



Flavor Explore: Rapid Prototyping and Evaluation of User Interfaces

Shi Qiu^{1(✉)}, Liangyi Du², Ting Han², and Jun Hu¹

¹ Department of Industrial Design, Eindhoven University of Technology, Eindhoven, The Netherlands
{SQU, J. Hu}@tue.nl

² Department of Design, Shanghai Jiao Tong University, Shanghai, China
duliangyi@mail.sjtu.edu.cn, hanting@sjtu.edu.cn

Abstract. Expert-evaluation methods, such as cognitive walkthrough and heuristic evaluation, are widely used in user experience studies due to the reason that it can identify usability problems early and fast and find areas for improvements. In this paper, we present evaluation of Flavor Explore, a high-fidelity prototype, aiming at searching for delicious food and nearby restaurants. A task-based evaluation was conducted to evaluate primary search tasks of the prototype, which included five experts with the background in interaction design and two potential users. Twenty two usability problems were then identified. The findings were helpful for improving the design and for further research.

Keywords: Expert evaluation · Rapid prototyping · Heuristic Evaluation
Cognitive Walkthrough · Usability problems

1 Introduction

Rapid prototyping and expert evaluation are important methods in the development of interactive systems. In the process of rapid prototyping, designers often use sketches, wireframes and interactive prototypes to visualize the concepts. Designers iteratively develop the high-fidelity prototypes to simulate user interface is often helpful and represent the final product. At this point, it makes sense to have experts identify usability problems with these prototypes before exposing to users. In a rapid process of developing systems such as web and mobile applications, expert evaluation may be the only method available before these system go online [1]. There are many forms of expert evaluation, including Cognitive Walkthrough [2, 3], Guidelines Review [4], Consistency Inspection [5] and Heuristic Evaluation [6–8]. Among these methodologies, Cognitive Walkthrough (CW) and Heuristic Evaluation (HE) are widely used in industry. In practice, sometimes a hybrid approach is adopted that combines CW and HE together in a task-based evaluation. In this paper, we present Flavor Explore, a high-fidelity prototype, aiming at searching delicious food and nearby restaurants. After implementation of the prototype, a task-based evaluation which combines CW and HE was adopted to evaluate primary search tasks. The procedure includes these steps: (1) define the target users and their purpose by use scenarios; (2) define three primary search

tasks they will attempt; (3) walk through each task step-by-step; (4) look and identify usability problems on a set of heuristics; (5) explain where in the user interface the problem is, how severe it is and possible design improvements. Five experts with background in interaction design and two potential users followed this procedure to complete the evaluation. Twenty two usability problems were then identified.

2 Prototype

Brainstorming sessions were conducted based on SET factor analysis ((Social trends (S), Economic forces (E), and Technological advances (T)). We tried to find out design opportunities for conceptualization. This approach, focusing on the product concept and quickly finding out its market opportunity, is widely used in innovative industrial product development [9]. Based on the brain storming results, we positioned our design opportunity on creating a mobile application named Flavor Explore, aiming at users to search for delicious food and nearby restaurants.

2.1 Persona and Scenarios

We clarified and visualized our concept of Flavor Explore by using persona and scenarios, which were developed at early stages of the concept design. Instead of considering user behavior and experience through formal analysis and modeling of well-specified tasks, scenario-based design is a relatively lightweight method for envisioning future use possibilities [10]. Here we identify three typical scenarios of Flavor Explore in our design: searching for restaurants (Fig. 1(a)), recommending dinner menu (Fig. 1(b)) and suggestions for nutrition (Fig. 1(c)).

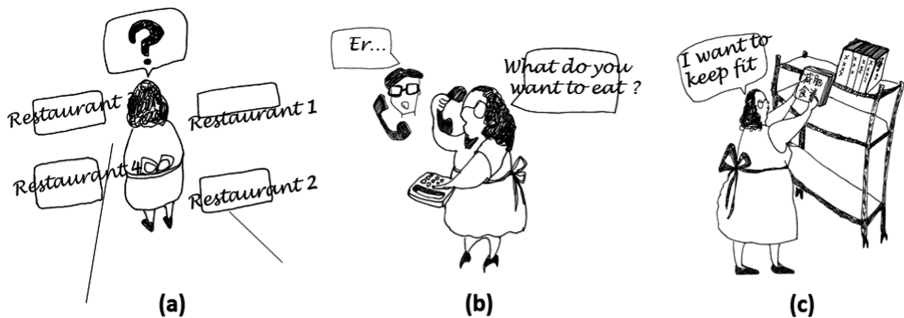


Fig. 1. Three typical scenarios of Flavor Explore: (a) searching for restaurants, (b) recommending dinner menu, and (c) suggestions for nutrition.

