

A Coding Framework for Usability Evaluation of Digital Health Technologies

M. Ebnali^{1,2}, L. R. Kennedy-Metz^{3,4}, H. M. Conboy⁵, L. A. Clarke⁵, L. J. Osterweil⁵, G. Avrunin⁵, C. Miccile¹, M. Arshanskiy⁴, A. Phillips⁴, M. A. Zenati^{3,4}, R. D. Dias^{1,2}

¹ Human Factors and Cognitive Engineering Lab, STRATUS Center for Medical Simulation, Brigham and Women's Hospital, Boston, MA, USA

² Department of Emergency Medicine, Harvard Medical School, Boston, MA USA

³ Medical Robotics and Computer Assisted Surgery Lab, Division of Cardiac Surgery, VA Boston Healthcare System, Boston, MA, USA

⁴ Laboratory for Advanced Software Engineering Research, Manning College of Information and Computer Sciences, University of Massachusetts Amherst, MA, USA
rdias@bwh.harvard.edu

Abstract. Several studies have reported low adherence and high resistance from clinicians to adopt digital health technologies into clinical practice, particularly the use of computer-based clinical decision support systems. Poor usability and lack of integration with the clinical workflow have been identified as primary issues. Few guidelines exist on how to analyze the collected data associated with the usability of digital health technologies. In this study, we aimed to develop a coding framework for the systematic evaluation of users' feedback generated during focus groups and interview sessions with clinicians, underpinned by fundamental usability principles and design components. This codebook also included a coding category to capture the user's clinical role associated with each specific piece of feedback, providing a better understanding of role-specific challenges and perspectives, as well as the level of shared understanding across the multiple clinical roles. Furthermore, a voting system was created to quantitatively inform modifications of the digital system based on usability data. As a use case, we applied this method to an electronic cognitive aid designed to improve coordination and communication in the cardiac operating room, showing that this framework is feasible and useful not only to better understand suboptimal usability aspects, but also to recommend relevant modifications in the design and development of the system from different perspectives, including clinical, technical, and usability teams. The framework described herein may be applied in other highly complex clinical settings, in which digital health systems may play an important role in improving patient care and enhancing patient safety.

Keywords: Usability Study, Decision Support System, Digital Health.

1 Background

With the rise of medical information and technological advancements, digital health technology offers a variety of benefits to support healthcare providers and patients, improving compliance with standards of health quality, cost, and practice [1]. Therefore, accelerating the acceptance and engagement of digital health technology has been recognized as a national policy priority [2]. In 2009, the Health Information Technology (IT) for Economic and Clinical Health (HITECH) Act under the American Recovery and Reinvestment Act was initiated to expedite the coordination and delivery of American healthcare through health IT, including the adoption of digital health technologies.

Although a growing body of literature suggests that higher adoption of digital health systems is associated with safer and higher-quality care, [3] several pieces of evidence report unintended adverse effects of these technologies on clinical workflows due to poor usability issues [4]. The design of digital health technology is a complex process because of the inherently complex nature of clinical procedures that are mostly characterized by dynamic, non-linear, interactive, and interdependent collaborative activities, with uncertainty in outcomes[5]. Dealing with this complexity demands following an extensive set of design and usability requirements, constraints, and safety measures. Poor usability of digital health technology may result in substantial increases in medical error and associated costs, decreased efficiency, and unsatisfied users [6]. Recent reports have highlighted the importance of using cognitive engineering and human factors approaches for the design and empirical assessment of technology used in clinical settings [7].

Focus group and interview techniques have been widely used as common qualitative approaches to capture users' feedback and to obtain in-depth insights on the usability of digital health technologies [8]. Focus group and semi-structured interview discussions are carefully planned and designed to obtain the perceptions of the individuals on a defined area of interest, facilitated by a moderator to keep the focus of the discussion. In these methods, participants are invited to discussion sessions to communicate their comments voluntarily in a safe and supported manner. This approach may lead to uncovered issues that researchers might have been unable to plan in advance. Moreover, focus group methodology is a cost-efficient way of evaluating user experience, as several subjects can be interviewed at the same time.

