2009 Ninth International Conference on Intelligent Systems Design and Applications

Qualitative and Task Analytic Methods to Support Comprehensible Intelligent System Design

Melanie C. Wright, PhD, Noa Segall, PhD, Jonathan B. Mark, MD, Jeffrey M. Taekman, MD
Duke University Medical Center Department of Anesthesiology
Durham, NC, USA
melanie.wright@duke.edu

Abstract—**We argue that a critical component of designing comprehensible intelligent systems is finding the right applications for intelligence and designing intelligent solutions toward those applications. While we do not refute the value of good attention to later stages of human-centered design such as the application of human interface design principles and usability testing as methods for improving comprehensibility, there must also be significant attention to understanding problems in the context of use and how intelligence systems can best address those problems. In light of supporting naturalistic decision-making, we present a review of task analytic and qualitative research techniques that may be useful for better understanding problems in context that will support the design of more comprehensible intelligent systems.**

Keywords - human-centered design; knowledge elicitation; ecological interface design; situation awareness; grounded theory; naturalistic decision-making

I. HUMANS AS PROBLEM SOLVERS

Researchers studying humans in true work environments support a model of *naturalistic decision-making* in which the focus of the worker is not on comparing alternatives and choosing the best course of action, but is instead on the recognition of familiar situations and the generation of a workable plan for addressing the situation [1]. The *recognition-primed decision model* was developed by Gary Klein following 25 years reviewing critical incidents and interviewing experts in a wide variety of dynamic, safety-critical roles such as firefighters, military leaders, nuclear power plant operators, and nurses [1]. In the recognition-primed decision model, a decision-maker begins with a situation requiring action in a dynamic or a changing context. Based on the assessment of that situation, the decision-maker makes a judgment as to whether or not the situation is typical or familiar. For example, if an anesthesia provider notes mild hypertension and tachycardia at incision, she may immediately recognize this as insufficient depth of anesthesia. Based on this typical situation, she will assess her goals, important informational cues, and expectancies. For example, her goals may be to keep the patient comfortable, pain-free, and unaware. Important cues may include the relationship of current values of blood pressure and heart rate to baseline or expected values. She would expect the blood pressure and heart rate to return to baseline or expected values upon administration of more anesthetic agent. She will then act in a typical manner and increase the delivery of anesthetic agent. If the situation is not recognized as typical, the anesthesia provider will have to work a little harder at diagnosing the situation. She may seek out further information to identify features she can recognize or build a story that might fit the situation. She will then evaluate whether or not expectancies based on that assessment are met. For example, if the hypertension and tachycardia did not resolve with more anesthetic agent, she might examine the depth of anesthesia using a brain function monitor to determine whether the patient was unusually "resistant" to anesthesia and act appropriately.

Klein notes that when poor decisions are made in naturalistic environments, it is not the result of decision biases, but rather inadequate expertise or inadequate knowledge [1]. Successful and comprehensible intelligent systems will be those that are designed to complement rather than replace the exceptional contextual awareness, pattern recognition, and creative generalization capabilities of human workers. As technology advances, we become increasingly capable of designing systems that are more intelligent and context aware. However, research has shown that automation of high-level decisions and/or unreliable automation can lead to problems in human-system performance [2-4]. Unless the system can reliably replace any human involvement, human-system performance is best served when intelligence is used to enhance information presentation. The human operator can then take into account additional contextual information not available to the intelligent system and make an appropriate judgment or decision. Intelligence can be used to manipulate information salience in a number of ways. Examples include sorting the highest likelihood solutions to the top of a list or graying/fading unlikely options or old, unreliable information. Providing access to additional information such as the rationale to support different alternatives or the degree to which evidence supports each option allows the human user to then consider that information in light of additional information available.

II. HUMAN-CENTERED DESIGN APPROACH

The process of determining when, where, and how to use intelligence to enhance information presentation can be hard work. Human-centered design is an approach in which systems are designed around the needs and capabilities of the users, as opposed to being driven by the available technology. Design efforts that have incorporated a human-centered design approach have shown significant human-

978-0-7695-3872-3/09 $26.00 © 2009 IEEE
DOI 10.1109/ISDA.2009.173

system performance improvements in aviation, military systems, and health care environments [5, 6].

