Evaluation Methods for Service Design

Xinlei Chen

Abstract

Service design is not a direct path but requires repeated creation and improvement in order to make progress. Suitable evaluation methods can be used to assess the results of each step (concept plan, mock-up, prototype, etc.) to avoid moving in the wrong direction. This article introduces ideas and evaluation methods used in service design.

Keywords: human-centered design, evaluation methods, prototyping

1. Introduction

The idea of human-centered design[*1] is the basic concept in service design. Everything begins with gaining an understanding of users and having empathy with them and proceeds in steps—defining the problem, finding a solution, developing a concept plan, producing mock-ups, and creating prototypes—until a solution is finally produced and the design is realized. However, this process is by no means linear and involves repetitions of creation, improvement, and selection. As such, decisions have to be made every day (concept A or concept B, design 1 or design 2, etc.). Correct decisions are necessary in order to proceed in the right direction, and the correct basis must be obtained to make them. One important way to obtain the correct basis is by conducting evaluations. Evaluation of service design has the following characteristics in contrast to conventional evaluation of existing services.

(1) Diversification of what is evaluated

In evaluating existing services, the service is usually evaluated in its completed form. However, during the service design process, most evaluations are done on intermediate results such as a concept plan, rough mock-up, or prototype. When such results are evaluated, it is important to carefully study how the service is shown to users and what is actually being evaluated in order to obtain correct results.

(2) Novelty of what is evaluated

To understand whether a service will be effective or acceptable, it can be compared with existing services. However, for very novel services, there is often no existing service to compare it with. In such cases, other effective evaluation methods must be studied.

(3) Repetitiveness of evaluation

Until now, evaluations were basically done before improvements were made (when problems were found in an existing service), or when a prototype or service was completed (checking the effect of improvements). Since service design is iterative, it is more effective to do evaluations at each step and to feed the results back into the next step.

(4) Acceleration of feedback

The amount of time that can be spent on each step in service design is limited, so evaluations must be done quickly. In some cases, there is insufficient time to gather enough participants. Therefore, a way to obtain evaluation results in a shorter time period is needed.

NTT Advanced Technology has been providing consulting services supporting research and development and improvement of products and systems based on human-centered design for over 20 years.

---

*1 Human-centered design: A development approach for making systems more user-friendly and usable. A detailed definition is given in ISO 9241-210.

Table 1. List of evaluation methods.

Design phase	To be evaluated	Evaluation objectives	Evaluation methods
(1) Understanding users	–	–	–
(2) Defining issues	–	–	
(3) Studying solutions	- Scenarios - Storyboards - Concept videos	- Checks concept validity	- Interviews - Surveys
(4) Designing	(Rough prototypes) - Sketches - Wireframes - Mock-ups - Paper prototypes	- Checks function acceptability - Verifies information architecture (navigation, categories, layout, labeling, etc.) - Verifies usability (effectiveness, efficiency)	- Interviews - Expert reviews - User testing - Card sorting - Simple evaluations

	(Advanced prototypes)
	- Advanced visual mock-ups
	- Interactive prototypes
		- Checks service acceptability
		- Verifies information architecture
		  (hierarchical visual structure etc.)
		- Verifies usability
		  (effectiveness, efficiency, satisfaction)
		- Checks price acceptability
			- Interviews
			- Expert reviews
			- User testing
			- Simple evaluations

(5) Development	- Real prototypes	- Checks effectiveness of use in real or simulated environments - Checks usability in real or simulated environments	- Field testing - Performance measures

During that time, we have accumulated know-how regarding characteristics of user behavior, various interface design patterns, and evaluation methods such as surveys. In collaboration with people from the Universal User Experience (UX) project at NTT Service Evolution Laboratories, we have applied this knowledge to service design and have developed simple evaluation methods that can be performed efficiently and effectively.

2. Evaluation methods

It is important to select evaluation methods suited to the type of results and evaluation objectives for each stage of service design. A list of evaluation methods is given in **Table 1**. The details of evaluation in each of the phases are described below.

2.1 Solution study phase

The main evaluation objective in this phase is to verify the solution concept. The various concept plans being studied are expressed in easy-to-understand text (scenarios), illustrations (storyboards), and video (concept videos). These are used to explain the solution to users, ask their opinions directly, and verify which concept plan evokes the best response. It is desirable to hear opinions from as many users as possible, but actual costs (time, money, etc.) are usually limited, so methods that integrate less expensive web surveys can also be used.

2.2 Design phase
(1) Rough prototype

In this phase, rough prototypes are created, from sketches of the idea to wireframes, mock-ups, and paper prototypes. The main functions and screen framework of the service can be seen, so the following areas should be verifiable before performing graphical and interaction design.

1) Function acceptability: Evaluate whether the main functions to achieve the user's objectives are included, and whether these functions seem attractive. An effective evaluation method is to have users experience the prototype and then conduct interviews with them.