2 Current Challenges

Focus group methods have many advantages over other usability evaluation methods, however, as with any research methodologies, there are limitations. Some of these limitations can be overcome by well-structured planning and moderation, but other issues are inevitable and unique to this approach. Despite the widespread use of these qualitative methods, few guidelines exist for analyzing the data gathered from participants. Compared to structured questionnaires and quantitative experimental approaches, data collected during usability focus groups and interview sessions are often diffi-

cult to assemble and analyze. Annotating and coding qualitative data can be time-consuming and complicated for most digital health technologies given their inherent complexities. Although extensive previous literature provides frameworks and guidance on designing and conducting usability focus groups and interview sessions [9], scarce literature exists on extracting and coding usability data to effectively inform system design and development improvements based on a human-centered approach.

3 Proposed Coding Framework

Few studies evaluating the usability of digital health technology have attempted to combine qualitative coding with usability theories and principles [10]. Although these studies provide insights on the categorization of usability issues, there is still no standard methodology on how to integrate usability principles in coding this type of qualitative data. In this study, we aimed to develop a coding framework for the systematic evaluation of user feedback comments generated during focus group sessions, underpinned by fundamental usability principles and design components. Integrating usability principles and design components in coding and analyzing data generated during focus group sessions may help to systematically improve the degree of shared understanding between users with different roles, identify the extent of overlap in their comments, and the communication content generated during the sessions. Developing a shared understanding of the usability components of digital health technology can better ground the system on effective communication, leading to improved outcomes related to patient care coordination, teamwork, and care continuity.

4 Codebook Development

In usability studies, there are several well-established principles that can be utilized to categorize issues and comments collected during focus groups. Most studies attempting to evaluate usability via focus groups suffer from poor reproducibility of evaluations due to variation and subjectivity in codes, and a lack of standard reporting [10]. To tackle this challenge, we developed a codebook system that systematically evaluates each comment against usability principles for digital health technology. Moreover, since usability focus group sessions are primarily focused on different parts of a system interface and its specific features, this codebook incorporates categories related to design functions and elements. Furthermore, we included a coding category to capture the user's clinical role associated with each specific comment, offering a better understanding of role-specific challenges and perspectives, as well as the level of shared understanding across the multiple clinical end-users.

The codebook allows researchers to code each users' feedback based on three main categories guided by prompt questions: *a) Usability Components*: Which usability principle is targeted by the feedback? *b) Design Components*: Which design function and element are targeted by the feedback? *c) Clinical Role*: Which clinical role provided the feedback? The following sections provide details on each of these categories.

ries, helping researchers to categorize comments extracted during focus groups targeting digital health technology usability.

4.1 Usability Principles: Which usability principle is targeted by the feedback?

The main goal of usability focus group sessions for qualitative studies is extracting information about system issues and gathering insights on a prototype and/or system. Based on the human-centered design perspective, the Ten Heuristic Principles developed by Jakob Nielsen [11] have been identified as standard principles to identify and categorize usability issues, including 1) Visibility, 2) Match, 3) Control, 4) Consistency, 5) Error, 6) Recognition, 7) Flexibility, 8) Aesthetic, 9) Recovery, and 10) Help.

4.2 Design Components: Which design function and element are targeted by the feedback?

Usability principles alone are not sufficient to accurately categorize the issues associated with the usability of digital health technology, and other significant factors related to system design need to be integrated into the coding system. Qualitative usability studies aim to collect insights about how users interact with the product or service. This interaction, as described by scholars in the design community, happens at the interface level or the product ‘front-end’ which enables user interaction through communication and conversation. Therefore, this codebook incorporates usability principles to categories related to design functions and design elements [12].

Design Functions. Most of the comments expressed during usability focus group sessions target a particular user interface [12]. Although these interfaces come in various forms such as buttons, pop-ups, radio buttons, panels, etc., they can generally be associated with a higher-level category of components. To date, though there are some available taxonomies, there is no well-accepted framework to categorize user interfaces [13]. For example, in previous research, *Kamaruddin et al.* proposed that the types of interface design consist of four main categories with separate features: presentation interface, conversation interface, navigation interface, and explanation interface [14]. According to these schemas, we considered the following categories for coding comments expressed during usability focus group sessions: input controls components, navigational components, informational components, and container components. These categories have been used frequently in the design community in order to describe the support provided to users by each individual interface during specific tasks. Each category encompasses user interfaces with common functions in the system (Table 1). This categorization can be helpful for analyzing the results from focus groups and interview sessions where the research team will gain better insights into function-wise usability issues in particular. Moreover, evaluation of function-wise usability issues across various iterations can be helpful to better compare versions of designs in various interactions throughout the prototype life cycle.

Table 1. Design functions of user interfaces.

