RESEARCH REPORT SERIES
(Survey Methodology #2016-05)

A Systematic and Multidisciplinary Approach to Developing an Evidence-based Framework of User Interface Design for Mobile Survey Instruments

Lin Wang
Christopher Antoun
Russell Sanders
Elizabeth M. Nichols
Erica L. Olmsted-Hawala
Jonathan M. Katz
Brian Falcone
Ivonne Figueroa

Center for Survey Measurement
Research and Methodology Directorate
U.S. Census Bureau
Washington, D.C. 20233

Report issued: October 20, 2016

Abstract

This paper describes a systematic and multidisciplinary approach to developing evidence-based standards and guidelines of user interface design for mobile survey instruments. The approach includes five major components: project team, systematic literature review, mobile user model, standards development, and guidelines development. A multidisciplinary project team was assembled to ensure adequate human resources. Systematic literature reviews were conducted on mobile user interface design and on mobile survey instrument design. Based on the results from the reviews, decisions were made to develop standards for basic mobile user interface elements and guidelines for mobile survey instrument user interface. Methodologies for developing the standards and guidelines were described in the paper. This approach emphasizes the importance of systems engineering principles, multidisciplinary collaboration, methodological rigor, and practical problem-solving.

Keywords: mobile, smartphone, user interface, UI, survey, Web survey, standards, guidelines, evidence

Suggested Citation: Lin Wang, Christopher Antoun, Russell Sanders, Elizabeth M. Nichols, Erica L. Olmsted-Hawala, Jonathan M. Katz, Brian Falcone, Ivonne Figueroa. **(2016). A Systematic and Multidisciplinary Approach to Developing an Evidence-based Framework of User Interface Design for Mobile Survey Instruments.** *Research and Methodology Directorate, Center for Survey Measurement Study Series (Survey Methodology #2016-05).* U.S. Census Bureau. Available online at <http://www.census.gov/srd/papers/pdf/rsm2016-05.pdf>.

A Systematic and Multidisciplinary Approach to Developing an Evidence-based Framework of User Interface Design for Mobile Survey Instruments

Abstract

This paper describes a systematic and multidisciplinary approach to developing evidence-based standards and guidelines of user interface design for mobile survey instruments.

The approach includes five major components: project team, systematic literature review, mobile user model, standards development, and guidelines development. A multidisciplinary project team was assembled to ensure adequate human resources. Systematic literature reviews were conducted on mobile user interface design and on mobile survey instrument design. Based on the results from the reviews, decisions were made to develop standards for basic mobile user interface elements and guidelines for mobile survey instrument user interface. Methodologies for developing the standards and guidelines were described in the paper.

This approach emphasizes the importance of systems engineering principles, multidisciplinary collaboration, methodological rigor, and practical problem-solving.

1. Introduction

A survey instrument is a tool for obtaining respondent-reported data through a scientific protocol. The instrument includes a set of questions presented in certain mode - on paper, on a computer (including a mobile device) screen, or via telephone. The paper and computer modes could be self-administered or interviewer-administered while telephone mode could be interviewer-administered only. A mobile survey instrument is a survey instrument that a respondent can interact with on a mobile device to answer survey questions. With the growing use of smartphones [1], many surveys can now be administered through mobile devices.

A key concern with mobile survey instrument development is how to optimally design instruments' user interface so that respondents' ability to enter responses correctly is maximized. During the process of completing a survey, various errors could be introduced into the survey responses. These errors can be characterized in a theoretical model, the Total Survey Error Framework [2], as shown in Figure 1. The user interface of a mobile survey instrument directly influences respondents' perception (information input) and action (information output) during survey completion, and may consequently introduce measurement error. To minimize measurement error, it is thus crucial to optimize the mobile survey instrument's user interface.

