Research article

Testing a model to help analysts understand, evaluate, and make inferences about health technology change

Cara Stitzlein[1], Penelope Sanderson[1,2], Cristina Beltran Orihuela[1,3], Leanne Jack[1,3] and Bala Venkatesh[3]

[1] *The University of Queensland, Brisbane, Australia*
[2] *National ICT Australia, Brisbane, Australia*
[3] *Princess Alexandra Hospital, Brisbane, Australia*

Abstract

Background: Implementing new technologies in a hospital is difficult. The healthcare environment is complex, with many factors for decision makers to consider before a new system is introduced. **Aim:** The aim of this study was to evaluate a modeling approach designed to help healthcare stakeholders evaluate the impact of electronic information technologies on existing clinical work. We performed an initial test of an information flow and workflow model, evaluating its effectiveness at providing descriptions, supporting evaluation, and supporting conjectures. **Method:** Hospital staff were interviewed and a model of information flow and workflow of a critical care work context was developed from the interview data. We then used a structured set of tests to probe non-healthcare participants' ability to understand the model's notation and evaluate the work activities. **Results:** An analysis of participants' responses indicated that although participants did not understand some model symbols initially, once provided with a legend they could satisfactorily describe the information flow and workflow represented in the model. The participants' verbal reports of work activities and their perceptions of the notation indicate which parts of the model were difficult to understand and suggest ways of improving this type of representation. **Conclusions:** Initial support was found for the adequacy of our modeling approach. Results from this study are being used to refine the model's notation and structure. Next steps in the research are discussed, which include a test with healthcare professionals to see if our approach helps stakeholders make decisions about the implementation of new technologies.

Background

The use of new information technologies in healthcare can 'reinvent' how the health system functions and how healthcare is delivered [1]. Policy makers and clinicians expect technologies such as computerized information systems to improve the management and sharing of clinical information, and to optimize the continuity of care [2]. The failure to consider how clinical work is performed prevents the effective design and implementation of these technologies, as stated by Patrick in his 2011 report about the deployment of a health information system in New South Wales [3]. Patrick reported that the system could not be effectively used by clinicians, and that information was not shared efficiently. These consequences caused disruption to patient discharges from the hospital department. A better awareness of existing clinical practices and work routines may have prevented some of the problems associated with the implementation of the system.

In this research, our goal is to help stakeholders determine the impact of technological innovation on work practice. The research involves modeling the work of a hospital environment, with a focus on the information flow and workflow of physical, computerized, and human systems. Our models differ from other diagramming approaches because they include contextual information about the work being performed [4]. Such detail may help stakeholders (e.g., information officers, medical directors, and analysts) reason about the issues involved in the potential change, what types of technologies are currently needed, and what impact technologies will have on existing work practice. To support stakeholders' reasoning about the effects of implementing computerized technologies, models should meet three criteria: they should (1) convey an adequate description of the work; (2) help people evaluate the effectiveness of existing information flow and workflow; and (3) help people conjecture about the impact of technical or other change on future work.

As part of previous work, we developed a model notation and structure that identified how information traveled between hospital entities, and how it was transformed by people using paper-based and electronic information systems. Additional model notation specified areas where the work routine was flexible ("variability in practice" or ViP circles) and provided some information about the context of work (cloud symbols). The model of information flow and workflow was based on an analysis of interviews with healthcare professionals in the Intensive Care Unit and Blood Bank at a tertiary hospital in Queensland, Australia [5]. The model represented a non-urgent blood transfusion request process for a patient in the Intensive Care Unit, and was developed in several iterations with Microsoft Visio™ (Figure 1).

Corresponding author: Cara Stitzlein. Email – cara.stitzlein@uqconnect.edu.au

There is no standard framework for the evaluation of conceptual models; however, a model is often improved with an empirical test of its quality [6]. The use of non-standard symbols also requires evaluation, as this distinguishes our modeling approach from other ways of diagrammatically representing work processes [4,7]. The symbols must be effective if models built with them are to support stakeholders' abilities to understand the work and reason about change in the work system. In this paper, we report initial findings from a test of whether the models meet the three criteria outlined earlier: (1) whether people can describe the work adequately; (2) whether people can evaluate the effectiveness of the current work; and (3) whether people can make conjectures about the impact of change.

Methods

Study Design

We created a series of tests to assess how well the model meets the criteria of supporting description, evaluation, and conjecture. The University of Queensland's Human Research Ethics Committees granted approval for the study (approval no 11-PSYCH-PHD-38-JJ).