2) Information architecture suitability: Evaluate the comprehensibility of the overall screen layout, navigation, categories, labels, and other aspects. Expert review by a specialist in human-centered design and user testing are often used. Card sorting is also useful for evaluating the appropriateness of categories.

3) Usability: Evaluate whether objectives are achieved (effect), whether there are wasteful procedures, and whether objectives can be achieved by the shortest possible path (efficiency). There are indices for usability satisfaction (whether users feel any discomfort), but evaluating it at this stage is not recommended because prototypes at this stage are visually incomplete, and this could introduce some subjective bias. Expert reviews and user testing are

common methods to evaluate usability, and the *Wizard of Oz* method is often used for user testing, where in place of a computer or smartphone, a human operator runs the prototype according to user operations. If time is limited, a simple evaluation can be done in three days (not including recruiting of participants), combining expert reviews and user testing with two participants.

(2) Advanced prototype

During this phase, design of overall screen graphics (style, color scheme, icons, etc.) and detailed animations (smartphone screen window swipes and motions, etc.) are prototyped. Thus, visual and subjective acceptability can be evaluated.

1) Service acceptability: Check the overall attractiveness of the service, including appearance, and disposition toward using it. Detailed views can be gained in interviews with users and through questions about the conditions and environment after they have experienced the prototypes.

2) Information architecture suitability: Evaluate visual aspects that could not be evaluated with the rough prototype (e.g., how color and font size in headings make the text structure easier to understand). An expert review by a specialist and user testing can be used for evaluation.

3) Usability: Evaluate whether user interface (UI) elements on the screen are recognizable in terms of effectiveness and efficiency (e.g., whether it is clear that a given element is clickable as a button). Evaluations such as expert reviews and user testing or a simple evaluation combining them can be used here as well.

4) Price acceptability: Pricing is usually evaluated through interviews, but effective answers cannot be obtained by simply asking "How much do you think you would pay to use this?" This is because users tend not to give high prices, and users have differing standards, so responses tend to vary greatly. From our experience, we obtain the most useful responses when we ask interested users, "We plan to charge around xxx yen. Do you think you would want to use it?"

2.3 Development phase

In this phase, real equipment can be used in the evaluation based on real data. Effective field testing can be conducted with real target users in a real environment. The effectiveness of the overall service can be judged by having users actually experience it,

observing changes in behavior before and afterwards, determining conditions through journaling, and understanding users' views through interviews. These methods can be used even if there is no existing service with which to compare, and significant evaluation data can be obtained. If quantitative data are needed, user behavior not obtainable from service logs can be understood from performance measurements. Examples of specific measurements are as follows.

- Effect: Achievement rate (task completion rate)
- Efficiency: Work time or numbers of errors, help requests, steps, etc.
- Satisfaction: Subjective evaluation of difficulty, comprehensibility, ease of training, consistency, and other factors

3. Points for improving evaluation quality

(1) Thorough preparation

Preparation before evaluation may be the most important step. Before planning an evaluation, the following should be reconfirmed within the design team and summarized in a document.

- Applicable service: Product name, features, main functions, related tasks, enterprise strategy, development state, user environment
- Relevant users: User class, basic attributes and characteristics, skill characteristics, usage objectives

(2) Appropriate restraint

What can and must be evaluated during each design phase is fixed. It is important not to rush ahead, and to carefully follow the plan. Missing things that need to be evaluated can be a problem, but it is even more dangerous to obtain reference data for things that cannot be evaluated. This is because there is a danger that referring to inaccurate data can have more influence on decisions than expected. If relevant evaluation results cannot be obtained, it is better not to do the evaluation.

(3) Gathering suitable participants

It is important to gather participants that are similar to the anticipated users, but it is even more important to gather users that will be interested in the service. Users with different motivations will not only have different levels of acceptance of and satisfaction with the service, but they will also have different levels of acceptance regarding usability. Those that are interested, even if their literacy is low, will not consider small difficulties troublesome, and they will learn quickly. Suitable conditions and questions (if recruiting

through the web) must be set when recruiting participants to understand the motivations of participants accurately beforehand.

4. Future prospects

Recently, the number of our projects has been increasing, and they range from requests for evaluation only to requests for total support from prototyping through to evaluation. A feature of NTT Advanced Technology is that we have teams for both development (web, mobile, IoT (Internet of Things), etc.) and UX design. They communicate easily, so they can move quickly from rough prototyping to producing real prototypes. We will continue to utilize this strength to meet the demands of customers and to provide higher quality development and evaluation, and we will use our entire capacity to support service design for our customers.

Xinlei Chen
UX Consultant, NTT Advanced Technology.
She received an M.E. from the Department of Information Management Science at the Graduate School of Information Systems, the University of Electro-Communications, Tokyo, in 2002. She joined NTT Advanced Technology in 2002 and has been working as a UX consultant. She is also a Certified Human Centered Design Professional.