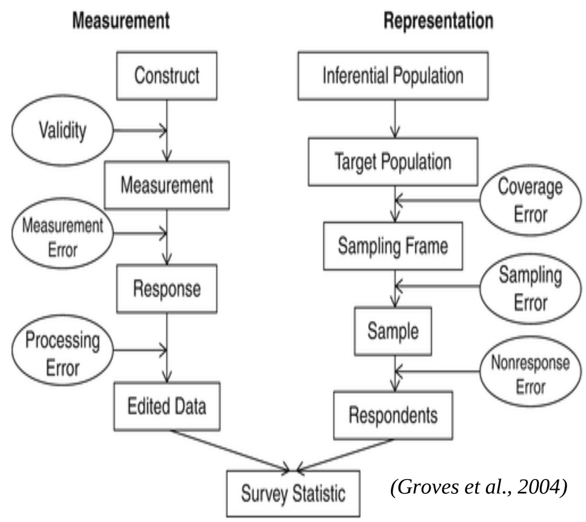


Figure 1. The Total Survey Error Framework

Because of the nature of a small screen size and touch interface, we are challenged with usability in mobile survey instruments. Usability issues concern effectiveness, efficiency, and satisfaction [3] with respondent's interaction with the instrument. Though research has been done on improving user interface for Web survey instrument designed for large screen, e.g., a desktop computer [4], much work is yet to be done for the user interface of mobile survey instruments. In particular, there are no standard user interface design specifications that are based on empirical evidence for mobile survey instruments. Inadequately mobile survey instrument user interface could result in errors in survey responses, prolonged time in completing a survey, and even breakoffs (i.e., the respondent quits the survey before it is completed). In addition, it could be costly to re-design and re-develop user interfaces [5, 6, 7]. To address this important issue, we engaged in an initiative to develop an evidence-based framework of user interface design for mobile survey instruments through a systematic and multidisciplinary approach. The framework will consist of standards and guidelines based on empirical evidence. The initiative is still in progress. In this paper, we will describe the major components of the systematic and multidisciplinary approach to this initiative.

2. The approach

A mobile survey instrument is a piece of software as well as a tool for collecting survey data. Completing a survey on a mobile device consists of a series of human-computer interactions. The respondent and the instrument comprise a closed-loop human-machine system in which information exchanges between the two components take place. It thus requires knowledge from multiple relevant disciplines to optimize the mobile survey instrument, including survey methodology, mobile software development, user interface design, survey instrument design, user research, and human factors research.

To effectively and efficiently carry out this study, we followed the principles of systems engineering [8, 9, 10] and implemented a strategy of phased progress. Our strategy started with assembling a multidisciplinary project team. The project team first conducted a systematic literature review to become informed on the state of mobile user interface design in the domain of survey instruments. Based on the literature review, the team drafted an initial plan for developing standards and guidelines and then improved the plan through multiple iterations of extensive discussion among team members. In the following sections, we will go through this workflow and discuss technical details in each phase.

2.1. Project team

Based on the multidisciplinary nature of this project, we assembled a team consisting of the following roles: a project manager, survey methodology researchers, user researchers, human factors researchers, user interface designers, mobile application developers, and statisticians. One team member may assume one or multiple roles. One researcher also served as a technical lead. The project manager and the technical lead co-directed the project. The project manager was responsible for planning, progress monitoring, budget, and coordination. The lead researcher was responsible for research design and implementation. Table 1 shows the functions covered by each role.

Table 1. Project team members' roles and functions

Role	Function
Project manager	Project planning and monitoring
Lead researcher	Research design and project implementation
Survey methodology researcher	Mobile survey user interface design
User researcher, Human Factors researcher	User experience evaluation, usability testing
User interface designer	Mobile user interface design
Mobile software developer	Mobile software development
Statistician	Statistical analysis

2.2. Literature reviews

The purpose of the literature review was to obtain the current knowledge of mobile device user interface design and mobile survey instrument design, and to identify areas where additional information is needed. Considering the fact that software user interface design and survey instrument design are traditionally two disciplines with little overlap, we decided to conduct two literature reviews: one on basic user interface design for mobile devices – features like touch target size, and the other on mobile survey instrument design – features like response options. To ensure that the review was to be as objective as possible, we adopted the methodology of systematic literature reviews. For both reviews, we first defined the scope of review, information sources to be searched, literature search period, and search terms. Table 2 summarizes the parameters used in the two reviews. The reviews were limited to the literature in English language.

