PERGAMON

Computers in Biology and Medicine 32 (2002) 141–149

Computers in Biology and Medicine

www.elsevier.com/locate/compbiomed

Evaluation in the design of health information systems: application of approaches emerging from usability engineering

Andre Kushniruk *

Department of Mathematics and Statistics, Information Technology Division, York University, 4700 Keele St., Toronto, Ont., Canada M3J 1P3

Accepted 1 November 2001

Abstract

This paper examines the role of evaluation in the design of health care information systems. A framework is presented for considering evaluation in the context of software development processes, in particular, the systems development life cycle (SDLC). Variations on standard design methodologies are then discussed, including methods based on rapid development and continual evaluation of prototype systems. Usability testing is presented as a key method for conducting evaluations during iterative system development. The emergence of design methodologies, where evaluation is viewed as a central part of the development cycle is also discussed. Evaluation methodologies are then considered along a continuum, ranging from studies involving a high degree of experimental control to observational approaches. A full cycle approach to evaluation of health care systems is argued for, involving deployment of new methods across the SDLC. Implications for future work exploring the integration of design and evaluation processes in health informatics are discussed.

Keywords: Health information systems; Technology evaluation; Usability engineering; Usability testing; User interfaces; Systems analysis and design

1. Introduction

In the field of health informatics, summative evaluations have been conducted to evaluate the impact of a wide variety of systems in clinics and other health care environments. Numerous evaluations have assessed the effects of health information systems on dependent measures such as health outcomes, length of stay, cost-effectiveness, quality of care and other measures [1]. These studies typically apply randomized clinical control trials methodologies. Such evaluations are analogous to the testing of drugs and other discrete health care interventions where there is a clearly defined

* Tel.: +1-416-736-2100x33980; fax: +1-416-736-5757.
E-mail address: andrek@mathstat.yorku.ca (A. Kushniruk).

0010-4825/02/$ - see front matter
PII: S0010-4825(02)00011-2

independent variable (e.g., presence or absence of a drug, or in the case of health informatics studies, presence or absence of an information system). Although these types of evaluations are necessary in order to ensure that systems that are developed are both safe and effective, greater emphasis is needed to ensure that the *process* of system design is effective. In addition, it can be argued that we will not understand the full meaning of outcomes we obtain from the use of information technology, unless we can understand outcomes in terms of the design decisions that lead to the outcomes.

A wide range of methodologies have been developed in the general software industry for providing frameworks for designing information systems. Life cycle models provide a foundation for guiding the development and evaluation of complex systems [2,3]. Traditional system development life cycles (SDLC) presuppose a set of fixed stages for system development, with evaluation of the system predominantly occurring in the final stages. Such approaches have proven difficult to successfully apply in health care where information needs may be hard to precisely determine. The possible reasons for this are that the health care environment is often complex and characterized by missing information, shifting goals and a great deal of uncertainty. Health care decision making processes are complex, poorly understood and consequently difficult to model effectively in the analysis of the SDLC. Health care decisions are subject to a level of uncertainty not found in traditional business environments and consequently health care technology and the knowledge on which it is based is often very volatile. In fact, even before decision making processes are understood, they may change within the time span of the traditional SDLC.

In recent years, a number of software engineering methodologies have been developed that focus upon deploying evaluation methods *throughout* the software life cycle—from an initial needs analysis through to design and implementation, in addition to summative evaluation conducted upon completion of the system. For purposes of developing an evaluation framework this paper will initially consider methodologies in terms of the traditional SDLC, followed by methods involving rapid iterative development, based on a usability engineering approach. The traditional SDLC is characterized by the following phases: (1) planning, (2) analysis, (3) design, (4) implementation, and (5) maintenance/support [2,3]. In this model each phase consists of similar activities that are somewhat sequential, with one phase typically being completed prior to beginning the next phase.

A wide range of iterative design methods have appeared, where the distinction between the phases has become blurred and the role of continual evaluation is emphasized. For example, methodologies such as rapid application development (RAD) and various approaches to prototyping have become increasingly popular as the basis for system development [3]. Using such approaches, system requirements are quickly converted to a prototype system that is iteratively revised, based on evaluations, until an acceptable product is developed and the system is completed [3]. Continual evaluation is a central component of these design methodologies. Along these lines, a greater emphasis on iterative formative evaluation, closely integrated with the design and development process, has been called for in the health care information technology industry [4]. In conjunction with this, deployment of a wider range of methods for acquiring information relevant to design in health informatics evaluations is also argued for.

2. Evaluation in the health information system development life cycle

Evaluation in health informatics spans a continuum from project planning to design and implementation. Fig. 1 relates a number of approaches that have been used in the evaluation of health

