On effective prototype sessions for the e-book of the future

Cindy Berghuizen, Chiel Peters, Mary Gouseti,
Omar Pakker, Fabiën Tesselaar

December 5, 2013

1 Introduction

In this document we define our theory on prototyping (for e-books of the future). First we define aspects that can make a session typically bad or good. Then we provide two experiment setups to test these theories. Finally, we describe the experiment we conducted and present its results.

2 Theory

In this section multiple subjects are stated that will likely negatively impact a prototype session. These subjects are based on the theory from Kahneman's book, *Thinking fast and slow*.

2.1 Bad Target Group

The participants in the prototyping session are of vital importance. Their backgrounds will determine the kind of feedback that will come out of the session. Two categories of participants will most likely not give useful feedback. These are:

- Insiders. Participants that already have a stake in the product

- Fanboys/girls. Participants that already like the product, because of the company or any other particular reason.

The insiders are bad participants because they are biased. They can only observe the inside view. Kahneman states the exact problem on the inside view:

He's taking an inside view. He should forget about his own case and look for what happened in other cases.

In a prototype session all the cases of the inside view are probably already covered by the prototype; however the outside view will render cases that give useful feedback. Having only insiders in the prototype session will thus hurt your feedback results.

The second group of participants is fanboys/fangirls. These participants will most likely fall for a different kind of fallacy, namely answering the easier question. Instead of looking at the product from an unbiased perspective they will already like product. This is because, they replace the question *How much do I like this prototype?* with an easier question: *How much do I like the company or their earlier products?*. This based perspective will negatively influence their ability to find new cases or flaws in the prototype.

2.2   Cognitive Ease & Affect

"When you are in a state of cognitive ease, you are probably in a good mood, like what you see, believe what you hear, trust your intuitions, and feel that the current situation is comfortably familiar."

"The affect heuristic is an instance of substitution, in which the answer to an easy question (How do I feel about it?) serves as an answer to a much harder question (What do I think about it?)."

Setting an environment in which the subjects experience cognitive ease may lead them to be more receptive of the prototype. To accomplish that, the environment of the experiment will be comfortable and the participants may be offered food or listen to music or even asked to smile throughout the experiment. According to the first quote of this section, they will not be critical towards the prototype, they will accept what is suggested by the experiment and the prototype will seem familiar. Moreover, they will feel more comfortable using it than they would have. Finally, if they have to answer the questionnaire in the end of the experiment, they will have associated the memory of using the prototype with their good mood so they will be even more positive towards the prototype assuming they use the affect heuristic.

2.3   Deplete System 2

Ego depletion: *"if you have to force yourself to do something, you are less willing or less able to exert self-control when the next challenge comes around"* .

Ego depletion might be a way to let the participants easier accept the prototype as a good one. First let the participants do something that takes a lot of effort, watch a movie, try not to focus on something distracting etc. Then the prototype will be shown. If the only parts of the system shown are the parts that are good (as explained in 1.6) the participants will easier accept the system to be good overall and do not criticize it.

2.4 Anchoring

A bad prototype session would be to bias participants to like the prototype. One way this can be achieved is the *Anchoring as priming effect.* Before starting the prototype session the participants are shown fake results from a hypothetical prior prototype session. These results show that a separate group really like the prototype and on average gave very positive feedback on it. Although the new participants should not be affected by these results, the anchoring effect suggests that they are more likely to give more positive feedback as well. In the worst case the new participants will exactly match the fake results resulting in a completely useless prototype session.

2.5 Availability

"The availability heuristic, like other heuristics of judgment, substitutes one question for another: you wish to estimate the size of a category or the frequency of an event, but you report an impression of the ease with which instances come to mind."

Availability can be used to influence the results of the prototype session in a way that is desirable. For instance, the answer of a participant to the question whether she would use the prototype in her everyday life can be manipulated; if different situations that the participant could use an eTextbook, such as while traveling by train, or when moving to a summer house, are mentioned before the question, the answer would be more positive. Similarly, if situations that she could not use an eTextbook are mentioned before the question, the answer would be more negative. This would happen because the references before the question change the ease that examples of using eTextbooks in everyday life come to the mind of the participant.

2.6 Priming / Framing on the good features / WYSIATI

Based on the WYSIATI principle, if we only show the good parts, the features that work well and are easy to use, the participants will likely give more positive feedback. The parts of the prototype that do not work well or are forgotten will be completely ignored. Therefore the participants will not notice the faults and limits of the system and will believe that the software works perfectly.