For the review of basic mobile user interface design, we conducted four rounds of search with different combinations of search terms as shown in Table 2. For each round, the abstracts of the first 50 articles in the search results were independently screened by two reviewers in an attempt to select appropriate articles for further review. The two reviewers compared their selections and reconciled their differences. For each agreed-upon article, one of the two reviewers conducted a full article review. The review revealed the following areas that may impact the quality of survey response [e.g., 11, 12, 13, 14, 15]: touch target size, touch target location, touch target spacing, navigation, color vision, perceived aesthetics.

Likewise, similar methods were used to conduct the review of mobile survey instrument design. The review highlighted areas that may affect how a respondent completes a mobile survey, including layout of questions and response options, response types (e.g., radio button, check box), grid design (a type of question and response options design), items grouping [e.g., 16, 17, 18].

Table 2. Literature search parameters

Literature search parameters	Review on mobile user interface	Review on mobile survey user interface
Review scope	User interface design for mobile devices	Questionnaire design for mobile Web
Information sources	Guidelines; Web sites of major industry, professional, and academic institutions; Computers & Applied Sciences Complete; PsycINFO; Google Scholar; EBSCO Host	Web survey methodology bibliographic database: www.websm.org
Search period	2000-2015	2007-2015
Search terms	mobile user interface design, mobile interface design, smartphone user interface design, smartphone interface design	mobile web, smartphone web, mobile web surveys, smartphone web surveys

Based on the results of the literature review, we identified a list of user interface components (Table 3) that are pertinent to mobile survey instruments. Those components were classified into two groups. One group concerns the basic operations of a mobile device, such as touching a target icon. The other group is associated with survey questionnaire design layout, such as the orientation of response options. We consider the former group to be the basic mobile user interface elements, while the latter group is comprised of larger building blocks that are made up of the basic elements. Based on the functional characteristics of the two groups, we decided that the user interface design framework would consist of two parts: standards for the basic elements of mobile device operation, guidelines for the building blocks of mobile survey instruments. The standards are mandatory for the software developers to adhere, whereas the guidelines are recommendatory.

Table 3. User interface variables pertinent to mobile survey instruments

Basic mobile user interface elements	Width of a square target touch area
	Space between adjacent touch areas
	Character size
	Typeface
	Height of text field box
	Radio button diameter
	Check box width
	Foreground/background contrast in gray scale
Mobile survey user interface components	Date input format
	Maximum number of response options
	Optimal orientation of response options
	Optimal format of grids
	Optimal format of navigation buttons

2.3. Mobile user model

We started the process of developing the standards and guidelines by constructing an Information Processing Model of Mobile Device Operation (MoDO), as shown in Figure 2, and corresponding Mobile User Model (MUM), as outlined in Table 4.

The MoDO illustrates the information flow from perceiving visual information displayed on a mobile device screen to the human brain - a presumed information processing device – and to the index finger that touches the target. In MoDO, three critical factors have direct implications for survey data collection: a respondent's vision, index fingertip size and movement, and cognitive ability.

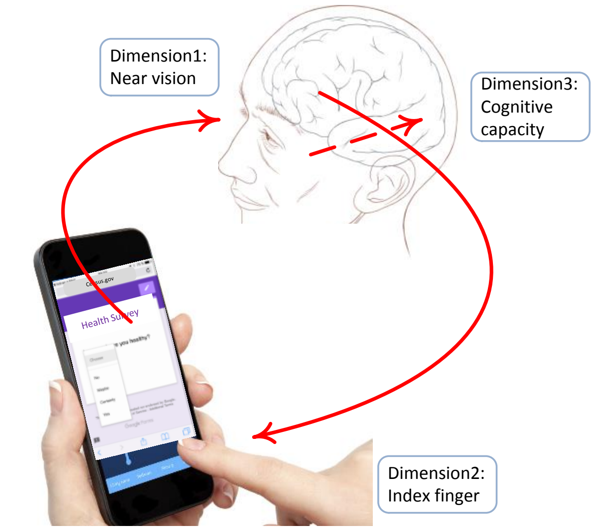


Figure 2. Information processing model of mobile device operation (MoDO)