Component	Description	Examples
1. Input controls	Allow users to input information into the system.	A button allowing the user to select options
2. Navigational	Enable users to move around a system or a website.	Tab bars Scroll bars Next/ back UIs
3. Informational	Used to share information with users.	Descriptions Icons, feedback Pop-up Messages
4. Container	Designed to hold related content and interfaces together.	Image carousel Frame of a window

Design Elements. We have also incorporated design elements into the codebook to have a more elaborated view of each feedback from a design perspective. Previous studies have established various design elements based on different perspectives and research frameworks. Most of these elements focused on the fundamental design components of the interface, which target a wide range of visual, audio, and content aspects. Based on these design foundations, a list of principles and elements are categorized into a shorter list of design elements [13]: color, imagery, typography, language, location, and audio. In line with these studies, the healthcare usability literature also suggests that color, imagery, position, and text style are the main design elements in digital health technology, which contribute to the ability of the user to accurately interpret and use the interface [15]. Incorporating these elements in the coding process of data generated during focus group sessions can help researchers to have a more accurate and detailed evaluation of users' feedback. Each comment can be evaluated based on these six basic design elements (Table 2).

Table 2. Design elements of user interfaces.

Component	Description
1. Color	One of the most imminent elements of a design; It is used to differentiate items, create depth, add emphasis, and/or help organize information. It can stand alone, as a background, or be applied to other elements, like imagery or typography.
2. Imagery	Can be in different styles: shapes, illustrations (image, video, animation), 3D renderings, etc. Defined by boundaries, such as lines or color, they are often used to emphasize a portion of the page.

3.	Typography and Text Style	Can be used in different ways in the context of an app or a website and mainly refers to which fonts are chosen, their size, alignment, and spacing.
4.	Language	Covers the meaning and tone of words used in the product.
5.	Position	Can significantly impact the usability of a system such as the readability of design.
6.	Sound	Used to notify the user about a situation or for avoiding hazardous events.

4.3 Clinical Role: Which clinical role provided the feedback?

In each focus group, participants with different clinical roles and backgrounds were invited to the sessions. Getting their unique perspective and integrating it with data collected from other participants can be useful for creating a shared understanding of usability issues across different parts of the design. The clinical role was considered as a code category for the person in that role providing feedback.

The codebook described in the previous sections allows researchers to code each participant's comment based on three main categories guided by interviewers' prompts. Adhering to this systematic coding framework allows researchers to assess saturation of usability issues in general, across design components, and roles in particular. Furthermore, since usability focus group data is analyzed one group at a time, following a systematic coding approach enables researchers to aggregate the findings across multiple focus groups, informing design modifications. It also allows the quantification and analysis of system issues by specific usability principles and/or roles, which may identify areas of design and development to focus on in future iterations.

4.4 Data-driven Design Modifications

After categorizing the user comments based on usability principles, design components, and clinical roles, researchers and design and development teams can frame a list of potential changes to address the raised issues. Due to limitations in time and costs, implementing all the changes is often not feasible. The prioritization of which system modifications should be made is one of the most important challenges in designing and/or re-designing a digital health technology. As a part of the proposed framework, we created a voting system to facilitate this process and quantitatively inform system modifications, as well as their level of prioritization (Table 3).