Variations of human or user centered design processes have been adopted by successful consumer product designers and system designers, with specific human factors methods integrated at appropriate points in the design process [7] (see Figure 1). These processes generally involve initial steps that include gathering information about potential users and the requirements of the users, followed by conceptual design, and then detailed design and prototyping. Throughout the process, human factors methods designed to gather information about user needs and human-system performance are used to optimize the human-system design and reduce the potential for errors from early conceptual design through final design details. Ultimately, system evaluation is completed within environments that incorporate critical contextual details that can influence the success or failure of system design.

This model places an emphasis on including users in a variety of steps in the design process. This does not mean that design should follow user preferences. A typical example would elicit input from users regarding functional requirements of a system, feedback from users on preliminary design, and users as participants who are observed in laboratory and field studies. Subjective feedback from users is important, but generally more important are the qualitative observations and objective measures or outcomes that are observed during task performance or system evaluation. While some companies may have fairly formal processes for user-centered design, it is generally flexible as a design method. Relevant human factors methods (e.g., a variety of knowledge elicitation, task analytic, and evaluative techniques) will be chosen based on the specific goals of the product or process design.

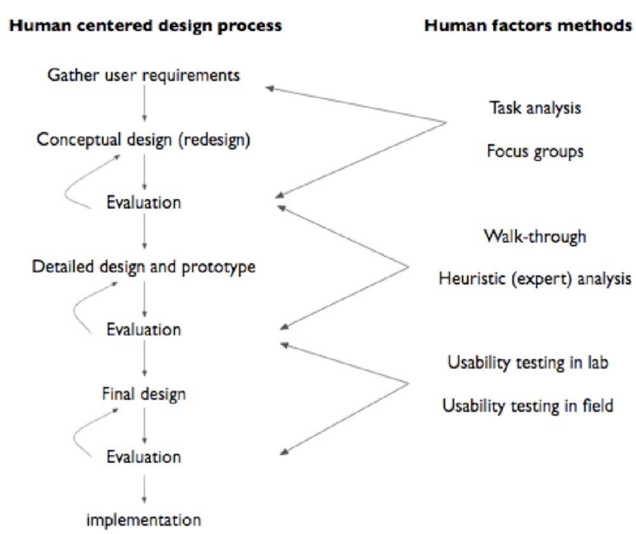

Figure 1. Human centered design process and example human factors methods

Although human-centered design places an important emphasis on early data collection and analysis of the work domain, human-system goals, or functional requirements, this important phase is frequently skipped or shortchanged in system or product design and development. It is uncertain whether this is because these techniques are not well known, are difficult to apply, or their benefits are under-appreciated. Without a clear understanding of information needs and appropriate applications for intelligence system design, researchers and designers can spend significant effort on applications that, for complex socio-technical systems, will have no practical application or comprehensibility.

Expert systems designers have historically used knowledge elicitation methods for acquiring and representing domain expertise that can then be translated into expert system rules or logic [8]. These methods include techniques such as interviews, protocol analysis, and observations. Expert system researchers have also developed and explored more automated methods of attaining knowledge, for example through the use of knowledge base tools, machine learning induction, and data mining [9, 10]. Integration of techniques such as interviews and data mining provide the advantage of combining higher levels of abstraction and intelligence associated with human operators with precise and explicit knowledge from data [9].

In this paper, we present approaches toward knowledge acquisition for the purposes of targeting where and when expert systems are most likely to be beneficial (as opposed to knowledge acquisition for the development of detailed rules or algorithms). These methods are intended to provide initial insight into the work environment and to identify promising areas in which to pursue more detailed expert system design. This is particularly important in the development of expert systems for complex problems such as intensive care and anesthesia decision support. In the following sections, we present two formal design processes that place significant attention on early attention to system goals and user requirements: ecological interface design and designing for situation awareness. We conclude with a third alternative: using a variety of human factors techniques for collecting rich data in context and analysis of this data using qualitative methods such as grounded theory or content analysis.

A. Ecological Interface Design

Perhaps the most commonly referenced formal human-centered design method is *ecological interface design (EID)* [5, 11, 12]. EID is a theoretical framework for the design of interfaces for complex human-machine systems [13]. An "ecological" approach to design is focused generally on the relationship between humans and the environment. The goal of EID, however, is more specific and is intended for the specific problem of designing human-computer interfaces for complex sociotechnical systems [13, 14]. The main purpose of EID was to support the "knowledge worker", whose primary function is to engage in intellectual work that requires discretionary decision-making, in adapting to change and novelty. EID was formulated originally for application in power plant process control.