We constructed a MUM with relevant physical and mental characteristics as defined in Table 4, i.e., vision, index fingertip and finger movement, and cognitive capacity. The rationale for using this MUM is that, if a person like MUM can successfully complete a mobile survey whose design is based on the standards and guidelines developed in this project, anyone who has better physical and mental capabilities than the MUM can do at least as well. Our respondent population is the U.S. population. This MUM represents a less educated and relatively senior person, for example, a 60-year-old person.

Table 4. Characteristics of Mobile User Model

Dimension I: Near vision (for reading)	Binocular habitual visual acuity – around 20/20
	Normal contrast sensitivity
	Color blind
Dimension II: Index finger	Index fingertip touch area breadth – 13 mm
	Index finger mobility – not good but able to operate touch screen device
Dimension III: Cognitive ability	Language – fluent English speaking
	Education – Eighth grade or equivalent

2.4. Standards development

Standards are being developed for three categories of basic mobile user interface elements: (1) touch target size and spacing, (2) text entry and display, and (3) graphics luminance and color. For categories 1 and 2, the following approach is taken to establish the standards: (1) Define a list of basic mobile user interface elements; (2) define respondent performance metrics for each element; (3) design behavioral experiments for mobile user interface elements; (4) conduct the experiments to collect respondent performance data; (5) perform statistical analysis of respondent performance data; (6) establish standards based on the results of the statistical analysis. For category 3, graphics luminance and color, standards are to be formed through literature review because considerable research has been done in this area and evidences are available. Table 5 lists seven variables for which standards are to be established through collecting empirical evidence. The seven variables cover all the basic user interface elements to be used in mobile survey instruments. For example, a square target may represent either a check box or a touch target.

Table 5. Basic mobile user interface parameters for standards

Parameters / Parameter Combinations	Standard to be established
Width of square touch target	Minimum width of a square touch target
Spacing surrounding a square touch target	Minimum space surrounding a square touch target with a given width
Diameter of round touch target	Minimum diameter of a round touch target
Spacing surrounding a round touch target	Minimum space surrounding a round touch target with a given diameter
Height of a text field	Minimum height of a text field
Vertical spacing above and below a text field	Minimum vertical space above and below a text field
x height of text display	Minimum x height of text display

Five human behavioral experiments were designed to generate empirical evidence for five parameters: the optimal combination of square touch target and spacing (Experiment 1), the optimal combination of circle touch target and spacing (Experiment 2), text field height for text entry (Experiment 3), text field height for text editing (Experiment 4), and character height for text display (Experiment 5).

For each experiment, we defined experimental factors (independent variables) and respondent performance metrics (dependent variables), designed an experimental paradigm, developed a data analysis strategy, and specified participant inclusion criteria based on the MUM. In the following paragraphs, we will use Experiment 1 as an example to describe these major experimental design components.

The purpose of Experiment 1 is to establish optimal combinations of square touch target width and spacing between targets. For example, for a square target with the width of 6 mm, the experiment will determine the minimum space between two targets such that a MUM could successfully touch this particular target without touching any neighboring targets with a success rate of at least 90%.

Two experimental factors are target width (Width) and space surrounding the target (Space). There are 10 levels of width, ranging from 2 mm to 11 mm, with 1-mm increment. For each target width, there are certain levels of spacing, starting from 0 mm and, with 1-mm increment, ending with a value M dependent on the target width as defined in this equation:

$$\text{Width} + (M \times 2) = 17 \text{ or } 18$$

There are total of 65 combinations of target width and space.

Three trial-level metrics are used to measure participant performance: success (Success), task completion time (Time), and perceived task difficulty (Difficulty). See Table 6 for the definition of each measure.