Persona

Emilia is a housewife and her son is a high school student, who has dinner at home with Emilia almost every day. Emilia likes travelling and tasting delicious food in her leisure time. She also concerns about a balanced diet and wants to keep fit.

Scenario 1: Searching for Restaurants

Emilia goes to Hong Kong for her holiday. When she first goes shopping in Mong Kok, she feels tired and hungry. She wants to find a restaurant with local special food. There are many kinds of restaurants on both sides of the street and she cannot decide which one is her favorite. She wants to know nearby restaurants conveniently. In the street, she takes out her mobile phone and starts the camera. From the interface of the camera, she can see information of the nearby restaurants on semi-transparent layers floating in the real-street scene. She checks recommendations and decides to go to Rose Restaurant. In the restaurant, she is interested in a snack looks like dumplings and she wants to know more about it. She takes a photo of it and starts searching. Then she finds its name is *shaomai*, a traditional Chinese food which is worth trying. In this scenario, Emilia uses AR (Augmented Reality) Search, a concept search that based on Augmented Reality technology on mobile photos [11], to find out information of the nearby restaurants. She also uses Image Search to know the name of the snack.

Scenario 2: Recommend Dinner Menu

Emilia is about to prepare dinner at home. She talks with her son by phone and she asks him: “what do you want to have for the dinner?” Her son says: “This week we have three times to take beef with tomatoes. Maybe we can try something different...” Emilia accepts her son’s idea and starts flavor Explore to search some recommended cookbooks. In this scenario, Emilia finds out new cookbooks through “Top Recommendations” in Flavor Explore.

Scenario 3: Suggestions for Nutrition

Emilia is now 45 years old. She is no longer slim and not satisfied with her body shape. She wants to lose weight and keeps fit. She tries to keep a balanced diet with lower calories. When she searches for new cookbooks in Flavor Explore, she used “Sort by Calories” to find out most healthy cookbook. In this scenario, Emilia finds out healthy cookbooks through “Sort by Calories” in Flavor Explore.

Based on the scenarios, we proposed three features of Flavor Explore as follows:

1. *Comprehensive search* Combine text, image and AR search to help the user find out food and restaurants.
2. *Top Recommendations* Recommend new and popular cookbooks to the user.
3. *Healthy cookbooks* Use “Sort by Calories” to find cookbooks with lower calories.

2.2 Rapid Prototyping

Sketches of the interface wireframes were created based on three features of Flavor Explore. The wireframe is a visual guide that represents the skeletal framework of an interface. It often lacks color or graphics and primarily focus on the functionality, behavior and priority of content [12, 13]. We were fast to draw the wireframes on a piece of paper and to clarify elements of the user interfaces (Fig. 2(1)). The paper-version wireframes provides a rapid way to visualize design ideas in an early stage. It helps to organize ideas and modify the design concept in an iterative design process. In order to

make the prototype more interactive, we used Axture¹ to create an interactive HTML prototype. The axure prototype is shown in Fig. 2(2). The final high-fidelity prototype was created by Adobe Flash Catalyst² (Fig. 2(3)). The high-quality images of the user interface in PSD format could be imported into Adobe Flash Catalyst conveniently.



Fig. 2. (1) Wireframes of user interfaces, (2) the interactive HTML prototype created with axure, and (3) high-fidelity prototype created with Adobe Flash Catalyst.

3 Evaluation

Five experts from two online companies and two potential users were invited to evaluate the prototype. The evaluation scope was limited to three search tasks: Image Search, AR Search and Text Search which were completed in the high-fidelity prototype. The task flow of three search tasks is shown in Fig. 3.