Description

1. **Symbol description test.** Participants wrote a description of what they thought 32 symbols meant, when the symbols were viewed in isolation. Participants rated their confidence in each of their responses on a 5-point scale.

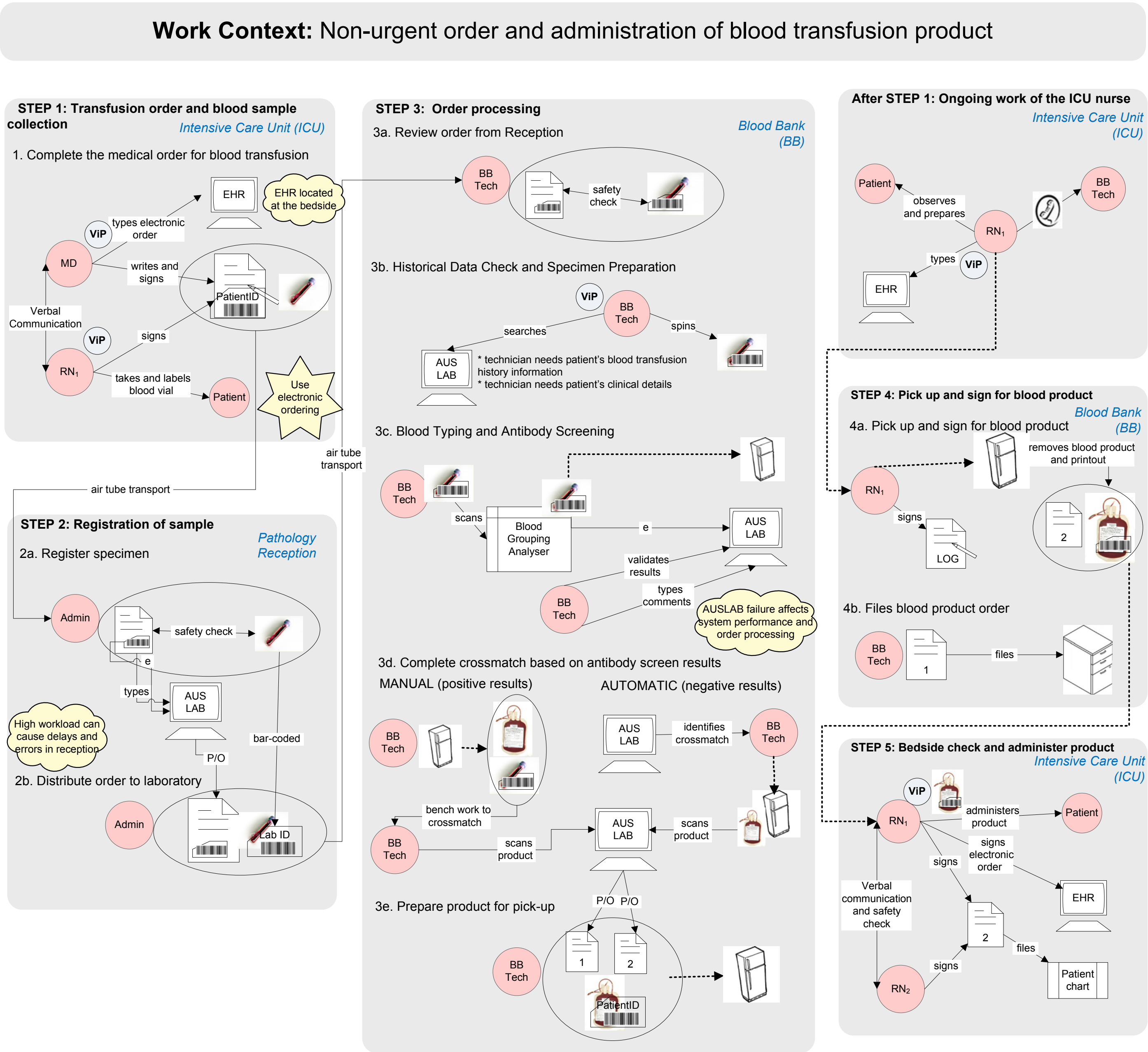

Figure 1. *A model of information flow and workflow represents a blood transfusion process. Cloud and star symbols were added by researcher when testing whether the model supports evaluation and conjecture about information flow and workflow.*

2. **Symbol combination test**. Participants were provided with a legend (Figure 2) for the symbols, and wrote a description of what they thought 13 different combinations of the symbols meant. They rated their confidence in each of their responses on a 5-point scale.
3. **Model walkthrough test**. Participants viewed a model of information flow and workflow of a critical care work situation for five minutes (Figure 1). Then participants provided a vocal overview of the entire process and a detailed description of selected steps.

