

User Experience Registry for Generalizable Usability Assessment

Lin Wang
U.S. Census Bureau
Washington D.C., USA
lin.wang@census.gov

Usability assessment is a crucial component in human-centered design (HCD) throughout the systems development lifecycle (SDLC). Sound usability assessment ensures users to operate the system effectively, efficiently, and with satisfaction. However, in the reality of frequent releases and tight development schedules, formative testing is often reduced and summative testing eliminated. To address these challenges, this position paper proposes a framework for a user experience registry as a solution to aggregating data collected from formative testing, with the goal of improving the robustness and generalizability of the findings based on such data.

User experience, Usability assessment, Usability evaluation, Usability testing, Database, Registry

1. INTRODUCTION

Usability assessment is a crucial component in human-centered design (HCD) throughout the systems development lifecycle (SDLC). Sound usability assessment ensures that the user operates the system effectively, efficiently, and with satisfaction. Conceptually, there are two forms of usability assessment, formative testing and summative testing. These two types of testing go hand in hand: the formative testing is to be conducted iteratively during SDLC for identifying usability problems and investigating their root causes, while summative testing is to be conducted after the system is developed for benchmarking users' performance that is to be generalized to the target user population. However, in the reality of frequent releases and tight development schedules, formative testing is often reduced and summative testing eliminated. Consequently, usability issues are not adequately addressed and user experience is compromised. In addition, due to small sample size (usually 5-10 users) and the qualitative nature of formative testing, it is difficult to generalize findings from a small group of test participants to a large target population, potentially resulting in inferior external validity and poor cost effectiveness. To address these challenges, this position paper proposes a framework for a user experience registry as a solution to aggregating data collected from multiple formative usability tests, with the goal of improving the robustness and generalizability of the findings based on such data. We will also present a hypothetical design for a user experience registry for web survey instrument usability as a case study.

2. THE FRAMEWORK

A user experience registry is an organized system that aggregates user data (user performance and covariates) collected from various formative usability assessments to evaluate the effectiveness, efficiency, and satisfaction with using a particular system (e.g., an online communication application like *Skype*) by the specified user population to achieve specified goals. The infrastructure of a registry consists of two major components: A repository for storing the data, and a set of analytical methods or tools that are used to analyse user experience.

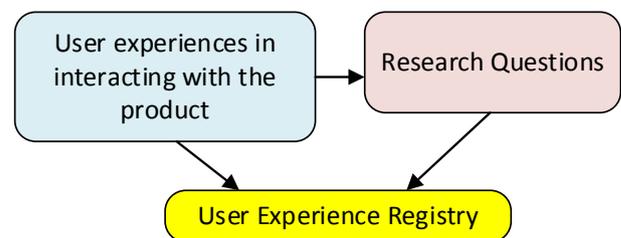


Figure 1: *The relationships among user experiences, research questions, and user experience registry: Research questions come from user experiences; research questions guide user experience registry development.*

Designing a user experience registry starts with formulating research questions. Basically, there are

three types of questions: What are the usability deficiencies? What are the root causes of the deficiencies? What are the remedies to the deficiencies? A registry can be designed to address some or all of these questions, by analysing the data collected during usability testing regarding effectiveness of, efficiency of, and satisfaction with completing tasks using a particular system. Fig 1 depicts the progression from user experiences, to research questions, and to user experience registry. After the questions are defined, we construct a human operator model (Fig 2). Based on the model, we conceptualize the registry with four information modules: (1) product information, (2) usability use cases, (3) operation environment factors, (4) user characteristics, (5) user performance. When designing a registry, the following aspects should be carefully considered:

