Interlibrary Loans and Journal Article Requests

Notice Warning Concerning Copyright Restrictions:

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted materials.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One specified condition is that the photocopy or reproduction is not to be *"used for any purpose other than private study, scholarship, or research."* If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use," that user may be liable for copyright infringement.

Upon receipt of this reproduction of the publication you have requested, you understand that the publication may be protected by copyright law. You also understand that you are expected to comply with copyright law and to limit your use to one for private study, scholarship, or research and not to systematically reproduce or in any way make available multiple copies of the publication.

The Stephen B. Thacker CDC Library reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

Terms and Conditions for items sent by e-mail:

The contents of the attached document may be protected by copyright law. The CDC copyright policy outlines the responsibilities and guidance related to the reproduction of copyrighted materials at CDC. If the document is protected by copyright law, the following restrictions apply:

- You may print only one paper copy, from which you may not make further copies, except as maybe allowed by law.
- You may not make further electronic copies or convert the file into any other format.
- You may not cut and paste or otherwise alter the text.

Using Storytelling to Elicit Design Guidance for Medical Devices

BY KIM GAUSEPOHL, WOODROW W. WINCHESTER III, JAMES D. ARTHUR, & TONYA SMITH-JACKSON

Storytelling allows more freedom to uncover details and fills gaps where privacy concerns prohibit information gained through direct observation.

The Institute of Industrial Engineers (IIE) Council of Fellows identified the re-engineering of health care delivery as a grand challenge for engineers (IIE, 2007). Part of the challenge involves the design of usable medical devices. Usability directly affects practitioners' ability to perform diagnostic tasks efficiently, effectively, and safely. Poor usability may facilitate error and undermine patient safety. Practitioners may become "second victims" of patient injury from the resulting emotional and professional consequences (Wu, 2000).

In this article, we address the potential benefits of storytelling as an elicitation method in a domain in which opportunities for observation are limited. We provide practical recommendations for conducting storytelling sessions based on our experiences.

Designing Usability to Address Medical Errors

Medication errors are one of the most common types of health care errors. Devices such as the infusion pump improve safety through controls that regulate medication delivery. However, the design may not prevent administration errors, such as delivering an incorrect volume or rate. In fact, the design may facilitate operator error (Lane, Stanton, & Harrison, 2006), and a single "slip-of-the-finger" error can deliver unsafe levels of medication (Husch et al., 2005).

Prevention through design is a human factors strategy to "design out" characteristics that contribute to error (National Occupational Research Agenda, 2009). Design standards recommend a focus on usability to identify error opportunities. The International Electrotechnical Commission (2007) standard for the application of usability engineering to medical devices suggests that designers create a usability specification to identify potential hazards and errors associated with usability. Device usability is validated against testable requirements listed in the specification to ensure that the design reduces opportunities for error.

User-Centered Design and Storytelling

Standards also recommend a user-centered design (UCD) process. UCD begins with understanding the context of use, which is defined as "user characteristics, tasks, equipment, and a physical and social environment in which a product is used" (International Organization for Standardization, 1999). Standards warn designers that usability is affected by the context of use. For example, an infusion pump designed for use in a hospital may be viewed as too bulky in an ambulance.

Ethnographic methods such as observations and interviews help designers understand a product's context of use. However, designers must overcome domain-specific obstacles to using these recommended methods within health care facilities (Martin, Norris, Murphy, & Crowe, 2007). Although standards require user research during requirements gathering, patient privacy regulations often prevent observations. Thus, designers may resort to using self-report methods, though important contextual information may be lost, as the practitioner is removed from the environment and must recall and express information in a way that is understandable to the designer. The inability to observe practitioners,

FEATURE AT A GLANCE: Medical device designers must understand the complex context of use within a health care environment to ensure product usability. Designers must overcome domain-specific obstacles during usability research, such as patient privacy standards, which prevent designers from observing practitioners in context. In this project, we investigated storytelling as an alternative elicitation method for medical device requirements when direct observations are limited or not possible. While gathering requirements for an infusion pump, we compared the types of information elicited by focus groups, interviews, and storytelling sessions. Several advantages and implications for the use of storytelling in usability research are discussed.

KEYWORDS: usability, narrative, health care, design, infusion pump, thematic analysis

coupled with the limitations of self-reports, impedes understanding of the context of use.