EID draws on the theoretical concept of an abstraction hierarchy [15] that is used to represent constraints on the work domain. An *abstraction hierarchy* is intended to provide a hierarchical representation of the work domain or the work environment that provides several (usually around five) different levels that describe the work domain in a unique way. For example, Hajdukiewicz et al. present examples for the human body and cardiovascular system that map five levels of abstraction: purposes, balances, processes, physiology, and anatomy [12]. Each level is a different way of describing the same system. In addition, each level of the hierarchy places constraints on lower level functions. That is, there is a means-end relation between adjacent levels of the tree. EID emphasizes that the abstraction hierarchy is derived from a *work domain analysis* rather than a task based analysis and is independent of any particular worker, automation, event, task, goal, or interface.

Following the development of an abstraction hierarchy, EID then applies Rasmussen's skills, rules, and knowledge taxonomy [16] in communicating to the user the constraints defined by the work domain abstraction hierarchy. These concepts are used to guide system analysis and interface design using three general principles that support the three different levels of cognitive control [13, 14].

- To support skill-based behavior, users should be able to directly manipulate or act directly on the interface. Where operators must provide control, design the perceptual motor control aspect of the interface so that it maps to the intended action with minimal transformation.
- To support rule-based behavior, the interface should provide a consistent one-to-one mapping between work domain constraints and perceptual information. Object displays that integrate several directly measurable variables into a single, more meaningful (i.e., goal-relevant) variable are an example. A goal of this principle is to minimize situations where the interface mis-represents the true state of the system due to unknown interpretations or transformations between the system and the information presented.
- To support knowledge-based behavior, the work domain should be represented in the form of an abstraction hierarchy that would serve as an external mental model.

In general, the interface design should encourage use of the lower levels of cognitive control (skill- and rule-based behavior) since they involve fast, effortless processing that is less error prone, while supporting knowledge-based behavior that is crucial for novice users and for managing unexpected problems [13].

B. *Designing for Situation Awareness*

Situation awareness (SA) refers to an individual's awareness and understanding of the information available in his or her environment. Endsley has formally defined SA as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future" [17]. In Endsley's theoretical model, SA is a construct that is distinct from and proceeds decision making and the performance of actions [17]. SA as defined in this model includes only that portion of a person's knowledge that pertains to the state of a dynamic environment. Background knowledge, experiences, and established rules are static knowledge sources that fall outside of the definition of SA, though they may influence the SA development. Another key detail is that SA is continuously changing as the environment changes, either due to the decisions and actions of the individual or due to other outside influences.

Endsley, Bolte, and Jones [18] present an approach to human-centered design that focuses on designing for SA. Like EID, this approach includes design methods and design principles. Similar to EID, designing for SA starts with an analysis of the work environment. However, the focus in this analysis is on determining SA requirements. SA requirements are the operator's dynamic information needs with respect to attaining their goals. Endsley et al. promote a method known as goal directed task analysis [18]. A *goal directed task analysis* focuses on identifying the basic goals of the operators, the decisions required to accomplish those goals, and the information necessary (or SA requirements) for each decision. The focus of a goal directed task analysis is independent of current system technologies or capabilities. A goal directed task analysis is generally completed using a series of interviews in which multiple operators are queried to obtain a hierarchy of goals, sub-goals, decisions, and information requirements (for example, see Figure 2).

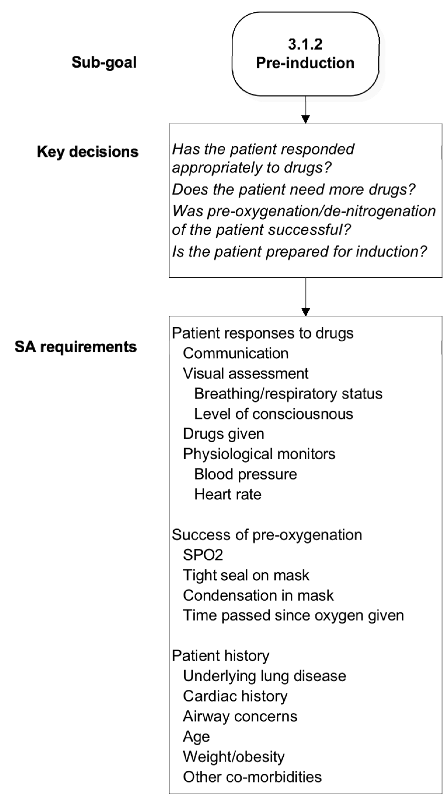

Figure 2. Sample portion of a goal directed task analysis for a nurse anesthetist

Following the goal directed task analysis, designing for SA then requires attention to SA-oriented design principles in the design, followed by evaluation using SA measures. Like other human-centered design approaches, the process is iterative, with the results of the evaluation informing re-design. Endsley et al. [18] provide 50 design principles intended to support the development and maintenance of SA, some paralleling the principles of EID. For example, SA design principles include: (1) directly support comprehension and (2) organize information around goals. In EID, direct mapping to work domain constraints is intended to support comprehension and goals. Other design principles that derive directly from a model of SA include: (1) design systems to support SA rather than decisions, (2) provide support for projection (e.g., trend displays), and (3) make both the known and the unknown apparent.