3 Good prototyping

We believe a good prototype session has the following:

- Prototype session closely related to reality

- Intervene / let the participants experiment

- Priming creativity

3.1 Outcome

A good prototype session will provide you with useful information. The participants will experiment with the software / object, in the best case integrate it in their daily life for a little while. While experimenting and using the device they will find out what works well, what does not work, what they miss and what parts are not used at all. This way they can provide better feedback and critique. Also, a successful prototype will uncover the non-functional requirements. For example: the participants may like the system but you find out that they will not buy it. Now work can be done to find out why people won't buy the system so you improve these factors and make sure the system will be a success.

4 Suggested Experiments

4.1 Anchoring

Setup   In this experiment we will use anchoring to influence the results of the prototype session. As described in the theory section 2.4 the participants will be required to fill out a form which will start with a few general questions and then will ask how much they like each of the new features (on a scale 1 to 10) that will be mentioned in the requirements document. On these questions the averages of a *"hypothetical"* prior experiment will be stated as well.

To measure the anchoring effect we will hand out two versions of questionnaire. The questions will remain the same; however the *"hypothetical"* averages we provide along with the questions will differ. One questionnaire will contain a high anchor 'average' and the other one a low anchor 'average'. Then after the experiment we will measure the difference between the questionnaires to see if the anchoring had any effect on our participants.

Expected Outcome   If other factors, such as a small experiment group, don't affect the results of this experiment then the difference between the high anchor average and the low anchor average should be significant.

4.2 Depletion of System 2

As stated before: *"If you have to force yourself to do something, you are less willing or less able to exert self-control when the next challenge comes around"* . This experiment will focus on this idea and attempt to influence the results.

Setup   For this experiment, 2 questionnaires will be created and 2 groups will answer one of them. One of the questionnaires will contain a large amount of boring and obvious questions before the questions we actually care about. The other questionnaire will contain only the questions we want answers on. The idea behind this is that the group that has to answer a great amount of boring questions will get annoyed and start to force themselves to go on and answer the questionnaire in such a way that they get it over with as soon as possible.

In contrast, the other group will have just the core questions and therefore will not need to force themselves through a long and boring questionnaire.

Expected Outcomes We expect the group that will have to force themselves to finish the questionnaire to do it in a manner to get it done as fast as possible. Therefore we expect that they will be more likely to answer with the first possible option. In contrast we expect the group that will not have to force themselves to finish the questionnaire to think their answers more deeply and critically.

5 Anchoring Experiment

5.1 Setup

This is the experiment we conducted to test the anchoring theory. As described in sections 2.4 and 4.1, the participants are required to fill out a questionnaire in which they rate the presented features on a scale 1 to 10. This survey starts with a few lead-in questions and then moves on to the features, after the features there is a open text question for feedback and comments and the final question asks the participants to give their overall rating for the eTextbook. The questions display the averages of a 'hypothetical' prior experiment.

To measure the anchoring effect we created two questionnaires, both using the same questions. While the questions are the same, the averages for the questions are swapped; question 'x' is rated high in one and low in the other. This setup can be seen in Figure 1. After the experiment we measure the average and median of the answers for the high and the low one to see if the anchoring had any effect on our participants.

Our results show that the anchoring effect is obvious in many cases. Only one question has an inconclusive difference and the deviation cannot be contributed to the anchoring effect for sure.

5.2 Results

This experiment was intended to be an anchoring experiment. However, by checking the results we also saw other theories from Kahneman book put in practice. The results presented in the following sections correspond to 40 participants and the answers were collected from 4/12/2013 to 5/12/2013. The survey is still online and we continue gathering data.

5.2.1 Anchoring

The way to introduce the anchors in out survey was a challenge. Our goal was to present it without drawing a lot of attention on it, so a participant would not reject it. We introduced the survey by explaining that its goal is to test the results of a previous prototype session and we provided the average for every feature. We decided to show the current score on the side of each question because people are familiar with ratings since they are used a lot on the internet.

1 Introduction

Betekenis: publiek

E-textbooks are nothing new and have gained widespread acceptance under a specific audience. One of the many examples is the Ipad schools that are created in the Netherlands [7]. These schools are determined to run their full educational program on an electronic device. It would seem that the e-textbook has already developed and formed a place in society. However the current capabilities of the e-textbook are heavily underused. Most students still buy the original textbooks. The purpose of this document is to examinate the requirements for the future e-textbook. One that offers significant advantages over its printed counterpart.

3.

**Feature Click & Translate: By clicking on a word you will see the translation in a preset language. (see picture)
How high do you rate this feature?**

Current score: 9.1

1   10
○ ○ ○ ○ ○ ○ ○ ○ ○ ○

3.