Storytelling has been used successfully as an ethnographic research method for data gathering and analysis within the social sciences (Mishler, 1986; Riessman, 1993). Specific to health care, storytelling has also facilitated understanding of practitioners and workplace demands (Wolf & Zuzelo, 2006). Storytelling can serve as a robust self-report method by supporting the retrieval of contextual cues from long-term memory. Given that designers have adapted other ethnographic methods, the use of storytelling warrants further investigation.

Although standards require user research during requirements gathering, patient privacy regulations often prevent observations.

Similar to Flanagan's (1954) critical incident technique (CIT), storytelling provides a framework for collecting and analyzing retrospective reports. The purpose of CIT is to collect "observed incidents having special significance and meeting systematically defined criteria" (Flanagan, 1954, p. 327). Similarly, the purpose of storytelling is to collect narrative, which is a personal account of experience. Narrative can be viewed as a structured ordering of events (Labov & Waletzky, 1967), a representation of character and action (Riessman, 1993), or a production of storyteller and listener (Mishler, 1986). We operationally define *narrative* as a representation of personal experience formed by content and structure.

In contrast to interviews, which encourage a question-and-answer discourse, storytellers are allowed more freedom in directing the conversation. Although it is possible to elicit stories during interviews, the ability to do so depends on the skills of the interviewer (Riessman, 1993). One of storytelling's potential benefits during requirements gathering is the ability to elicit practitioners' experiences embedded within contextual information that may be lacking in interview responses.

Objectives of Our Research

The objective of this exploratory research was to investigate the use of storytelling as a requirements-elicitation method for medical devices. We expected that storytelling would enable designers to capitalize on the inherent narrative nature of the health care domain, given that stories are used extensively to communicate medical knowledge (Hunter, 1991). We compared the information retrieved from storytelling sessions with information elicited from open-ended interviews. Empirical studies that explore the types of information acquired from different elicitation methods are needed to provide the requisite insight to choose an appropriate method.

Method

We conducted focus groups, interviews, and storytelling sessions to determine the effectiveness of storytelling as an elicitation method. We included focus groups and interviews because design standards identify these as common elicitation methods.

Ten registered nurses participated in requirements-gathering sessions for an infusion pump. Infusion pumps pose a design challenge because of their ubiquitous use in hospitals, the diverse user base, the multiple uses, and the potentially conflicting needs of patients and practitioners. According to the Food and Drug Administration's (2010) Infusion Pump Improvement Initiative, 87 infusion pumps were recalled between 2005 and 2009, with many problems identified as design and engineering flaws.

We followed recommendations for triangulation and used more than one elicitation method in each group to gather a broad range of user requirements (Garmer, Liljegren, Osvalder, & Dahlman, 2002). We initially conducted separate focus groups to bring project stakeholders together to talk about user needs (Wiklund, 1995). We created balanced groups on the basis of practitioner demographics to ensure similarity in gender, age, and experience. We assigned elicitation methods to each group to allow for comparisons: (a) FG&I, focus group followed by individual interviews; and (b) FG&S, focus group followed by individual storytelling sessions.

Scripts were created to ensure session similarity. We developed the focus group questions to elicit high-level goals and requirements. Interview and storytelling questions were developed to elicit usability requirements (see Table 1). To ensure that we prompted practitioners to discuss all aspects of usability, we also designed questions to gain information about the usability components of efficiency, effectiveness, and satisfaction (see Table 2). Negative correlates of these components that result from a lack of usability (inefficiency, error, and worker stress) were elicited as well. The questions were controlled to allow for similarity between interviews and storytelling sessions.

TABLE 1. SUMMARY OF ELICITATION TECHNIQUES USED

Method	Purpose	Time in minutes
Focus group Group 1, $n = 5$; Group 2, $n = 5$	High-level goals and requirements	48.2 ($SD = 1.1$)
Interview Group 1, $n = 5$	Usability requirements	39.4 ($SD = 5.8$)
Storytelling session Group 2, $n = 5$	Usability requirements	30 ($SD = 4.5$)

TABLE 2. EXAMPLE PROMPTS TO ELICIT USABILITY INFORMATION

Usability Category	Prompt	Interview/Storytelling Questions
Efficiency	Please tell me a story about efficiency and infusion pumps. Be sure to include *who, what, why, where, when, how* details in your story.	*Who* is affected by the efficiency of the infusion pump? *What* functions, features, or qualities affect the efficiency of the infusion pump? *Why* is using an infusion pump more efficient than other types of medication delivery? In which *locations or contexts* are infusion pumps more efficient? *When* do you determine if your use of the infusion pump has been efficient or not? *How* should an infusion pump work so that your work is as efficient as possible?