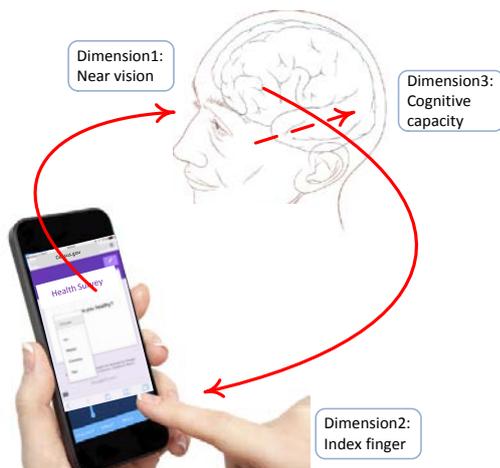


Figure 2: A sample operator model of information processing of mobile device operation. The information flows, and is processed, from the device screen to vision, to brain, to motor action.

Target user population. The target user population is the population to which the findings from the registry are meant to apply. This is the population who are the potential users of the same type of systems. For example, *WhatsApp* and *WeChat* belong to the same type of social media systems, while *WhatsApp* and *Twitter* do not. Data collected from usability testing of *WhatsApp* or *WeChat* can be included in the same registry, but not for that of *WhatsApp* and *Twitter*.

User sampling. Since the registry aggregates data collected from multiple formative usability tests, all the tests should use the same inclusion/exclusion criteria for recruiting test participants. The inclusion/exclusion criteria must be in line with the definition of the target user population. The sampling method (e.g., convenient sampling) for each test needs to be documented, as a potential covariate.

Performance measures. To empirically address the research questions, we need to translate the abstract constructs of effectiveness, efficiency, and satisfaction into measurable and uniform performance indicators that can be applied to all the usability testing of the same type of systems. Performance indicators for effectiveness (e.g., task completion) and efficiency (e.g., time on task) can usually be objectively quantified. Though there are existing questionnaires for measuring the feeling of satisfaction, one often finds that there isn't a one-size-fits-all instrument that can adequately measure users' satisfaction with various systems. We thus need to develop additional satisfaction measurement that is particularly tailored to a specific system, for example, a difficulty rating scale for *Google Map*.

Covariates. In addition to the system itself, user's interaction with the system can be influenced by personal characteristics and environmental factors, for example, age, language, education, etc. These are covariates that ought to be taken into account in assessing the system's usability.

Data elements and database. Performance measures are to be collected at person level, analysed at sample level, and generalized to population. Person-level data will be stored in the database. Each data variable stored in the database is a data element. Each data element needs to be explicitly defined. For example, for the measure of task performance success, there must be clear criteria for a success.

Analysis plan. A data analysis plan is the core component of the registry. The plan is a roadmap to addressing the research questions using the registry data. Due to the nature of heterogeneity in the user experience data, particularly the qualitative data, the method for analysing a particular performance measure needs to be well-thought and clearly defined. For example, the coding scheme for responses to debriefing questions needs to be documented with unambiguous rules.

Baseline performance. Assessing improvement in usability over the course of SDLC is one of a registry's main purposes. The data collected from initial formative usability testing on a system can serve as the baseline performance to be compared with results from later usability tests on the same but revised system for usability improvement.

Confidentiality. Since the registry usually contains personal identifiable information, such as date of birth, a rigorous access control must be applied to the registry. The control must comply the laws in the registry's country of residence.

3. A HYPOTHETICAL CASE STUDY

A survey instrument is a tool for obtaining self-reported data from respondents through standardized questions. Those questions may be presented on paper, on a computer screen via the Web, or verbally by a person or via telephone. Conducting surveys on the Web has become an important mode of survey data collection, thanks to the ubiquity of Internet and smartphones. For Web-based survey instrument, usability is crucial because it has a direct impact on survey data quality, called measurement error. Since Web survey instruments' basic functionalities are similar, a user experience registry for Web survey instruments would be a good approach to aggregating respondents' experience collected from web survey usability testing and consequently to help the survey methodology community improve Web survey design based on generalized respondent performance findings.