Methods of evaluating the design are focused toward measuring SA. Indirect measures of SA are similar to other measures used for evaluation in human-centered design and can include subjective measures, observation, verbal protocol analysis (asking participants to speak out loud their actions), and performance measures. Direct measures of SA include subjective rating measures and objective methods that involve querying operators regarding their current perception, comprehension, and projection of the state of the system. Queries may be completed after a session is complete, during "freezes" in a simulated session, or real-time in either field or laboratory environments [18, 19].

III. Qualitative Research and Grounded Theory in Human Centered Design

We have found that work domain analysis and goal-directed task analysis may be difficult to apply for the purposes of intelligent system design. Development of an abstraction hierarchy, originally intended for a complex physical system, is difficult to apply in industries such as health care where constraints are less easily defined [11, 14]. In applying goal directed task analysis, it can be difficult to elicit articulation of goals or to achieve consensus on goals (as opposed to articulating tasks or describing events in sequence). Other task analytic methods such as verbal protocol analysis [20], structured interviews or focus groups, and critical decision methods [21] may be easier to implement but may also provide less of a direct link between results and design. As a means of analyzing these data toward design purposes, we suggest an approach in which *grounded theory* or other content analysis techniques are applied to data collected through a wide variety of methods such as focus groups and observation. Using these techniques allows researchers and designers to select the data collection method (or methods) that is most appropriate to the specific application and work environment.

Grounded theory emerged from the work of Barney Glaser and Anselm Strauss in studying the social organization and temporal order of dying [22]. They shared their methods with a goal toward developing systematic methodological strategies, comparable in rigor to quantitative methods, that social scientists could adopt for studying other topics [23]. We contend that identifying applications for intelligent system design to support human-system performance is well suited to grounded theory methods. These methods provide an approach for analyzing qualitative data from a variety of sources such as interviews and observations and identifying emerging themes, concepts, or theories that will direct efforts in intelligent system design.

Grounded theory is founded on several key research principles [23, 24]. First, *constant comparison* refers to the ongoing analysis of data as it is collected. Data is compared to data as the data is collected. The benefit of this method for the purposes of *generating theory* is that early data can drive decisions about where and how to look for further data. This is particularly useful in the context of studying human-machine systems as it is often unclear where intelligent systems or other new tools will be most beneficial. Second, *saturation* refers to the situation that is achieved when continued data collection reveals no new information regarding the environment or application under study. While the concept of saturation can make it difficult to predict how much time and effort may be required to study a particular problem, in the absence of statistical methods, it is important for providing evidence that a problem has been studied to sufficient depth. The specifics of coding and analysis based on grounded theory techniques are described in section B below.

A. Collecting Rich Data

Whether following grounded theory in detail or using other qualitative analysis techniques, there are a variety of methods for collecting qualitative data to identify appropriate applications and implementations for intelligent system design. Examples include: (1) structured interviews, (2) critical decision interviews (experts recall a specific incident that was non-routine, challenging, or difficult), (3) observation of task performance in the field, in simulation, or on video, (4) retrospection, and (5) verbal protocol analysis of ongoing, recorded, or simulated task performance [21, 25]. Focus group interviews provide the advantage of accessing the input of multiple participants at the same time. In addition, focus group interviews can create a synergy among participants that leads to a rich discussion of issues that may not be apparent in single person interviews [26]. Data collection methods are generally observational and the type of information that is collected varies depending on researcher goals. Because there are advantages and disadvantages to the different techniques, multiple techniques may be used to meet specific goals.

While interviews of various types can provide rich qualitative data, they may fail to yield an understanding regarding passive use behaviors or critical contextual details that may be needed for developing computer algorithms for intelligent system applications. In contrast, observations and computer-recorded use data may provide a contextual understanding of the information users access, but provide no insight into why information is accessed. For example, is data of interest to users because it is important, or perhaps, because it is inaccurate? One means of combining the rich contextual data provided through observation of use and the rich qualitative data provided through interviews is to record use behavior then replay that recording for the provider, asking them to describe and rationalize their behaviors [27]. Ultimately, the quality of the data collected should be assessed for comprehensiveness regarding relevant individuals and contexts, range of views and actions, and temporal changes [23].