Each protocol was designed to adhere to Shefelbine, Clarkson, Farmer, and Eason's (2002) recommendation to ask "who, what, why, where, when" questions during requirements gathering. Our objective was to compare the types of content elicited via each elicitation method, so we created questions of who, what, why, where, when, and how for each of the usability categories (e.g., efficiency-inefficiency, effectiveness-error, satisfaction-stress) to ensure all story components were addressed during interviews. For example, one of our interview questions relating to medical device error was, "Who typically makes mistakes while using an infusion pump?" In storytelling sessions, we simply asked practitioners to tell us a story about the desired usability category, such as "Tell me a story about errors and infusion pumps," and reminded practitioners to include the who, what, why, where, when, and how components.

Results

We used a mixed-methods approach to analyze the 53 transcript pages to determine which elicitation combination resulted in the better information set. We operationally defined a "better" elicitation method in terms of quantity, quality, and time. We used *requirement themes* as the unit of analysis, which was operationally defined as a practitioner's statement(s) of a user need that contributes to the development of a requirement. This unit of analysis made it possible for us to analyze transcripts qualitatively prior to conducting a quantitative analysis.

Two data coders used thematic analysis to identify requirement themes. A data-coding judge reconciled intercoder disagreement. The content analysis allowed comparisons between the content contained within nonnarrative interviews and story narratives. In contrast to the typical inductive CIT approach of grouping incidents into newly formed categories, data coders categorized each identified theme using a requirements ontology created for this study. The ontology (see Figure 1) is based on definitions for medical device usability provided in design standards IEC 62366 (IEC, 2007) and ANSI/AAMI HE74 (American National Standards Institute & Association for the Advancement of Medical Instrumentation, 2001). The requirements ontology includes the usability components of efficiency, effectiveness, satisfaction, and context of use. We expanded the context-of-use category to include the domain-specific contextual factors of spatial, social, technological, hygienic, physical, and activity context (see Table 3 on page 22).

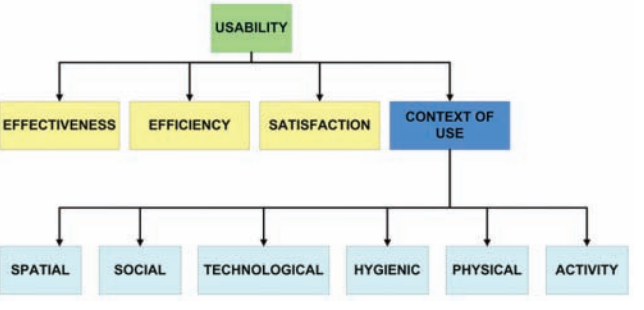

Figure 1. Requirements ontology used during thematic analysis.

Additional quantitative analyses was performed to determine which group provided the better information set.

Quantity. A better elicitation method elicits a greater quantity of information per individual. Quantity is represented as how many requirement themes an individual contributed for a given category. We calculated the total number of themes identified per practitioner, per category, as the union between the results of the focus group and individual follow-up session. Two-sample *t* tests were used to test for significant differences in the quantities of themes addressed per practitioner for each category.

No significant differences were found for any of the requirements categories. This result suggests that a similar quantity of information can be expected from either combination of methods.

Quality. A better elicitation method elicits greater quality of compiled information. We were interested in the compiled results for each group, given that designers value the

TABLE 3. CONTEXTUAL FACTORS FOR MEDICAL DEVICES (INTERNATIONAL ELECTROTECHNICAL COMMISSION, 2007)

Context	Example
Activity	Aspects of the activity, such as distractions and task complexity
Hygienic	Requirements for ensuring sterile conditions
Physical	Environmental factors, such as visibility and noise level
Social	Organizational factors, such as practitioner hierarchy and procedures for transitioning care between wards
Spatial	The activity's location (e.g., hospital ward)
Technological	The use of other technological equipment in conjunction with the device of interest

information set, not individual responses. *Quality* was defined as the breadth and depth of the compiled information per group. Breadth, the group's coverage of all possible requirements categories, indicates the comprehensiveness of the information. Depth (i.e., how many themes a group identified within each category) indicates information completeness.