Evaluation

4. **Evaluation of work test**. Participants gave two responses about ViP symbols: first, their perception about what the symbol meant about work performance, and second, alternative ways in which the work could be performed. These questions were asked for ViP symbols within the following symbol combinations:
   a. Verbal communication between the medical and nursing staff (Step 1).
   b. The search for patient information using a database (Step 3).
   c. The continuous activities of a nurse (ongoing work step).

For each of the three clouds in Figure 1, participants were asked if the additional information about the work helped them to understand the work any better, and they rated their response on a 5-point scale.

Conjecture

5. **Conjecture test**. Participants considered the possible changes with the introduction of an electronic ordering system in Step 1: “Transfusion order and blood sample collection” (see Star on Figure 1). The participant reported:
   a. Possible positive effects on information flow.
   b. Possible negative effects on information flow.
   c. Possible positive effects on workflow.
   d. Possible negative effects on workflow.

There was a short discussion about how the model structure would change, and whether the change would be beneficial or detrimental to the overall work process. Participants used a 5-point scale to rate the confidence of their judgment of how a new computerized system would affect the work associated with a blood transfusion process.

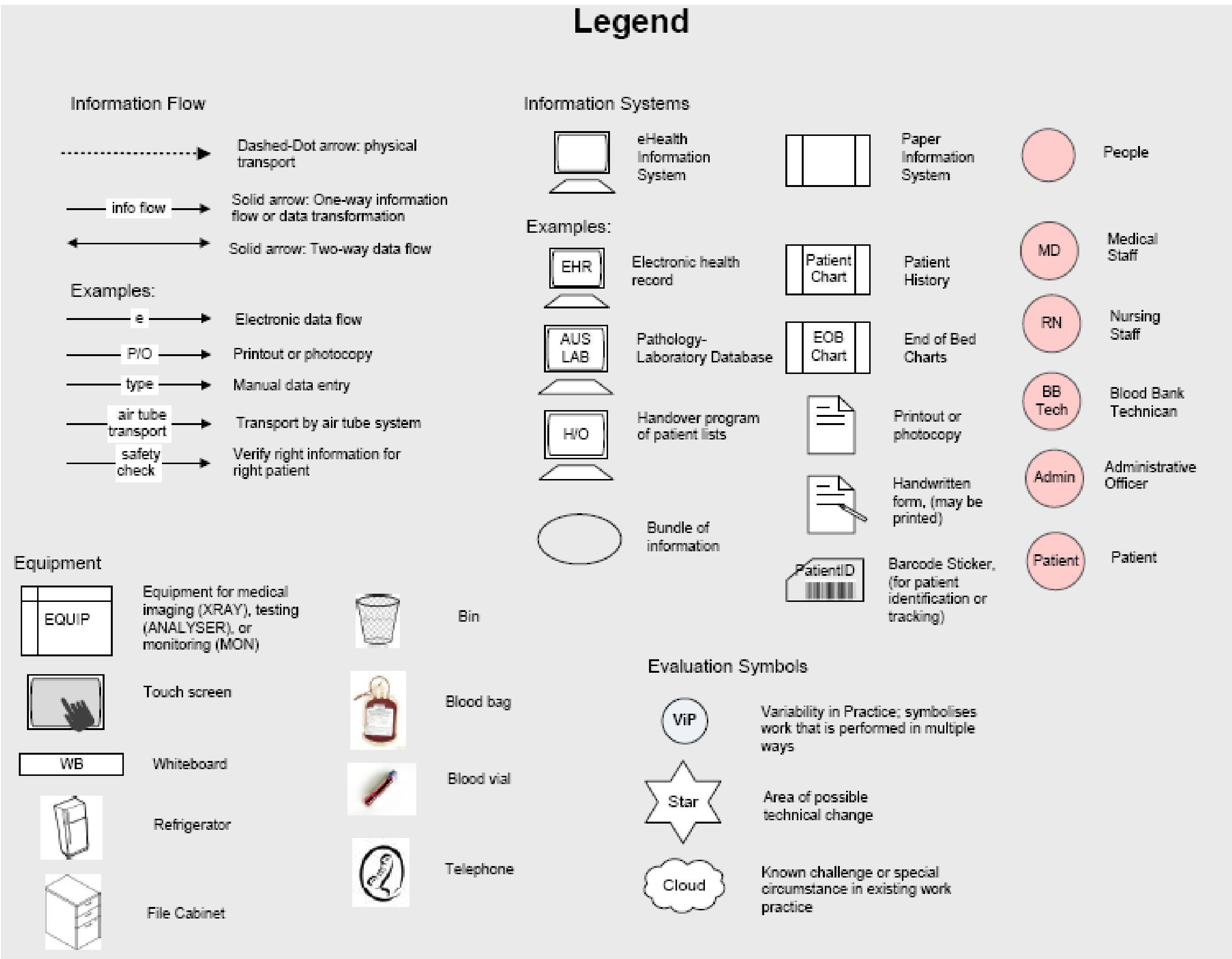

***Figure 2.** Legend given to participants with the model notation and intended meanings. Symbols are grouped into categories of (1) arrows representing information flow (top left), (2) different kinds of information systems (top right), (3) equipment (lower left), and (4) evaluation symbols (lower right).*

Materials

For the symbol description test, symbols were taken from the legend and were tested using a standardized form. Figure 3 shows an excerpt of the form. For the symbol combination test, symbol combinations were taken from the model in Figure 1 and tested with a standardized form. Figure 4 shows an excerpt of the form.

Item No.	Symbol	Symbol Description	Confidence Rating				
	—e→	(Electronic information flow, one-way)	Not at all Confident 1	Not very Confident 2	Somewhat Confident 3	Confident 4	Very Confident 5
	EHR	(Electronic health record program on this computer)	Not at all Confident 1	Not very Confident 2	Somewhat Confident 3	Confident 4	Very Confident 5

Figure 3. *Examples of items in the symbol description test. (Symbol descriptions were not shown when test was presented to participants).*

Item No.	Symbol	Symbol Description	Confidence Rating				
	MD1 MD2	(Grouping of two medical staff in a step)	Not at all Confident 1	Not very Confident 2	Somewhat Confident 3	Confident 4	Very Confident 5
	ViP, BB Tech, searches, AUS LAB	(There is variability in how the Blood Bank technician searches for information in the Pathology database)	Not at all Confident 1	Not very Confident 2	Somewhat Confident 3	Confident 4	Very Confident 5