B. *Qualitative Data Coding*

Grounded theory methods use coding techniques known as open coding and axial coding. Generally, analysis of qualitative data using grounded theory involves the coding of text-based transcriptions of interview or observation data. In open coding, incidents or issues with similarities are grouped together into themes or categories, which are named according to meaning. Through constant comparison, themes are renamed, reorganized, and redefined through an ongoing process of refinement. Using the themes identified through open coding, axial coding is a process in which analysts explore and define relationships between categories.

Other methods of coding qualitative data may be described as content analysis [28]. In conventional content analysis, codes are defined during data analysis and are derived from data as described in grounded theory methods. Directed content analysis is a method in which codes are defined before and during data analysis. These codes may be defined from theory or from relevant research findings. Directed content analysis may be appropriate for studies in which specific pre-conceived ideas are of interest. In some cases intelligent system designers and researchers may choose an approach that combines quantitative and qualitative data collection and analysis [29].

As is true with quantitative research, qualitative research methods must be evaluated for the degree to which researchers are exposing "truth" [26, 30]. Examples of steps can be taken to promote reliability, exhaustiveness and trustworthiness include:

- Collecting data from a variety of different individuals and in different settings
- Collecting data to saturation
- Using at least two independent coders
- Coders create, adhere to, and record rules regarding concepts and categories.
- Feeding back data interpretations to a subset of participants for agreement or disagreement, seeking dissenting opinions.
- Detailed documentation of all processes for collecting and coding data.

Example applications of grounded theory methods for studying health care work environments that may ultimately have applications for expert system design include observations of emergency room status board use [31] and clinical oversight of medical trainees [32]. While we have only begun to apply the combination of verbal protocol use-based data collection with grounded theory qualitative analysis in our lab, our experiences with a variety of interview, observation, task analytic, and qualitative research methods suggest great promise for this approach.

IV. CONCLUSIONS

Designing comprehensible intelligent systems requires a focus on the individuals who are expected to use and understand those systems throughout the entire design process. While many designers and developers focus on the evaluation and refinement of designs after initial prototypes have been designed, we propose that it is critically important to study intended users and use contexts early in the conceptual design of systems. Formal human-centered design methods have been successfully applied for these purposes. In addition, formal qualitative research analysis techniques can be applied to data collected in a variety of ways to study the needs of workers in target work environments.

ACKNOWLEDGMENT

Dr. Wright's effort was funded by the Agency for Healthcare Research and Quality (K02 HS015704-01). Funding for Dr. Mark and Dr. Segall were provided in part by the Veterans Affairs National Center for Patient Safety, Patient Safety Center of Inquiry (2007-2009).

REFERENCES

[1] G. Klein, Sources of Power: How People Make Decisions, Cambridge, MA: The MIT Press, 1998.

[2] M. R. Endsley, and E. O. Kiris, "The out-of-the-loop performance problem and level of control in automation," Human Factors, vol. 37, 1995, pp. 381-394.

[3] R. Parasuraman, and V. A. Riley, "Humans and automation: Use, misuse, disuse, and abuse," Human Factors, vol. 39, 1997, pp. 230-253.

[4] R. I. Cook, and D. D. Woods, "Adapting to New Technology in the Operating Room," Human Factors, vol. 38, 1996, pp. 593-613.

[5] A. Jungk, B. Thull, A. Hoeft, and G. Rau, "Evaluation of two new ecological interface approaches for the anesthesia workplace," Journal of Clinical Monitoring, vol. 16, 2000, pp. 243-258.

[6] D. H. Harris, "Human Factors Success Stories," in 28th Annual Meeting of the Human Factors Society, Santa Monica, CA: Human Factors and Ergonomics Society, available at www.hfes.org, 1984.

[7] IBM. "User-Centered Design," January 6, 2004; http://www-306.ibm.com/ibm/easy/eou_ext.nsf/Publish/570.

[8] A. Hart, Knowledge Acquisition for Expert Systems, New York, NY: McGraw-Hill, 1992.

[9] W. Browne, L. Yao, I. Postlethwaite, S. Lowes, and M. Mar, "Knowledge-elicitation and data-mining: Fusing human and industrial plant information," Engineering Applications of Artificial Intelligence, 3, 2006, pp. 345-359.