**Feature Click & Translate: By clicking on a word you will see the translation in a preset language. (see picture)
How high do you rate this feature?**

Current score: 5.1

1   10
○ ○ ○ ○ ○ ○ ○ ○ ○ ○

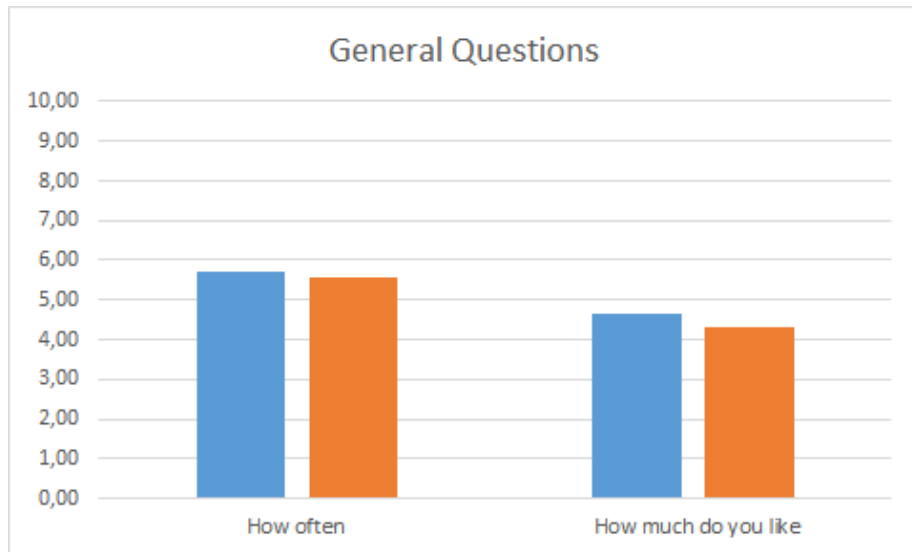Figure 1: The same question, different anchoring point per questionnaire.

Figure 2: The average of the answers to the general questions.

The choice of the anchors was tricky because we wanted them to be believable according to our intuition but we also wanted to preserve a distance between the anchors to get more noticeable results. We decided to approach this separately for each feature and play with different distances.

Although the features were the target questions of our anchoring experiments, in order to be consistent we used the anchors in the general questions too. In this case, we were more careful to provide more believable anchors in order to establish the trust of the participant to the anchors. In Figures 2 and 3 we can see the results of each survey and how the anchoring affects the results.

Anchoring Index. The anchoring effect is measured in percentages by dividing the differences between averages by the difference between the anchors and is called anchoring index. For example, the difference between the anchors 3 and 9 is 6 and the difference between the averages 5 and 7 is 2. Then, the anchoring effect is 2/6, or 30%. An anchoring effect of 100% means the subjects adapts the anchor point and an anchoring effect of 0% means the anchor has no effect on people.

In Figure 4 we see that there was almost no anchoring affect on the question about the dictionary feature (only 3.5%). This is probably because students already know the feature and like it. We will further discuss that in the following section.

The question about the interactive figures had a very strong anchoring effect, it was 53%. This is probably one of the examples that people in general do not have a high opinion of so they are easily influenced by the anchor.
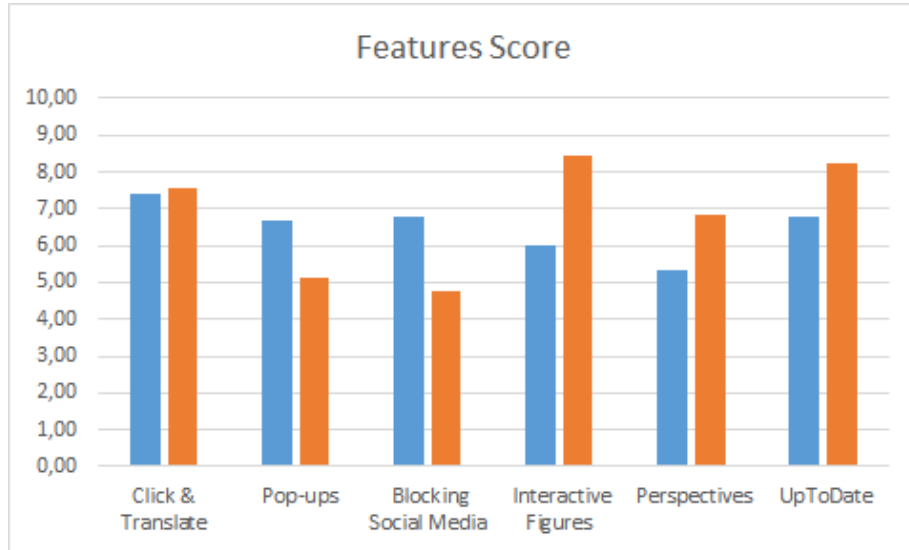
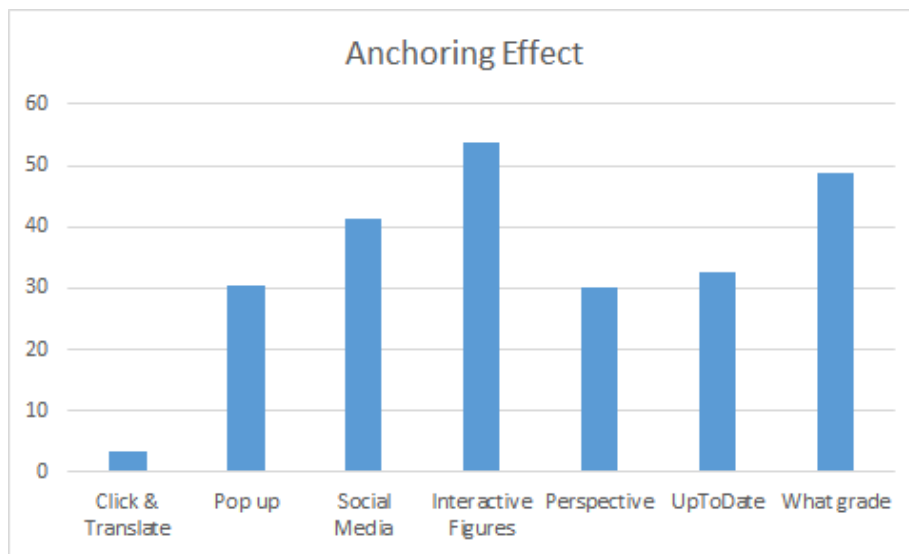Figure 3: The average of the answers to the feature questions.