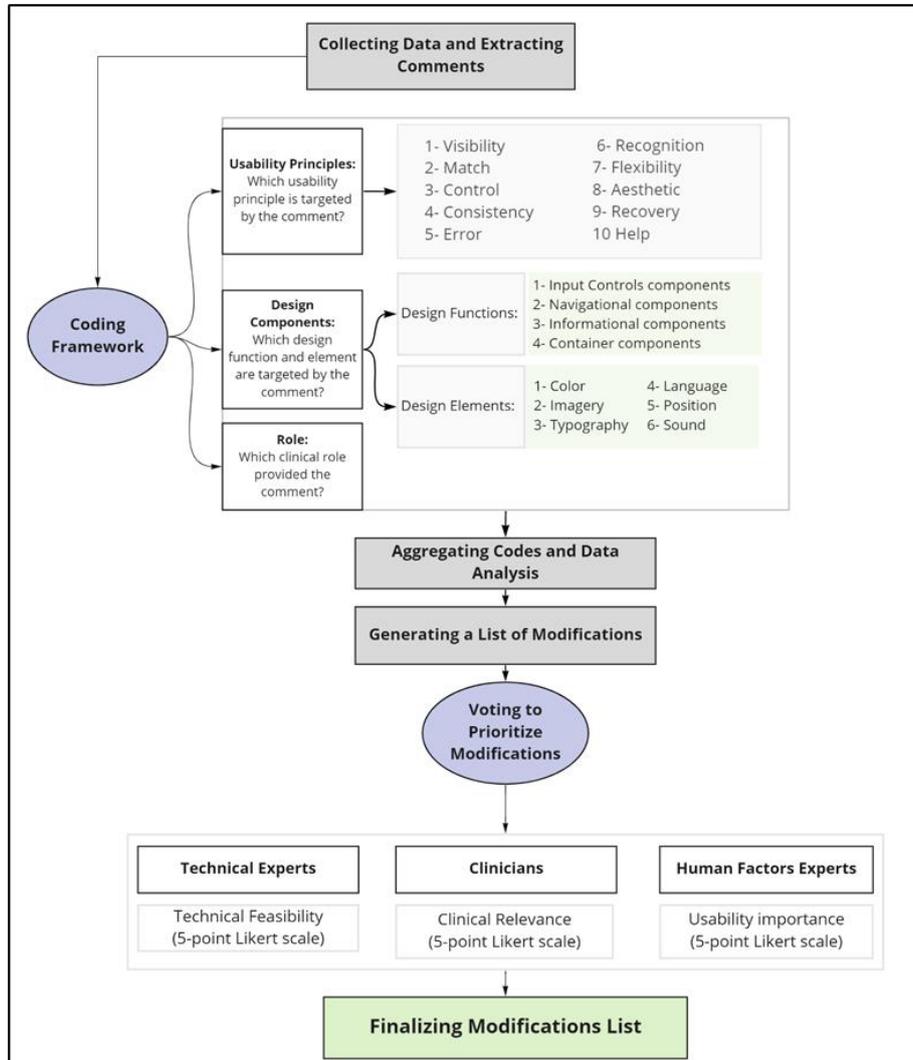


Fig. 1. Main categories and components of the proposed coding framework

This voting system evaluates each suggested modification based on three criteria (usability importance, clinical relevance, and technical feasibility), using a 5-point Likert scale. An advantage of this voting system is that it can incorporate multidisciplinary aspects from various experts, including not only the technical design and development team but also clinicians and human factors analysts.

Table 3. Decision criteria and rating for prioritization.

Criteria	Rating anchors	Description
Usability Importance	1: Very Important 2: Important 3: Moderately Important 4: Slightly Important 5: Not Important	Whether implementing the proposed change may prevent the user from completing a task or properly accessing information.
Clinical Importance	1: Very Important 2: Important 3: Moderately Important 4: Slightly Important 5: Not Important	Whether implementing the proposed change may negatively impact clinical goals and/or workflows.
Technical Feasibility	1: Very Feasible 2: Feasible 3: Moderately Feasible 4: Less Feasible 5: Not Feasible	Whether implementing the proposed change is feasible in terms of implementation time and costs.

5 Use Case

We applied the coding framework to evaluate an electronic cognitive aid (Smart Checklist) that was developed to guide cardiac surgery teams during common cardiac procedures in the operating room (OR) [16]. The Smart Checklist uses a carefully elicited and domain expert-validated process model to monitor the progress of an ongoing surgical procedure, determining the expected next tasks for each of the team members, thereby providing a context- and patient-specific perspective on each team role's task management. This study was approved by the Mass General Brigham (MGB) Institutional Review Board (IRB) and all research subjects completed an informed consent procedure.

5.1 Data Collection

Four multidisciplinary focus group sessions were conducted with OR team members, representing key cardiac surgery roles: cardiac surgeons, cardiac anesthesiologists, perfusionists, and OR nurses. Two interviewers (LKM and H.C) conducted focus group sessions via videoconferencing while demonstrating the Smart Checklist. HC presented a detailed demonstration of the Smart Checklist to convey its primary informational features, where they were located, and the various ways in which users could interact with the interface (e.g., clicking buttons, entering notes). Pre-determined prompt questions were interspersed by the researchers to gather targeted feedback from participants. User's comments were transcribed independently *post*

hoc by both interviewers after review of the focus groups recordings. For example, Fig 2 shows specific parts of the Smart Checklist in which participants were asked to discuss their preference over two alternative views: separated view vs. merged view. The merged view shows the next tasks for all teams in a single column and the separated view has different columns for each team or role.