Figure 4. *Examples of items in the symbol combination test. (Combination descriptions were not shown when test was presented to participants).*

For the model walkthrough test, evaluation of work test, and conjecture test, the model in Figure 1 was printed on a large sheet, so that the participants could easily see the notation. A version without cloud and star symbols was used for the model walkthrough test. For the evaluation of work test and the conjecture test, the cloud and star symbols were written on adhesive cutouts and placed on the model.

Procedure

Participants were run individually. The study took place in a quiet office setting on a university campus. Each test and question was presented in a standardized way to each participant. After completing the final test, the participant was asked to suggest alternate ways of representing symbols and relationships that the participant had found confusing. Audio recordings of the model walkthrough test, evaluation of work test, and conjecture test were captured with a smart pen, and a backup audio recording was captured by a voice recorder.

Participants

Participants (N=8) were university staff members and post-graduate students. There were five males and three females with an average age of 27.3 years (SD= 4.30). Participants had a background in either psychology (n=4) or engineering/information technology (n=4). Participants were given a small gratuity for their participation.

Data analysis

Quantitative analysis

The written responses on the symbol description and symbol combination tests were reviewed for how well the meaning of the symbols and the relationships in the model were inferred by participants. Descriptions were given a score of 100% for a fully correct response (the description matched the intended meaning), 50% for a partially correct response, and 0% for an incorrect response. Accuracy scores were assessed in the model walkthrough test through a review of the audio recordings. For the description to be scored 100% correct, the participants' overviews needed to include six major steps and locations. For descriptions of sub-steps, all symbol combinations needed to be included for a fully correct response. We analysed participants' ratings of their confidence by calculating an average based on the confidence ratings for each of their written descriptions.

The accuracy scores and confidence ratings for each item identified symbols that were easy or difficult to interpret. Scores for the 8 participants were averaged to evaluate the overall understanding of model notation. The general rule for considering the notation to be understandable was for the accuracy of describing an item, averaged across participants, to be at or above 80% and for confidence ratings, averaged across participants, to be at or above 3.0 on the 5-point scale. T-tests for independent samples were used to evaluate differences in accuracy and confidence between participant subgroups.