Breadth. We explored breadth to investigate whether practitioners discussed a broader range of themes as a result of the differences in treatments. The compiled sets for both groups covered all possible categories, so we performed further testing on the breadth of each individual's response. A two-sample *t* test was used to test for significant differences.

No significant differences for the breadth of information per practitioner were found. This finding suggests that similar breadth of information is expected from either elicitation combination (see Table 4).

Depth. We explored depth to investigate differences in the completeness of the compiled results. The total number of themes identified per group was calculated as the union between the results of the focus group and all individual follow-up sessions. We used Venn diagrams to discern differences visually, as statistical testing was not appropriate. We considered a 10% increase to be an indicator of significant differences.

We found depth differences for the context-of-use and social context categories. FG&S participants discussed 27% more distinct context-of-use themes than did FG&I participants (see Figure 2). Further exploration of context-of-use categories revealed that FG&S participants discussed 60% more distinct social themes (see Figure 3). FG&S participants gave a more holistic view of all those involved during patient care, including the interactions among employees,

TABLE 4. DEPTH: SUMMARY OF DISTINCT THEMES IDENTIFIED PER CATEGORY PER GROUP

Category	FG&I Themes	FG&S Themes	Overlap
Usability	422	440	118
Efficiency	90	92	40
Effectiveness	94	92	45
Satisfaction	194	176	57
Context of use	189	239	72
Spatial	21	24	11
Social	58	93	18
Technological	10	12	2
Hygienic	4	5	2
Physical	2	1	0
Activity	110	119	48

Note. FG&I = focus group followed by individual interviews; FG&S = focus group followed by individual storytelling sessions.

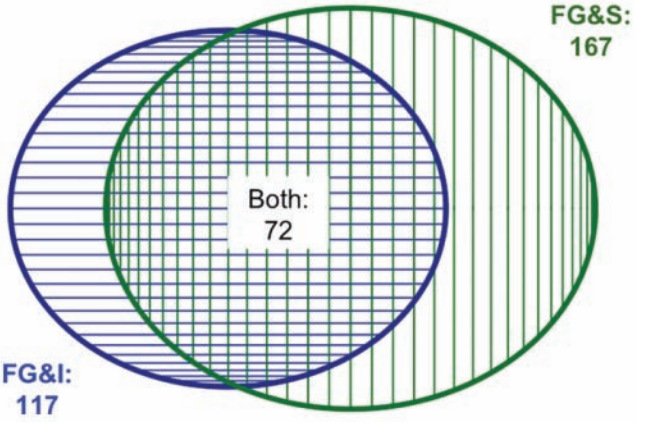

Figure 2. Depth: distinct context-of-use themes per group.

the responsibilities of employees, and the hospital's organizational culture.

An unexpected finding was that FG&I participants discussed 10% more distinct satisfaction themes. A possible explanation for this discrepancy may be a tendency for practitioners to focus on desires during interviews.

Time. Given practitioners' work demands and time constraints, we considered a better elicitation method to be one that required less time. We found significant differences between groups by using the Kruskal-Wallis test. On average, interview participants spent 30% more time – approximately 12 minutes – in follow-up sessions.

This finding has several implications. First, because we found no significant difference between the quantity and

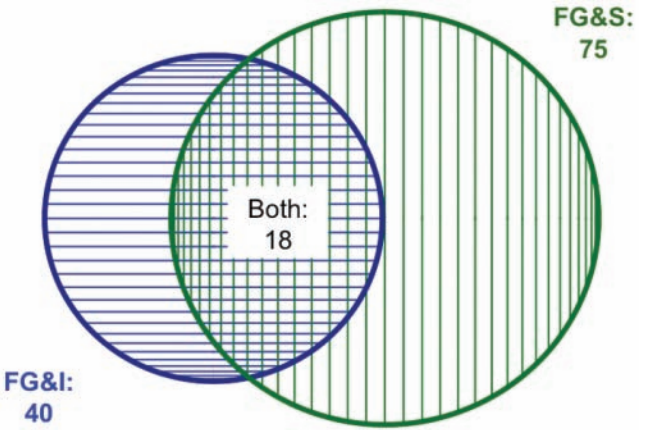

Figure 3. Depth: distinct social context themes per group.

breadth of themes identified between the two groups, we infer that FG&S participants contributed similar results in less time. An elicitation method that takes less time is valuable in the medical domain, where access to practitioners is limited.