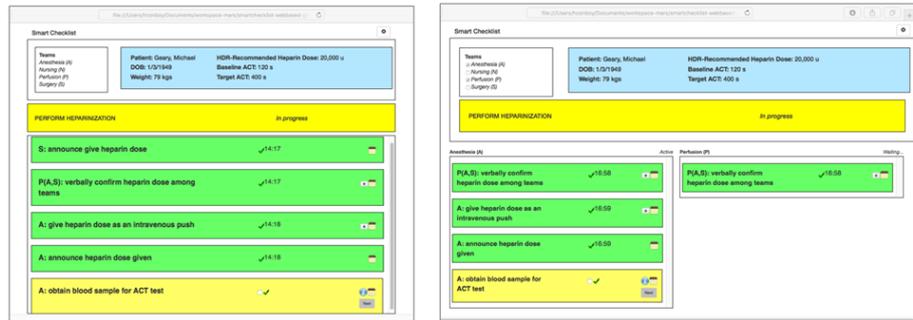


Fig. 2. Different views of the surgical steps within the Smart Checklist

5.2 Feedback Coding

Two independent coders (L.KM and H.C) used the coding framework to analyze the qualitative data generated in the focus groups. The coding of the 118 transcribed sentences was conducted via *Dedoose*, a web application for mixed methods research. First, the coders were instructed on coding components by providing examples of various types of usability principles and design components. Discrepancies between the coders were discussed during sessions until a consensus was reached. Then, the coders were asked to code each sentence based on the codebook (Fig. 1).

5.3 Results

A total of 12 subjects participated in four focus group sessions. Group 1: surgeon, anesthesiologist, perfusionist, nurse; Group 2: anesthesiologist, perfusionist, nurse; Group 3: surgeon, perfusionist, nurse; Group 4: anesthesiologist, perfusionist. To assess the inter-rater reliability (IRR) of the coding framework, we calculated Kappa coefficients. Analysis across all coding showed a moderate IRR with a Kappa coefficient of 0.53, p -value < 0.001 , and an overall percentage of agreement between coders of 82.3%.

A total of 18 modifications were suggested by the focus groups. The priority voting system was completed by 3 human factor experts, 4 OR clinicians (1 attending cardiac surgeon, 1 attending cardiac anesthesiologist, 1 perfusionist, and 1 scrub nurse), and 3 technical designers/developers. Table 4 shows the suggested modifications with the respective scores (median) across all three criteria.

Table 4. Priority scores by human factors (HF), clinicians (CL), and technical experts (TE).

Suggested Modifications	HF	CL	TE
Enable interface to update color scheme according to user selection	2.0	3.5	1.0
Integrate checklist with the post-procedure document generator	3.0	4.0	3.0
Embed numeric inputs into the step itself, rather than or in addition to in the pop-up dialogue box	2.0	3.5	3.0
Embed additional requisites corresponding to a step into the step itself, rather than or in addition to appearing in the pop-up dialogue box	2.0	3.5	3.0
In the Separated Team View, adapt the column width to the number of specialty teams involved in the surgical process	4.0	3.5	3.0
Increase default font size	4.0	3.0	3.0
Better distinguish the team primarily responsible for a step with distinct icons, border styles, etc.	5.0	3.0	1.0
Integrate checklist with voice-based support	4.0	3.5	1.0
Allow user to decide if the hierarchical steps should be the same color (e.g., yellow) as the steps with checkmark buttons/icons below them or a less saturated version of that color (e.g., lighter yellow)	2.0	3.0	4.0
Fix process header in place so it doesn't scroll as the checklist advances	4.0	4.5	1.0
Keep hierarchical steps hidden by default, but allow users to show them as desired	4.0	3.0	5.0
Include system timers for relevant steps (e.g., 3-minute timer after heparin administration)	4.0	4.5	1.0
Update steps reading announce X to announce 'X'	3.0	3.0	4.0
Better differentiate the header when the process is in progress from the display of a step that is in progress	4.0	4.0	3.0
Replace the suitcase icon to more accurately reflect the requisites corresponding to a given step	4.0	4.0	4.0
Indicate all steps related to a reported problem in a distinct way (e.g., all steps have red borders)	4.0	3.0	4.0

Include the ability to switch between Merged and Separated Views	4.0	3.5	5.0
Add a 'help' button to display a legend	5.0	2.0	5.0

6 Limitations and Future Directions

Even though the coding scheme was established based on well-established usability principles and design components and was tested to code usability data of a digital health system, it cannot be guaranteed that it will aid in the coding of all possible usability issues. Future studies are needed to further validate this codebook and voting system in the design and development process of other digital health technologies in additional clinical settings. Moreover, the coding system is specifically created to code data generated during focus group sessions, and future studies should evaluate the applicability of this framework on data gathered through other usability methods, such as verbal protocol and structured questionnaires.