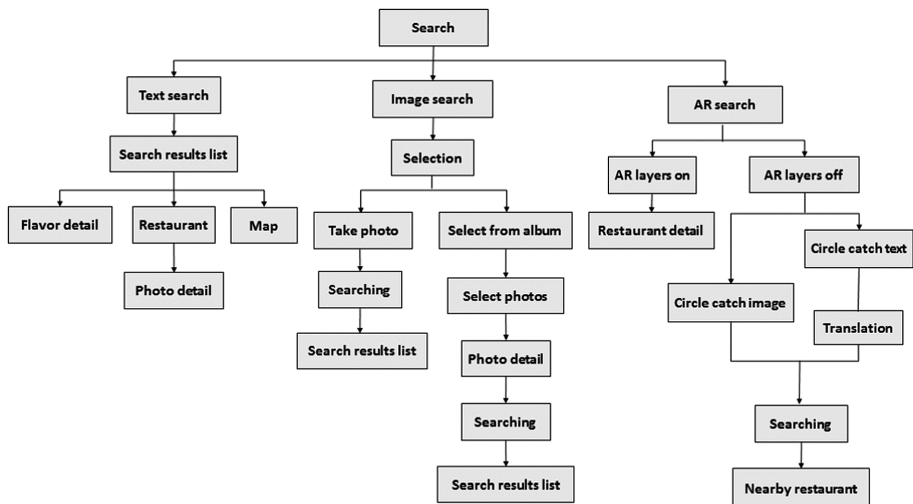


Fig. 3. Task flow of three search tasks in Flavor Explore

¹ <http://www.axure.com/>.

² <http://www.adobe.com/hu/products/flashcatalyst.html>.

3.1 Participants

Five usability experts participated, four of which were female. All of them had higher education qualifications and they had experience in evaluating average five mobile application projects. They were professional interaction designers and user researchers from online companies. Four were from Baidu (Chinese online company, primarily providing search engine services) and the other from Dianping (Chinese online company, primarily providing merchant information and consumer reviews). Two female users were invited to participate in evaluation. They were postgraduate students from School of Design, Hong Kong Polytechnic University. Both had sufficient experiences in using smart phones.

3.2 Setup

Two users participated in the evaluation in Hong Kong Polytechnic University in Hong Kong and five experts conducted the evaluation in online companies in Chinese mainland. Two users used a computer that installed the prototype previously. For the experts, we sent them the high-fidelity prototype with instructions by email before the test. The prototype was asked to be installed on their computers, which was more convenient for testing. The experts could choose to evaluate the prototype at home or at the office, where they should ensure surroundings as quiet as possible. Four experts evaluated at the office and one did it at home. They were asked to record a video with sound during evaluation. The video images contained the whole evaluation setting and partially the expert.

3.3 Procedure

Evaluation took 20–30 min. Table 1 shows the procedure. The pre-questionnaire included questions pertaining to users' demographics and their domain knowledge of smart phones. After that, the participant read the instruction to run the prototype on the computer. There were three search tasks (Image Search, AR Search and Text Search) for the participant to complete. Before doing each search task, the participant was provided with a use scenario and then decided which search method was most suitable for that scenario. The example scenario was: "*Assume you are on a journey abroad and look for 'Rose Restaurant' in the street. Try to find out remark stars of this restaurant.*" During the search task, the participant was asked to follow the think-aloud protocol. She needed to simply verbalize her thoughts when moving through the user interface. After completing the task, she was asked to evaluate it. After completing three tasks, the participant was asked to answer the question "*which search method do you prefer to choose? Please explain the reasons.*"

Table 1. The procedure of the test

No.	Procedure
(i)	The participant (only experts) downloads and installs the prototype on a computer
(ii)	The participant answers pre-questionnaire
(iii)	The participant conducts three search tasks in a sequence. After completing each task, participant answers a short questionnaire
(iv)	The participant answers an open question after completing three tasks

3.4 Measurements

Six evaluative principles were chosen from a categorization of heuristics and guidelines with twenty types [14]. These principles were easy to use, predict next step, clear icon metaphor, have no mistakes, same logical, and have feedbacks. After completing each task, participants were asked to give the weight to six principles and evaluated the task based on a seven-point scale from “strongly disagree” to “strongly agree”. The weight of each principle ranged from zero to one. Zero means never care while one means very care. For two users, the researcher explained the meaning of evaluative principles, weight and the seven-point scale to them and ensured that they could well understand before the test.

4 Results

Quantitative analysis was based on the result of the seven-point scale and divided into two categories: experts and potential users. The average score of experts in Image Search, AR Search and Text Search were 4.782, 2.132 and 4.942. Three experts were unable to finish AR Search: two experts used Text Search to complete the task which aimed to use AR Search and the other one just clicked the “Explore” tab. AR Search was scored lower than other two searches, which indicated potential usability problems. The average score of users in Image Search, AR Search and Text Search were 3.415, 0 and 7. Both of them did not complete AR Search and one of them did not complete Image Search. They finished Text Search without any difficulties, and the average score of Text Search was the highest one among three methods.