Figure 4: Anchoring index for each feature.

Overall we can conclude the anchors did effect the answers to the questions in almost all of the questions. Therefore, no anchors should be put when doing an prototype session. If people do not have a very strong opinion about the prototype they might adjust towards the anchor and the results will not be entirely true.

5.2.2 Group Dynamics & Strong Opinions

Figure 4 shows that the anchoring effect is not the same in all the questions. We noticed by talking with some participants that they are more drawn to the anchor when they do not have a strong opinion about a feature. This could be explained by group dynamics. The participant prefers to follow the average, the opinion of the group. On the other hand, when participants are strongly opinionated on a subject, we noticed that they also get affected by the anchor but in the opposite way. Instead of changing their answer towards the anchor they drive it away. When we asked why they did that their answer was that they thought that the rate of this specific feature was really unfair so they wanted to change the average towards their opinion.

5.2.3 Availability Heuristic

The last question of the survey was asking the participants to rate the whole eTextbook. In this question we noticed that the survey that ended with the bad results and was anchored also lower had lower ratings, whereas the other survey had a higher rating (see Figure 5). We believe that this result is a combination of the anchoring and the availability heuristic since the participants could easier recall the last features.

5.3 Conclusion

The conclusions of this experiment verified the following guidelines. In order to eliminate the influence from group dynamics the participants should not be able to exchange opinions during the prototype session. Additionally, information that indicate current rating or opinions about the prototype should not be given to participants to avoid anchoring.
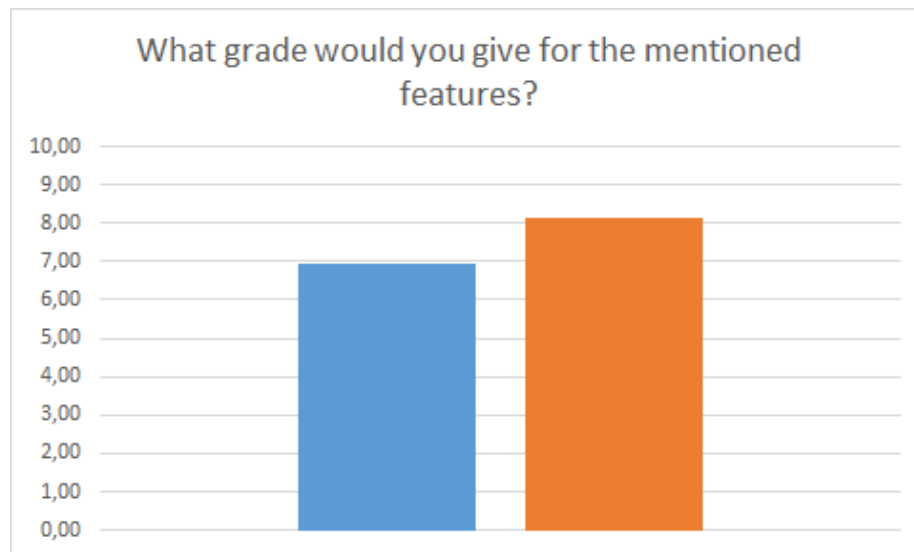
Figure 5: Overall rating of the features.