Table 6. Performance measure definitions

Metric (Abbreviation)	Definition
Success (Success)	A dichotomous indicator, 1 = hitting a target, 0 = missing a target
Task completion time (Time)	Duration between the onset of task and the end of the task
Perceived task difficulty (Difficulty)	Subjective rating of perceived difficulty level in performing the task, using a 5-point rating scale: *very easy*, *easy*, *neither easy nor difficult*, *difficult*, *very difficult*

The experiment calls for the participant to perform a touching task. The task starts with the participant holding a smartphone with one hand and placing the index figure of the other hand at a starting point on the smartphone screen. After the index finger is placed at the starting point, a square target imbedded in a 4x4 array of squares of the same size appears on the screen. As soon as seeing the touch target, the participant moves the index finger to touch the target. Once the finger lands on the screen, the array disappears and the task is complete.

Figure 3 shows a sketch of the array design. One performance of the aforementioned task consists of one trial. There are 260 trials in this experiment. The first 65 trials include all the combinations of touch target and space as defined earlier. The presentation order of the 65 combinations is randomly generated and then that order is repeated three more times, with the target square occupying a different location each time, amounting a total of 260 trials. The same randomization order and presentation is used for all participants.

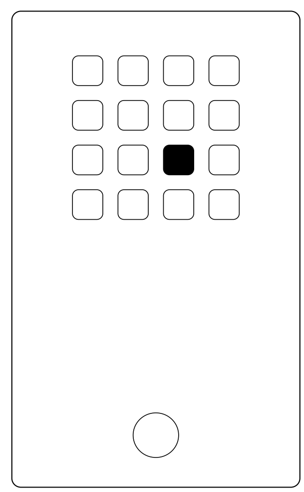


Figure 3. A sketch of the array design. A touch target could appear in one of the middle four locations. The solid square in the array represents a touch target.

Data analysis includes plotting a performance measurement curve for each target width as a function of space, and then identifying combinations of target width and space at which 90% of the participants perform successfully. Figure 4 shows the pilot data from one participant.

Since there is no historical data that could be used to estimate statistical power for participant sample size, we decided to use a sample size of 30 participants, which is usually considered a large sample in psychophysics research [19]. The sample has a balanced gender distribution of 15 males and 15 females to minimize potential gender bias in performance measures. Based on the MUM, we specified inclusion criteria for the participant sample: A participant would be between 60 and 70 years old, have binocular habitual near-vision around 20/20 and normal luminance contrast sensitivity (i.e., good vision for reading), have education of 8th grade or equivalent, speak fluent English, and have been using a smartphone for at least 12 months.

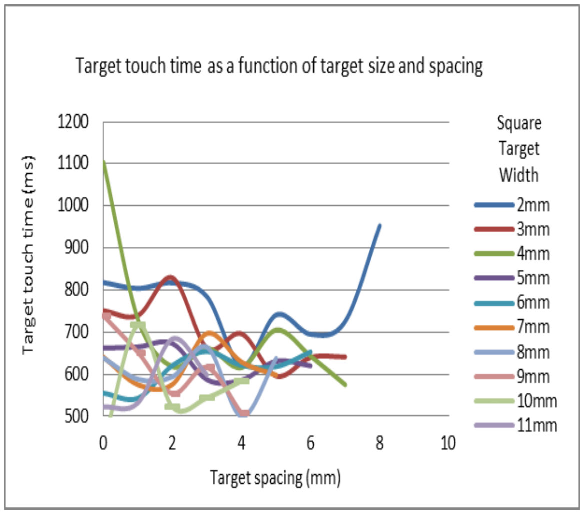


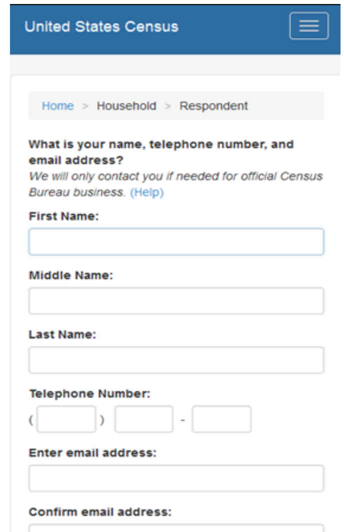
Figure 4. Pilot data of target touching time as a function of target size and spacing (one participant).