Twenty two usability problems were identified by experts and they also proposed suggestions for improvements accordingly. Two users almost kept silent during the evaluation and did not “think aloud”. Usability problems identified by experts were ranked by frequency based on the video analysis. As shown in Fig. 4, the usability problems of search result page (No. I-5) were reported for 8 times, while restaurant detail page (No. D-3) for 6 times. Other pages were reported less than 3 times. Table 2 illustrates the categorization of example usability problems and design proposals for improvements. For the open question, we collected experts’ comments about the preferable search method. Three experts proposed the search method was chosen based on the requirement of different scenarios. The example was “*my preferred search method depends on different scenarios of searching. For example: if I want to find a restaurant for lunch in a journey, I will probably choose AR Search. The reason is that I have no*

clear goal at that time. If the name of the restaurant is very clear and not nearby, key words search will be more suitable.” One expert preferred Text Search and the reason was “I prefer Text Search, because the keyword is very accurate. I cannot trust image search for the reason that some photos are difficult to recognize.” Other two experts preferred AR Search and Image Search. The reasons were: “my favorite search method is AR Search which looks novel and joyful.” “I will prefer Image Search, because I do not need to type anything and only need to take a photo. Then the search starts quickly, quite convenient for me.”

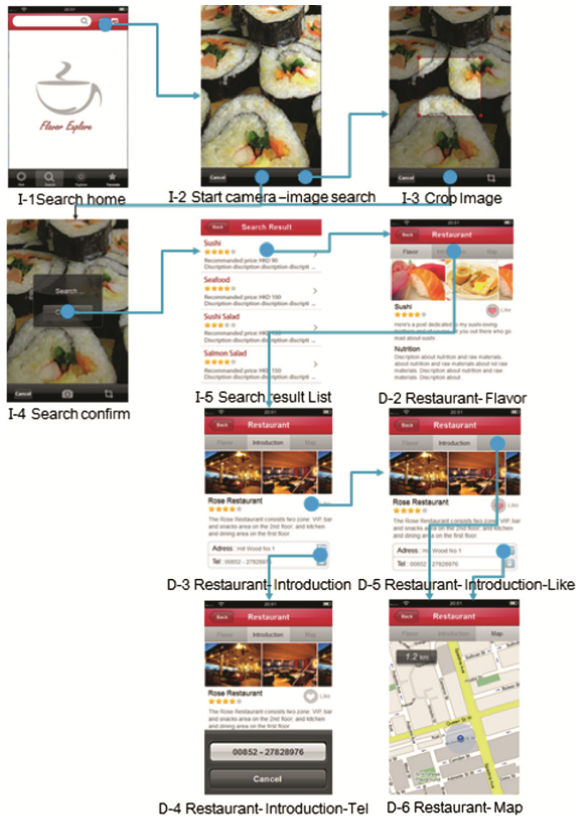


Fig. 4. The flow of the user interfaces

Table 2. Categorization of example usability problems and proposals of improvements

Page no.	Examples of usability problems	Suggestions of improvements
No. I-5 search result	The search result page does not have any pictures	The search result should contain pictures to help the user to make a decision
	The search result page has no ranking remarks, so it cannot help to decide which restaurant is most recommended	Add ranking remarks in search result page
	Input “Rose Restaurant” in Text Search, the search result has no restaurant information and all about food information	Give user a hint in the search box. For example: please input food name
	The search result page has problems to provide more options to the user and filter the search results	Provide different ordering rules to filter the results. For example: to provide ordering according to distance, price, taste and remarks
No. D-3 restaurant detail	Click the telephone number without any response	Increase clickable area rather than click a small icon
	The wording “Like” is not consistent with “My Favorite”	“My Favorite” changes to “Like”
	Unknown about the information of restaurant environment	Provide the information of restaurant environment
	The icon of “Like” is a heart shape and the icon of “Favorite” is a five-pointed star. They are not consistent in visual language	Make the icon of “Like” and “Favorite” consistent
	Difficult to find out telephone number of the restaurant and no response to click the item	Click the item to activate the phone app
No. A-1 AR search	AR virtual layers cannot provide discount information	Provide discount information on virtual layers
	The switch of the AR layers is confusing. Why need to close AR layers?	Provide a hint to the user to explain the function of AR switch

5 Discussion

5.1 Differences Between Experts and Potential Users

Comparing the experts to the users in average scores of search tasks, experts’ scores tend to be more moderate. The average scores of the users have a tendency of a dramatically difference between AR Search ($M_{AR} = 0$) and Text Search ($M_{TS} = 7$). The experts are proficient to complete search tasks and seldom influenced by the usability problems. They tend to evaluate more objectively. They are more patient and give helpful suggestions during the evaluation. However, the users are easily influenced by usability

problems. If they cannot complete the task, they tend to give up quickly. So the final score of the search task drop down dramatically. They are more sensitive to usability problems and even possibly exaggerate the severity of the problem compared with experts. In other words, a small usability problem for experts may become a big one for the end users. This phenomenon may be influenced by limited numbers of potential users. In this study, only two potential users were involved. If more potential users participate in the test, the result may be different.

