use a job site, if they have any experience with using the internet. To focus even more on the skills of the participants, only the scenario "Add a Résumé" was facilitated in the prototype.

As stated in [3], securing cooperation from all participants is important for the success of a prototyping session. Therefore only volunteers were selected for the user group, as it can be assumed that they are willing to cooperate. Students did not receive any compensation, therefore no bias was expected due to higher motivation through compensation [3].

Assessment of the remarks

During the prototyping sessions the observer had a checklist containing the seeded faults. When a participant noticed a particular fault the observer ticked the seeded fault on the list. The remarks that didn't relate to any of the seeded faults were written down.

All remarks were then judged by the researcher on relevancy. Almost every remark lead to a new identified fault. This fault could be new to the list or already have been remarked by another participant. Only remarks about visual elements of the prototype were ignored [11]. The newly found faults were also categorized by the researchers into the same categories as were used for the

seeded faults. Because all the identified faults were marked as relevant, the seeded and non-seeded faults were weighted equal.

Experiment Results

Existing faults in the prototype

Table 1 shows the distribution of the faults in the prototype. Originally we inserted 18 faults of which 13 were categorized as usability faults. After the assessment of the remarks made by the participants we added a total of 36 newly identified unique faults to the list. The majority of these faults, 31, were categorized as usability faults.

	Usability	Other types	Total
Seeded-faults	13	5	18
Non-seeded faults	31	5	36
Total	44	10	54

Table 1: Overview of unique faults in prototype, seeded and non-seeded

Faults per group

Figure 2 shows the average, minimum and maximum of found faults in both groups. Both groups found approximately the same number of faults. The group that knew the purpose found an average of 7,29 faults. The group that did not know the purpose found an average of 7,43 faults.



Figure 2: Number of found seeded and non-seeded faults

Faults per type

Figure 3 shows that a total of 41 usability faults were identified by the members of the group that knew the purpose versus a total of 42 found usability faults found by the group that did not know the purpose.

Discussion

The results found in the experiments do not support the hypotheses. A possible explanation can be that a high cognitive load was imposed on the participant during the experiment.

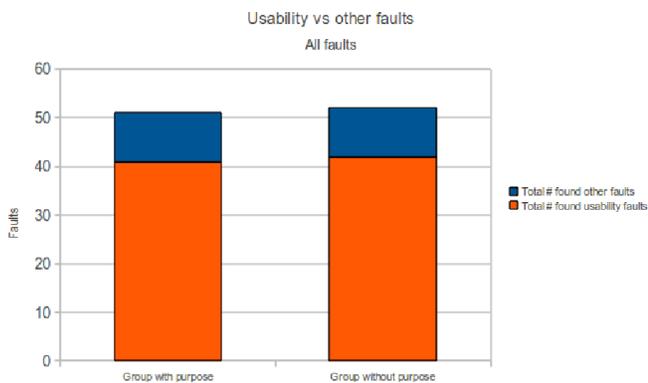


Figure 3: Number of found usability faults versus other faults

This high cognitive load can be caused by the combination of the instructions, the evaluation of the prototype and thinking aloud, the split-attention effect, the absence of extensive domain specific knowledge or the absence of extensive experience with prototyping [7, 12].

If the cognitive load exceeds the working memory resources, it is fatal to learning [7, 12]. This could explain the observations made during the prototyping sessions: most participants didn't really use the provided information and in case of frequently encountered problems schema automation occurred [12].

Threats to validity

The number of participants in this experiment was low, fourteen in total. This number is too low to make any well-grounded conclusions about the results.

The participants consisted of a specialized group, namely undergraduate students of software engineering. There is a chance that they look differently at software and user interfaces than users with a non-IT background.

Final words

In this article we investigated the effect of sharing the purpose of a prototyping session on the success of the prototyping session. Contrary to what we predicted sharing the purpose upfront did not lead to more identified faults concerning the purpose, nor did it lead to fewer identified faults of a different type. However, it is not clear how much other factors, in particular the cognitive load, have influenced the results, thus there is insufficient evidence to support or refute our hypotheses.

Future study should be directed to the influence of cognitive load when sharing the purpose.

References

[1] Hugh Beyer and Karen Holtzblatt. *Contextual design: defining customer-centered systems*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1998.

[2] Ellen Gottesdiener. *Requirements by collaboration: workshops for defining needs*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2002.

[3] Deborah Hix and H. Rex Hartson. *Developing user interfaces: ensuring usability through product & process*. John Wiley & Sons, Inc., New York, NY, USA, 1993.

[4] S. Houde and C. Hill. What do prototypes prototype? *Handbook of human-computer interaction*, 2:367–381, 1997.

[5] IEEE Computer Society. *Software Engineering Body of Knowledge (SWEBOOK)*. Angela Burgess, EUA, 2004.

[6] Jeff Johnson. *GUI Bloopers 2.0: Common User Interface Design Don'ts and Dos*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2nd edition, 2007.

[7] F. Paas, A. Renkl, and J. Sweller. Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38(1):1–4, 2003.

[8] J. Ratz. IEEE standard glossary of software engineering terminology. *IEEE Std 610121990*, 121990, 1990.

[9] M. Rettig. Prototyping for tiny fingers. *Communications of the ACM*, 37(4):21–27, 1994.

[10] R.L. Rosnow and R. Rosenthal. *Beginning behavioral research: A conceptual primer*. Prentice Hall Upper Saddle River, NJ, 2005.

[11] R. Sefelin, M. Tscheligi, and V. Giller. Paper prototyping-what is it good for? : a comparison of paper-and computer-based low-fidelity prototyping. In *CHI'03 extended abstracts on Human factors in computing systems*, pages 778–779. ACM, 2003.

[12] J. Sweller, J.J.G. Van Merriënboer, and F.G.W.C. Paas. Cognitive architecture and instructional design. *Educational psychology review*, 10(3):251–296, 1998.

[13] A. Timmerman. Prototyping, a stepping-stone to more successful prototyping. 2008.

[14] Martijn van Welie. Interaction design pattern library. <http://www.welie.com/patterns/index.php>. Retrieved November 21, 2010.

[15] D. Verhamme. Effective prototyping in embedded systems. 2009.

[16] T.Z. Warfel. *Prototyping: A Practitioner's Guide*. Rosenfeld Media, 2009.

[17] I.Q. Whishaw and B. Kolb. *Fundamentals of human neuropsychology*. Worth Publishers, 1995.