[10] B. Jeng, T. Liang, and M. Hong, "Interactive induction of expert knowledge," Expert Systems with Applications, 3, 1996, pp. 393-401.

[11] M. O. Watson, and P. Sanderson, "Designing for attention with sound: challenges and extensions to ecological interface design," Hum Factors, vol. 49, Apr, 2007, pp. 331-46.

[12] J. R. Hajdukiewicz, K. J. Vicente, D. J. Doyle, P. Milgram, and C. M. Burns, "Modeling a medical environment: an ontology for integrated medical informatics design," Int J Med Inform, vol. 62, Jun, 2001, pp. 79-99.

[13] K. J. Vicente, and J. Rasmussen, "Ecological interface design: Theoretical foundations," IEEE Transactions on Systems, Man, and Cybernetics, vol. 22, 1992, pp. 589-606.

[14] K. J. Vicente, "Ecological interface design: progress and challenges," Hum Factors, vol. 44, Spring, 2002, pp. 62-78.

[15] J. Rasmussen, “The role of hierarchical knowledge representation in decision making and system management,” IEEE Transactions on Systems, Man, and Cybernetics, vol. 15, 1985, pp. 234 - 243.

[16] J. Rasmussen, “Skills, rules, knowledge: Signals, signs, and symbols and other distinctions in human performance models,” IEEE Transactions on Systems, Man, and Cybernetics, vol. 13, 1983, pp. 257-267.

[17] M. R. Endsley, “Toward a Theory of Situation Awareness in Dynamic Systems,” Human Factors, vol. 37, 1995, pp. 32-64.

[18] M. R. Endsley, B. Bolte, and D. G. Jones, Designing for Situation Awareness: An Approach to Human-Centered Design, London: Taylor and Francis, 2003.

[19] M. C. Wright, J. M. Taekman, and M. R. Endsley, “Objective measures of situation awareness in a simulated medical environment,” Qual Saf Health Care, vol. 13 Suppl 1, Oct, 2004, pp. i65-71.

[20] L. Bainbridge, Verbal protocol analysis, 2nd ed., Bristol, Pa.: Taylor & Francis, 1995.

[21] R. R. Hoffman, B. Crandall, and N. Shadbolt, “Use of the critical decision method to elicit expert knowledge: A case study in the methodology of cognitive task analysis,” Human Factors and Ergonomics, vol. 40, 1998, pp. 254-276.

[22] B. Glaser, and A. Strauess, The Discovery of Grounded Theory: Strategies for Qualitative Research, Chicago: Aldine Publications Co., 1967.

[23] K. Charmaz, Constructing Grounded Theory, Thousand Oaks, CA: Sage Publications Inc., 2006.

[24] T. Kennedy, and L. Lingard, "Making sense of grounded theory in medical education," Med Educ, 2, 2006, pp. 101-108.

[25] R. Hoffman, A. M. Burton, N. Shadbolt, and G. Klein, “Eliciting knowledge from experts: A methodological analysis,” Organizational Behavior and Human Decision Process, vol. 62, 1995, pp. 129-158.

[26] K. E. Steinhauser, "Qualitative Research Methods: An Overview," Clinical Research Training Program: Health Services Research, Duke University Medical Center, 2009.

[27] A. W. Kushniruk, V. L. Patel, and J. J. Cimino, "Usability testing in medical informatics: cognitive approaches to evaluation of information systems and user interfaces," Proceedings : a conference of the American Medical Informatics Association / AMIA Annual Fall Symposium AMIA Fall Symposium, 1997, pp. 218-22.

[28] H. Hsieh, "Three Approaches to Qualitative Content Analysis," Qualitative Health Research, 9, 2005, pp. 1277-1288.

[29] L. D. Wolf, P. Potter, J. A. Sledge et al., “Describing nurses' work: combining quantitative and qualitative analysis,” Human factors, vol. 48, May 16, 2006, pp. 5-14.

[30] N. Mays, and C. Pope, "Qualitative research: Observational methods in health care settings," BMJ, 6998, 1995, pp. 182-4.

[31] R. L. Wears, S. J. Perry, S. Wilson, J. Galliers, and J. Fone, "Emergency department status boards: user-evolved artefacts for inter- and intra-group coordination," Cogn Tech Work, 3, 2007, pp. 163-170.

[32] T. J. Kennedy, L. Lingard, G. R. Baker, L. Kitchen, and G. Regehr, “Clinical oversight: conceptualizing the relationship between supervision and safety,” J Gen Intern Med, vol. 22, Aug, 2007, pp. 1080-5.