Qualitative analysis

For the model walkthrough test and evaluation of work test, verbal responses were analyzed to see how the ViP symbol was interpreted, whether alternate ways of work were reported and whether understanding was improved with additional information. For the conjecture test, a review of the verbal responses of participants was performed to see what positive and negative effects to changes in the information flow and workflow were reported. Similar responses were grouped for results reporting.

Results

Results are presented under the three criteria of description, evaluation, and conjecture. An overview of performance scores on the symbol description test and symbol combination test is provided in Table 1.

Description

Symbol description test. Overall, 14 of the 32 items met the criteria of being recognized with 80% accuracy or more and with a confidence level of 3 or more on the 5-point scale. Examples of correctly identified symbols were physical objects, such as a telephone, refrigerator, and barcode labels. Examples of items described incorrectly and with low confidence were the annotated arrows, the abstract symbols used to represent equipment and information systems, and details of the work. All participants described the variability in practice circle incorrectly as "very important person."

When looking at differences between participants based on their background (psychology or engineering/IT), both groups have an average accuracy of 69% (p=1.0), but the engineering/IT group gave statistically significantly higher confidence ratings (M= 3.38/5) than the psychology group (M= 2.70/5, $p<0.02$).

Symbol combination test. Each of the 13 combination items met the criteria of being described with 80% accuracy or more (average accuracy was 91%), and with a confidence rating of 3 or more on the 5-point confidence rating scale (average confidence was 4.13/5). There was no statistically significant difference in either accuracy (p=0.19) or confidence (p=0.61) between participants with psychology backgrounds and engineering/IT backgrounds for the symbol combination test.

Table 1. *Mean Completion Times and Performance Scores of Participants with Different Educational Backgrounds (Standard Deviation in Parentheses)*

Participant Background	Symbol Description Test			Symbol Combination Test		
	Mean completion time (min)	Mean accuracy score (%)	Mean confidence rating (/5)	Mean completion time (min)	Mean accuracy score (%)	Mean confidence rating (/5)
Psychology	9.5 (.60)	69% (.3)	2.7 (1.0)	11.8 (1.3)	93% (.1)	4.1 (.6)
Engineering/IT	10.0 (2.2)	69% (.3)	3.4 (.9)	10.0 (2.7)	89% (.1)	4.2 (.6)
All participants	9.8 (1.5)	69% (.3)	3.0 (.9)	10.9 (2.2)	91% (.1)	4.1 (.6)

Model walkthrough test. The average time taken by participants to provide an overview of the model was 1.9 minutes (SD=1.9). Seven participants described the multi-step process shown in Figure 1 by including the main steps and locations in the representation. One participant was scored with a partially correct response for failing to mention a major location of the work. All participants described the communication between medical and nursing staff in Step 1, although 3/8 participants were uncertain about how the step was initiated. Half of the participants expressed confusion about Step 3d, where work was performed in an automated or a manual process, depending on the outcome of the previous step. All participants accurately reported the work of the ICU nurse as happening in parallel with work in the Blood Bank.

Evaluation

When asked to describe ViP symbols in the evaluation test, participants reported alternate ways of work such as using different modalities of communication (6/8 participants), different information sources for retrieving information (5/8 participants), and different protocols to get the work done (4/8 participants). Five participants reported being confused about how the ViP notation was being used in the model. The symbol's placement and the participants' lack of clinical experience contributed to their confusion.

Cloud symbols provided information that was otherwise unknowable to participants about the context of work. Of the three clouds used in this study, no consensus existed between the participants if information was entirely helpful or unhelpful to the understanding of how work was performed. A common response from participants was that while the clouds provided more information, the information did not tell them how the work was affected by the information nor what kind of inferences they were expected to make from the information (e.g. which step changed or what solution may exist).

Conjecture

In the conjecture test, each participant gave examples of possible changes to information flow and workflow given the introduction of an electronic ordering system. When asked if they considered the change to be good or bad, six participants said that they were unsure, reporting that it was "hard to tell" and that it "depended on how it was implemented." Possible positive outcomes reported by participants included increased efficiency of the work (e.g., faster transport of request to other units) and greater worker satisfaction (e.g. less double documentation by medical staff). Almost all (7/8 participants) reported that safety checks needed to remain, but were unsure how this would be handled by the electronic system, and what safeguards would be in place to prevent "muddled information."