2.5. Guidelines development

A typical screen of a mobile survey instrument can be divided into these parts: question instruction(s), question stem, response options, navigation, supporting features, and general features that are applicable across the entire screen. Figure 5 is a screen shot of a mobile survey instrument that shows some of the parts.

Figure 6 depicts the workflow of guideline development, which will be described in details below.

Identifying topics. The first step we took in developing the guidelines was to identify the topics for which guidelines are in need. We iteratively conducted several brainstorming sessions among the team members to generate and select topics. In the first brainstorming session, each team member proposed a "wish list" of topics. After the meeting, the technical leader consolidated the proposed topics and drafted a new list of topics. In the second brainstorming session, the team reviewed the new list and decided a working list of topics. Topics were classified into 6 categories corresponding to the aforementioned screen parts. Table 7 lists the 6 categories and a sample topic for each category.



Source: 2016 Census Test

Figure 5. A sample mobile survey instrument screen.

Formulating research questions. For each topic, we formulated a specific research question. The answer to the research question becomes the guideline (See the example in Table 8). The research question is a type of comparison question which takes a form of "Is design A better than design B?" or "Is design A the best among all the designs?" The essence of generating a guideline is a series of comparative analyses through which a best available solution emerges.

Table 7. Topics for guidelines development

Category	Sample Topic
Question instruction	Layout
Question stem	Text color
Response	Response options orientation
Navigation	Optimal navigation method
Support features	Within-question Help link
General	Text-Field Labeling

Generating guidelines. The guidelines must be based on empirical evidence. We developed a two-step strategy to collecting the evidence. First, existing evidence is to be gathered through literature review. If no sufficient evidence is found, empirical studies will be conducted to collect evidence. To control guideline quality, we used a three-grade evidence strength rating system: two or more peer-reviewed studies being strong evidence, single peer-reviewed study in conjunction with at least two non-peer-reviewed reports being moderate evidence, two or more non-peer-reviewed reports being weak evidence. A guideline must be supported by at least two studies.

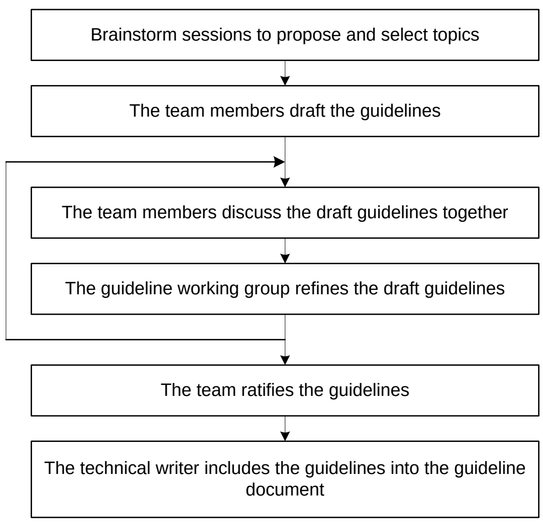


Figure 6. Guidelines development workflow.

To facilitate consistent practice across team members in this collaborative effort, we adopted a protocol template to document all the artifacts associated with each guideline development. Table 8 illustrates a sample template filled with information. Each guideline will also include an example to illustrate the use of the guideline (no shown in Table 8).

3. Summary

This paper described an approach to conducting a complex research and development project in a government environment. The approach highlights the importance of systems engineering principles, of multidisciplinary collaboration, of scientific rigor, and of practical problem-solving. This study is exploratory in nature, more work is to be done and more improvement is warranted.