In this hypothetical registry, we ask two research questions: (1) What are the usability deficiencies that result in error in self-administered survey data entry? (2) What are the usability deficiencies that result in break-off (a situation where a respondent drops the survey before it is completed)? These two questions concern survey measurement error and non-response error, as conceptualized in the model of Total Survey Error (Fig 3).

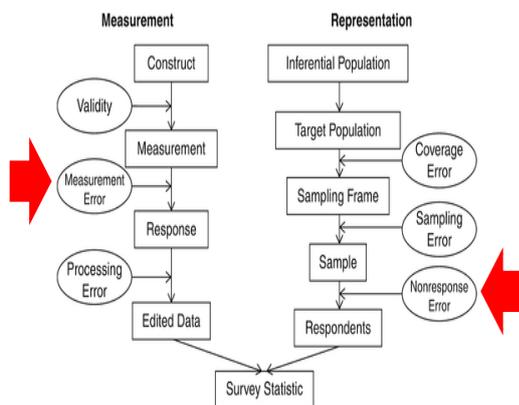


Figure 3: The model of Total Survey Error (Adapted from Groves et al, 2009).

The target user population is defined as individuals who are potentially able to respond to a survey in English on the Web. Based on this definition of target user population, we specify participant inclusion criteria for usability testing, as shown in Table 1.

Table 1: Participants Inclusion Criteria

Age	18-75
Language	Fluent in spoken and written English
Education	At least 8 years of schooling

Web experience	At least one year of surfing the Web, e.g., online shopping
----------------	-------------------------------------------------------------

Performance measures are designed to address the two research questions, as listed in Table 2.

Two categories of respondent characteristics are included as possible covariates: (1) Demographics – age, gender, native language, education; (2) Computer literacy – years of surfing the Internet, frequency of surfing the Internet, years of using a smartphone, frequency of using the smartphone.

Table 2: Performance Measures

Measure	Definition
Question 1: What are the usability deficiencies that result in error in data entry?	
Correctness	1=correct data entry, 0=incorrect data entry
Correction	Number of extra key strokes in data entry
Question 2: What are the usability deficiencies that result in break-off?	
Complete	1=survey completed, 0=survey not completed

The performance measures and covariates make up the data elements. A relational database schema will be developed to host all the data elements. Free-text comments by the participants during usability testing are also included in the database.

The analysis plan consists of two parts: Qualitative data analysis and quantitative data analysis. For qualitative data, the methodology of grounded theory will be applied to generate hypotheses. For quantitative data, an approach of exploratory data analysis followed by confirmative data analysis, if applicable, will be used.

For each Web survey instrument, its first formative usability test is marked as the baseline performance.

The registry will enforce data confidentiality in compliance with U.S. NIST Cybersecurity Framework.

4. SUMMARY

In this position paper, we proposed the concept of user experience registry as a solution to aggregating data collected from formative testing and to improving the robustness and generalizability of the findings based on this type of data. Though the notion is still in its infancy and much more thinking and experimentation is in need to nurture it into maturity, it is hoped that this paper will stimulate enthusiasms on exploring more effective methods of improving user experiences in interacting with modern technology.

5. REFERENCES

- Gliklich, RE, Dreyer, NA, eds. (2010) *Registries for Evaluating Patient Outcomes: A User's Guide* (2nd Edition). Agency for Healthcare Research and Quality, Rockville, MD.
- National Quality Registry Network. (2014) *What is a Clinical Data Registry?* American Medical Association.
- Workman TA. (2013) *Engaging Patients in Information Sharing and Data Collection: The Role of Patient-Powered Registries and Research Networks*. AHRQ Community Forum White Paper. AHRQ Publication No. 13-EHC124-EF. Rockville, MD: Agency for Healthcare Research and Quality.
- ISO 9241-11 (2015) *Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts*. International Organization for Standardization, Geneva.
- Groves, RM, Fowler, FJ, Couper, MP, Lepkowski, JM, Singer, E, Tourangeau, R. (2009) *Survey Methodology*. Wiley, New Jersey.