The time difference also suggests that storytelling's ability to elicit usability information was not fully revealed. One possible explanation for the time difference was the lack of follow-up questions during storytelling sessions. We designed the interview script to allow for spontaneous prompts, but we neglected to include this flexibility in the storytelling protocol.

Discussion

Designers may want to examine the results of our empirical comparison when considering the appropriateness of storytelling in their work. For example, they may use storytelling when time with stakeholders is limited or when building initial rapport. Designers may use storytelling when contextual factors heavily influence design decisions, though prompts relating to certain contextual factors may be required. For example, practitioners in both groups neglected to discuss the hygienic and physical contexts.

Storytelling may be best suited during the concept stage, when designers are still exploring the problem space. For example, designers may initially gather stories and then clarify information using other methods, such as observations and interviews.

Practical Recommendations for Designers

On the basis of our exploration, we provide the following practical recommendations for how to make a storytelling session go smoothly.

1. ***Advise participants of topics several days in advance.*** We noted a reactive effect whereby more experienced nurses expressed difficulty recalling specific events after story prompts. One practitioner summarized her difficulty at the end of the session: "It's hard to be really specific. Especially when you've been a nurse for a while. . . . I'll go home and think of a million examples." Similar to Flanagan's (1954) recommendation to advise participants of the incidents of interest prior to the CIT interview, we encourage designers to advise participants of the story themes several days in advance so participants have time to reflect and collect their thoughts.

2. ***Ask clarifying questions while keeping track of story progress.*** Practitioners frequently used medical jargon during storytelling sessions, but we did not ask clarifying questions because of fear of interrupting the story. We suspect that clarifying questions will act as prompts, sparking practitioners to tell additional stories. Although the collection of additional stories is beneficial, it is important for designers to keep track of story progress, as participants may abandon one story to tell a "story within a story." We suggest that designers make a written note of the practitioner's last statement when participants start to diverge from the original story. The last statement may be used as a prompt to encourage the practitioner to finish the original story.

Storytelling may be best suited during the concept stage, when designers are still exploring the problem space.

3. ***Emphasize anonymity throughout the session to ensure that participants share personal stories.*** When we emphasized anonymity only at the beginning of the process, we found that FG&S participants relayed many personal stories, but none provided stories of personal involvement with medical error. We altered our protocol to allow respondents to tell a story about a witnessed medical error when they displayed discomfort with the question. In retrospect, we should have emphasized anonymity repeatedly throughout the session to encourage participants to share first-person stories of medical error.

4. ***Encourage personal stories, but allow witnessed accounts.*** Some participants will be unwilling or unable to provide personal stories for some topics. Designers should initially prompt participants to tell first-person accounts but should allow participants to tell stories of witnessed events when the participant cannot contribute a first-person account.

The inclusion of witnessed accounts provides two benefits. First, it allows participants a means to elegantly opt out of telling an uncomfortable personal story. Second, it allows designers to collect information that may not be provided in first-person accounts. For example, in our collection of medical error stories, we found that these witnessed events provided valuable information regarding the organizational culture of the hospital, which increased our understanding of the social context.

Conclusion

The results of the empirical comparison provide support for the use of storytelling as an elicitation method for medical device requirements. Although we did not find significant differences in the quantity or breadth of requirements

categories, we found substantial differences in the depth of context-of-use and social requirements categories. This finding supports our initial hypothesis that storytelling would aid designers in the collection of elusive contextual information.

We encourage designers to consider the use of creative methods such as storytelling when learning about an unfamiliar domain. Given that requirements elicitation is more of an art form than an exact science, designers must be able to choose an appropriate method on the basis of the needs of the project.

References

American National Standards Institute & Association for the Advancement of Medical Instrumentation. (2001). *Human factors design process for medical devices* (ANSI/AAMI HE74). Washington, DC/Arlington, VA: Author.