5.2 Usability Problems and Modifications of AR Search

The design of AR Search causes some usability problems. From the video observations, we find one expert and one user try several times to click the “Explore” tab when they did the task of AR Search. It may indicate the wording “Explore” is more relevant to their knowledge of AR Search. The function of “Explore” in this prototype is to be based on LBS (local based service) technology, to provide the information of nearby restaurants to users. The interaction shares some similarities with AR Search. AR Search could be also a kind of “Explore”: when a user walks in the street, she can start her camera and check nearby restaurant information in semi-transparent layers. Therefore, we move the “AR” button from “Search” tab to “Explore” tab in the new design. The wording of “AR” is an abbreviation of Augmented Reality. It makes participants feel confused and it might be not necessary for them to know the exact meaning of AR. In the new design, AR button is located on the “Explore” tab and named “layer” which vividly explains its visual appearance. The text “layer” can also use the icon instead.

5.3 Limitations

Considering time and cost, the prototype was evaluated on the computers rather than mobile phones. Usability problems caused by the smaller screen might be neglected. In our future work, we will implement the prototype on mobile phones and conduct the evaluation accordingly. All participants are interaction and user experience designers. They might share the similar ways of thinking. In our future work, we will involve participants with different backgrounds. In this study, we only invited two users. The comparable results between users and experts might not be convincing and we need to recruit more users in the future study.

6 Conclusion

In this paper, we presented the evaluation of a mobile application of Flavor Explore. The implementation of the prototype was an agile procedure. We made the search task available first in the prototype and conducted the evaluation. The prototype was far from perfect, but aiming at shortening the time of development and trying to identify the usability problems in an early stage. In order to find out usability problems, we conducted a task-based evaluation for primary search tasks. Five experts and two users were invited to this study. Users tended to be more sensitive to usability problems. AR Search had

some usability problems and we proposed the design improvements based on the suggestions from evaluation.

Acknowledgements. This research is supported by the China Scholarship Council and facilitated by Eindhoven University of Technology. We also thank the support from Tina, Yangzi Li, Xueting Xie in Hong Kong Polytechnic University, Shuang Shi, Ying Chen, Fuhong Han and Jing Li from Baidu as well as Xuejiao Li from Dianping.

References

1. Petrie, H., Power, C.: What do users really care about?: a comparison of usability problems found by users and experts on highly interactive websites. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2107–2116 (2012)
2. Lewis, C., Polson, P.G., Wharton, C., Rieman, J.: Testing a walkthrough methodology for theory-based design of walk-up-and-use interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 235–242 (1990)
3. Lewis, C., Wharton, C.: Cognitive walkthroughs. In: Handbook of Human-Computer Interaction, vol. 2, pp. 717–732 (1997)
4. Dix, A.: Human-computer interaction. In: Liu, L., Özsu, M.T. (eds.) Encyclopedia of Database Systems. Springer, Heidelberg (2009). https://doi.org/10.1007/978-0-387-39940-9_192
5. Lazar, J., Feng, J.H., Hochheiser, H.: Research Methods in Human-Computer Interaction. Wiley, Hoboken (2010)
6. Molich, R., Nielsen, J.: Improving a human-computer dialogue. Commun. ACM **33**(3), 338–348 (1990)
7. Nielsen, J.: Usability Engineering. Elsevier, Amsterdam (1994)
8. Nielsen, J., Molich, R.: Heuristic evaluation of user interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 249–256 (1990)
9. Cagan, J., Vogel, C.M.: Creating Breakthrough Products: Innovation from Product Planning to Program Approval. Ft Press, Upper Saddle River (2002)
10. Rosson, M.B., Carroll, J.M.: Scenario based design. In: Human-Computer Interaction, Boca Raton, FL, pp. 145–162 (2009)
11. Höllerer, T., Feiner, S.: Mobile augmented reality. In: Telegeoinformatics: Location-Based Computing and Services, vol. 21. Taylor and Francis Books Ltd., London (2004)
12. Brown, D.M.: Communicating Design: Developing Web Site Documentation for Design and Planning. New Riders, San Francisco (2010)
13. Garrett, J.J.: The Elements of User Experience: User-Centered Design for the Web and Beyond. Pearson Education, London (2010)
14. Weinschenk, S., Barker, D.T.: Designing Effective Speech Interfaces. Wiley, Hoboken (2000)