7 Conclusion

In this study, we report the development of a comprehensive coding framework for usability evaluation of digital health technologies. In addition to incorporating relevant domains, such as usability principles, design components, and clinical roles, we have also provided a structured voting system to inform the prioritization of system modifications. The use case involving an electronic cognitive aid in the cardiac OR showed that this framework is feasible and useful not only to better understand distinct usability aspects that may be suboptimal, but also to recommend relevant modifications in the system from various perspectives, including the clinical team. The method and framework described herein may be adopted and applied to other highly complex clinical settings, in which digital health systems may play an important role in improving patient care and enhancing patient safety.

Acknowledgment

This work was supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health Under Award Number R01HL126896 (PI: Zenati). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. No conflicts of interest were declared.

References

1. Trinkley KE, Blakeslee WW, Matlock DD, Kao DP, Van Matre AG, Harrison R, et al. Clinician preferences for computerised clinical decision support for medications in primary care: a focus group study. *BMJ Health Care Inform.* 2019;26: 0.
2. Baker A. Book: *Crossing the Quality Chasm: A New Health System for the 21st Century.* BMJ. 2001. pp. 1192–1192. doi:10.1136/bmj.323.7322.1192
3. Hessels A, Flynn L, Cimiotti JP, Bakken S, Gershon R. Impact of Health Information Technology on the Quality of Patient Care. *Online J Nurs Inform.* 2015;19. Available: <https://www.ncbi.nlm.nih.gov/pubmed/27570443>
4. Furukawa MF, Eldridge N, Wang Y, Metersky M. Electronic Health Record Adoption and Rates of In-hospital Adverse Events. *J Patient Saf.* 2020;16: 137–142.
5. Ebnali M, Shah M, Mazloumi A. How mHealth Apps with Higher Usability Effects on Patients with Breast Cancer? *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care.* 2019. pp. 81–84. doi:10.1177/2327857919081018
6. Guo C, Ashrafian H, Ghafur S, Fontana G, Gardner C, Prime M. Challenges for the evaluation of digital health solutions—A call for innovative evidence generation approaches. *npj Digital Medicine.* 2020;3: 1–14.
7. Zenati MA, Kennedy-Metz L, Dias RD. Cognitive Engineering to Improve Patient Safety and Outcomes in Cardiothoracic Surgery. *Semin Thorac Cardiovasc Surg.* 2020;32: 1–7.
8. Brown W, Yen P-Y, Rojas M, Schnall R. Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology. *Journal of Biomedical Informatics.* 2013. pp. 1080–1087. doi:10.1016/j.jbi.2013.08.001
9. Leech NL, Onwuegbuzie AJ. An array of qualitative data analysis tools: A call for data analysis triangulation. *School Psychology Quarterly.* 2007. pp. 557–584. doi:10.1037/1045-3830.22.4.557
10. Willis G. Pretesting of Health Survey Questionnaires: Cognitive Interviewing, Usability Testing, and Behavior Coding. *Health Survey Methods.* 2014. pp. 217–242. doi:10.1002/9781118594629.ch9
11. Nielsen J. Enhancing the explanatory power of usability heuristics. *Conference companion on Human factors in computing systems - CHI '94.* 1994. doi:10.1145/259963.260333
12. Mosier JN, Smith SL. Application of guidelines for designing user interface software. *Behaviour & Information Technology.* 1986. pp. 39–46. doi:10.1080/01449298608914497
13. Kontio J, Lehtola L, Bragge J. Using the focus group method in software engineering: obtaining practitioner and user experiences. *Proceedings 2004 International Symposium on Empirical Software Engineering, 2004 ISESE '04.* IEEE; 2004. doi:10.1109/isese.2004.1334914
14. Mandel T. *The Elements of User Interface Design.* John Wiley & Sons; 1997.
15. Rind A, Wang TD, Aigner W, Miksch S, Wongsuphasawat K, Plaisant C, et al. Interactive Information Visualization to Explore and Query Electronic Health Records. *HCI.* 2013;5: 207–298.
16. Avrunin GS, Christov SC, Clarke LA, Conboy HM, Osterweil LJ, Zenati MA. Process Driven Guidance for Complex Surgical Procedures. *AMIA Annu Symp Proc.* 2018;2018: 175–184.