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, *51*(4), 28–35.

Food and Drug Administration. (2010). *Infusion Pump Improvement Initiative*. Retrieved from http://www.fda.gov/medicaldevices/productsandmedicalprocedures/GeneralHospitalDevicesandSupplies/InfusionPumps/ucm205424.htm

Garmer, K., Liljegren, E., Osvalder, A.-L., & Dahlman, S. (2002). Arguing for the need of triangulation and iteration when designing medical equipment. *Journal of Clinical Monitoring and Computing*, *17*(22), 105–114.

Hunter, K. M. H. (1991). Knowledge in medicine: Reading the signs. In *Doctors' stories: The narrative structure of medical knowledge* (pp. 3–26). Princeton, NJ: Princeton University Press.

Husch, M., Sullivan, C., Rooney, D., Barnard, C., Fotis, M., Clarke, J., & Noskin, G. (2005). Insights from the sharp end of intravenous medication errors: Implications for infusion pump technology. *Quality and Safety in Health Care*, *14*(2), 80–86.

International Electrotechnical Commission. (2007). *Medical devices—Application of usability engineering to medical devices* (IEC 62366). Geneva, Switzerland: Author.

International Organization for Standardization. (1999). *Human-centred design processes for interactive systems* (ISO 13407). Geneva, Switzerland: Author.

Institute of Industrial Engineers. (2007). *Grand challenges for engineering*. Norcross, GA: Author.

Labov, W., & Waletzky, J. (1967). Narrative analysis: Oral versions of personal experience. In J. Helm (Ed.), *Essays on the verbal and visual arts* (pp. 12–44). Seattle: University of Washington Press.

Lane, R., Stanton, N. A., & Harrison, D. D. (2006). Applying hierarchical task analysis to medication administration errors. *Applied Ergonomics*, *37*, 669–679.

Martin, J. L., Norris, B. J., Murphy, E., & Crowe, J. A. (2007). Medical device development: The challenge for ergonomics. *Applied Ergonomics*, *39*, 271–283.

Mishler, E. G. (1986). *Research interviewing: Context and narrative*. Cambridge, MA: Harvard University Press.

National Occupational Research Agenda. (2009). *State of the sector: Healthcare and social assistance*. Washington, DC: Department of Health and Human Services.

Riessman, C. K. (1993). *Narrative analysis* (Vol. 30). Newbury Park, CA: Sage.

Shefelbine, S., Clarkson, P. J., Farmer, R., & Eason, S. (2002). *Good design practice for medical devices and equipment: Requirements capture*. Cambridge, UK: University of Cambridge Engineering Design Centre/University of Cambridge Institute for Manufacturing.

Wiklund, M. E. (1995). Conducting effective focus groups. In M. E. Wiklund (Ed.), *Medical device and equipment design: Usability engineering and ergonomics* (pp. 69–80). Concord, MA: Taylor & Francis.

Wolf, Z. R., & Zuzelo, P. R. (2006). "Never again" stories of nurses: Dilemmas in nursing practice. *Qualitative Health Research*, *16*, 1191–1206.

Wu, A. W. (2000). Medical error: The second victim. *British Medical Journal*, *320*, 726–727.

Kim Gausepohl is a graduate student in the Grado Department of Industrial and Systems Engineering at Virginia Polytechnic Institute and State University. Her research interests include requirements elicitation, user-centered design, and the impact of medical device usability on practitioner and patient safety. She may be reached at kgausepo@vt.edu.

Woodrow W. Winchester III is a professor in the Grado Department of Industrial and Systems Engineering at Virginia Polytechnic Institute and State University. He received his PhD in industrial engineering from North Carolina A&T State University in 2005.

James D. Arthur is a professor in the computer science program at Virginia Polytechnic Institute and State University. He received his PhD in computer science from Purdue University in 1983. His teaching and research interests focus on concepts and issues that are inherent in software engineering.

Tonya Smith-Jackson is a professor in the Grado Department of Industrial and Systems Engineering at Virginia Polytechnic Institute and State University. She received her PhD in psychology and ergonomics from North Carolina State University in 1998.

This article was supported by Grant No. T01OH008613 from the Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health (NIOSH). Its contents are solely the responsibility of the authors and do not necessarily reflect the official views of NIOSH.