Discussion

The successful implementation of new healthcare technologies is affected by technical factors such as computer availability and system reliability, and contextual factors, such as the needs of the current health environment and the performance of clinical work. A tool that includes both technical and contextual factors could support reasoning about the whole change process, rather than just a part of it. We have presented a modeling approach and an initial verification of the resulting model, and hope that this type of data modeling can help solve current technology and healthcare dilemmas [1, 8-10].

The results from an initial test suggest that the model notation is sufficiently descriptive and understandable for a sample of non-healthcare participants, but only once the model's legend is available. This claim is supported by the increase in (average) accuracy scores in the symbol description test to the symbol combination test (from 69% to 91%), and the increase in (average) confidence (from 3.0/5 to 4.1/5). The finding that the engineering group was more confident, but not more accurate in the symbol description test suggests that educational background may be a factor in how the notation is perceived. The verbal reports from this sample suggest that naïve model users can identify relevant technical and work-related issues. The participants' responses are consistent with reported shortcomings of paper-based records such as double documentation and legibility issues, and perceived benefits of electronic systems, such as increased information security and accessibility [1].

There are limitations to this study. First, findings from a small sample may not generalise to a larger sample, and there is a risk that the findings may not generalise to the intended audience of healthcare stakeholders. Second, the method of calculating accuracy scores is based on a single researcher's scoring, and may not be fully reliable. Third, participants were not asked to provide a detailed description for all steps in the model. The assessment is therefore based only on portions of the model. We will extend the study to participants who have a background in healthcare and calculate inter-rater reliability scores for participants' responses in order to overcome some of these limitations.

Our next steps are to make changes to the less successful aspects of the modeling notation that were identified in the analysis. The changes should improve the effectiveness and the quality of the models. Then we will test new models with participants who have professional healthcare experience. We will assess whether such participants understand the notation, whether they believe that the models are correct representations of their work, and whether the models are at the appropriate level of abstraction for their purpose [10].

Once a model's accuracy is verified, we can then use it for its intended purpose, which is to help stakeholders make decisions about technology in healthcare environments. We hope the outcomes of our research will allow stakeholders to evaluate better whether a known technical change could lead to unintended negative consequences or whether it will probably yield important benefits (e.g., reduce transcription errors, provide faster information availability, or improve quality of care).

Acknowledgments

This research is supported by ARC Discovery Project DP0880920 to Sanderson and Venkatesh and by an International Postgraduate Research Scholarship to Stitzlein. We thank the participants in the model development and model test studies for their participation. NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program.

References

1. Whetton S. *Health informatics: A socio-technical perspective*. Victoria: Oxford UP; 2005.
2. National E-Health Transition Authority. *NEHTA annual report 2009–10*. Sydney (NSW): NEHTA; 2010 Nov. 63p.
3. Patrick J. *A Study of a Health Enterprise Information System*. University of Sydney (NSW): School of Information Technologies; 2011. 190p. Technical Report TR673
4. Leu LG. Cheung M. Webster TR. Curry L. Bradley EH. Fifield J. Burstin H. Centers speak up: The clinical context for health information technology in the ambulatory care setting. *Journal of General Internal Medicine*. 2007; 23(4): 327-328.
5. Stitzlein C. Sanderson P. Beltran Orihuela C. Jack L. Venkatesh B. Evaluating the impact of technological change in a critical care unit: towards a model to support stakeholder envisionment. *Proceedings of the 55th Annual Meeting of the Human Factors and Ergonomics Society*; 2011 Sep 19-23; Nevada, USA. In press.
6. Moody DL. Theoretical and practical issues in evaluating the quality of conceptual models: current state and future directions. *Data and Knowledge Engineering*. 2005; 55: 243-276.
7. Madison D. *Process mapping, process improvement, and process management: A practical guide to enhancing work and information flow*. Chico: Paton Press LLC; 2005.
8. Eason K. Changing perspectives on the organizational consequences of information technology. *Behaviour & Information Technology*. 2001; 20: 323-328.
9. Baxter G. Sommerville I. Socio-technical systems: From design methods to systems engineering. *Interacting with Computers*. 2011; 23: 4-17.
10. Berg M. Toussaint P. The mantra of modeling and the forgotten powers of paper: A sociotechnical view on the development of process-oriented ICT in health care. *International Journal of Medical Informatics*. 2003; 69: 223-234.

Cite this article as: *Stitzlein et al.* Testing a model to help analysts understand, evaluate, and make inferences about health technology change. *Ergonomics Australia – HFESA 2011 Conference Edition* 2011 11:44.