Table 8. A sample template for guideline development

Guide No	1
Importance	2
Category	General
Topic	Screen Orientation
Research questions	Should the survey instrument be designed for portrait or landscape display?
Guide	Design for portrait as most people hold their smartphone upright in portrait view rather than sideways in landscape view.
Evidence	Sahami Shirazi, A., Henze, N., Dingler, T., Kunze, K., & Schmidt, A. (2013, August). Upright or sideways?: analysis of smartphone postures in the wild. In Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services (pp. 362-371). ACM.
Evidence strength	2
Example	*(Omitted)*

4. Acknowledgement

TBD.

5. References

[1] A. Smith, "U.S. Smartphone Use in 2015", http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/, Pew Research Center, Washington DC, April 1, 2015, pp. 1-1.

[2] Groves, R.M., F.J. Fowler, M.P. Couper, J. M. Lepkowscki, E. Singer, and R. Tourangeau, Survey Methodology, Wiley, New Jersey, 2009.

[3] E. Frøkjaer, M. Herzum, K. Hornbaek, "Measuring Usability: Are Effective-ness, Efficiency, and Satisfaction correlated?", Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, New York, 2000, pp. 345-352.

[4] Couper, M.P., Designing Effective Web Surveys, Cambridge University Press, Cambridge, 2008.

[5] T. Haley, B. Ireland, E.Wojtaszek, D. Nash, and R. Dion, Raytheon Electronic Systems experience in software process improvement. Technical Report CMU/SEI-95-TR-017, Carnegie-Mellon Software Engineering Institute, Nov 1995.

[6] K. Butler and W. Lipke, Software process achievement at Tinker Air Force Base. Technical Report CMU/SEI-2000-TR-014, Carnegie-Mellon Software Engineering Institute, Sep 2000.

[7] First CeBASE eWorkshop: Focusing on the cost and effort due to software defects. NSF Center for Empirically Based Software Engineering, http://www.cebase.org/www/researchActivities/defectReduction/eworkshop1/, Mar 2001.

[8] International Council on Systems Engineering. What is Systems Engineering? http://www.incose.org/practice/whatissystemseng.aspx. Accessed June 10, 2016.

[9] A Guide to the Project Management Body of Knowledge, 5th Ed. Project Management Institute, Newtown Square, PA. 2013.

[10] Humphrey WS. Managing the Software Process. Addison-Wesley, Boston, USA. 1989.

[11] P. Parhi, A.K. Karlson, and B.B. Bederson, "Target Size Study for One-Handed Thumb Use on Small Touchscreen Devices",MobileHCI'06, ACM, Helsinki, Sep 2006.

[12] S. Komine and M. Nakanishi, "Optimization of GUI on Touchscreen Smartphones Based on Physiological Evaluation – Feasibility of Small Button Size and Spacing for Graphical Objects", Lecture Notes in Computer Science, Springer, 2013, pp. 80-88.

[13] L. Jefferson and R. Harvey, "An Interface to Support Color Blind Computer Users", CHI 2007, San Jose, California, USA, April 28–May 3, 2007.;

[14] C. Travis and P. Murano, "A comparative study of the usability of touch-based and mouse-based interaction", International Journal of Pervasive Computing and Communications, Emerald Group Publishing Limited, 2014, pp. 115-134.

[15] H. Hwangbo, S.H. Yoon, B.S. Jin, Y.S. Han, and Y.G. Ji,"A Study of Pointing Performance of Elderly Users on Smartphones", Intl. Journal of Human–Computer Interaction, Taylor & Francis Group, 2013, pp. 604–618.

[16] F. Funke, "HTML5 and mobile Web surveys: A Web experiment on new input types", LINK Institut für Markt- und Sozialforschung, Frankfurt. March 2013.

[17] P. Lugtig and V. Toepoel, "The Use of PCs, Smartphones, and Tablets in a Probability-Based Panel Survey: Effects on Survey Measurement Error", Social Science Computer Review, SAGE, 2016, pp. 78-94.

[18] A. Mavletova, "Data Quality in PC and Mobile Web Surveys", Social Science Computer Review, SAGE, 2013, pp. 725-743.

[19] Hays, W.L., Statistics, Holt, New York, 1988.