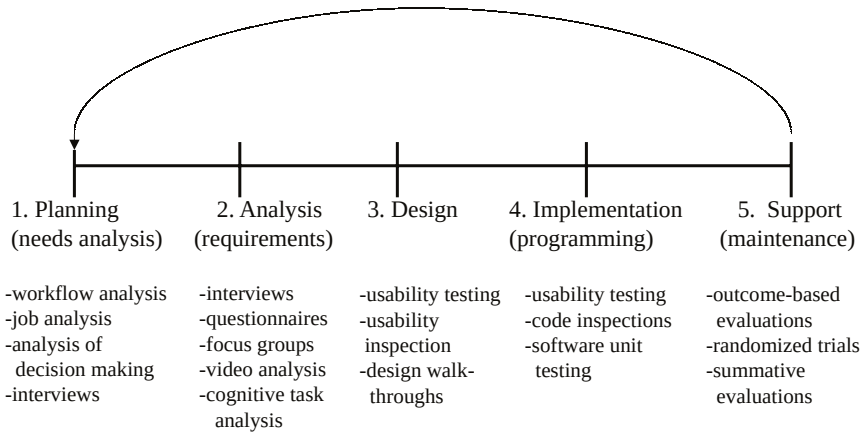

Fig. 1. The systems development life cycle (SDLC) in relation to evaluation methodologies.

information systems during the phases of the SDLC. For example, methods for analyzing work activity (including workflow analysis and job analysis) are potentially essential components in the initial assessment of user needs and system requirements early in system development. During the analysis phase a number of evaluative methods may be employed, ranging from traditional systems engineering methods (including observation, interviews and deployment of questionnaires) to formal evaluations of work places in terms of socio-cognitive factors (including methods of cognitive task and workflow analysis in the context of distributed collaborative activity [5,6]). During the design and implementation phases a number of methods from the field of usability engineering are becoming increasingly more important when conducting formative system evaluations. As can be seen from Fig. 1, formal summative evaluation comprises only one aspect of evaluation during the latter stages of the cycle. In the larger computer software development industry emphasis is being placed on evaluation methods applied to a number of different points in the SDLC. According to Friedman's and Wyatt's framework for discussing evaluation in health care, many of these formative approaches fall into the category of "subjectivist approaches" [1]. As will be discussed, the deployment and methods of analyses associated with these methods, in particular usability testing, have become increasingly refined and objectified over the past few years.

3. The role of usability engineering in the evaluation of health care information systems

Important developments in software engineering have emerged with variations in the SDLC based on principles of iterative design. A driving force in these advances has been the emergence of a wide variety of information systems. Increasingly, the traditional SDLC, with its fixed and rigid order of phases, is being identified as not being well suited to rapid development of highly interactive Web-based applications [7]. For example, Web-based patient record systems, decision support and educational applications constitute new forms of computer applications, that require innovative design methodologies. In contrast to traditional systems applications (e.g., payroll, billing, etc.) many newer applications may require continual, iterative feedback from potential end users and clients during both design and development. There are a number of reasons why such applications may require continued

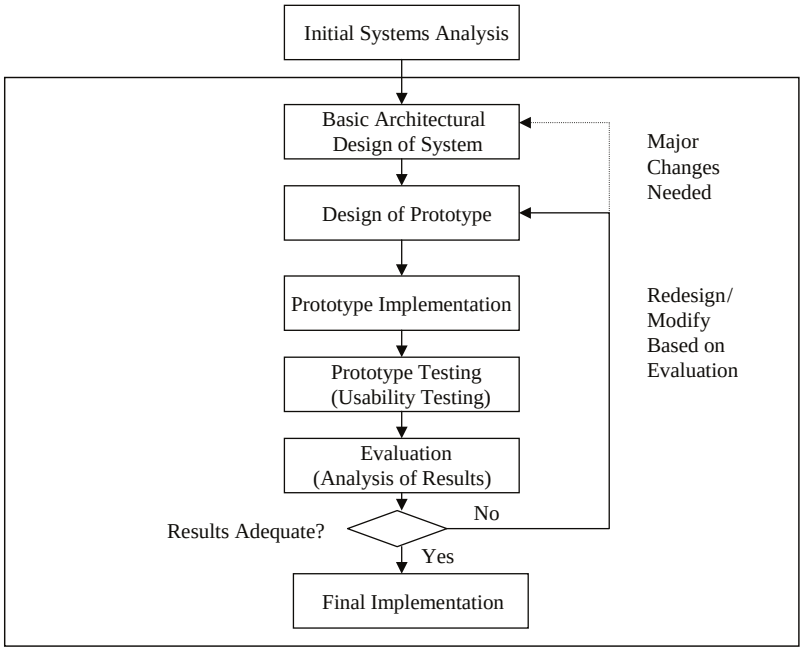

Fig. 2. Systems development based on prototyping and iterative usability testing highlighting the role of evaluation in system design.

user input, including the predominant role of the user interface in such highly interactive systems (particularly Web-based systems), as well as the exploratory nature of many advanced applications, where it may be impossible to fully specify what will be feasible or desired (in terms of features and functionality) prior to starting system design.

The role of evaluation in iterative design and testing has been highlighted in many newer design methodologies. Iteration in this context involves repeating or looping through the same development activities, successively refining the system in each cycle. Typically this involves initial development followed by evaluation and feedback into system design, leading to further cycles of evaluation and redesign until a satisfactory system arises. In contrast to traditional SDLC, analysis is not finalized at the beginning of system development, but recurs throughout the process.

Fig. 2 depicts a systems cycle based on the development of prototypes representing evolving models that will lead to the final software product. Prototype development and iterative evaluation are embedded in this process. As shown in Fig. 2, iterative systems analysis and design is similar to traditional design in that initial systems analysis is performed. However, the approaches diverge considerably after the basic architectural design has been mapped out. At this point a prototype is constructed. Prototypes vary from simple "mock ups" of the system design to partially working models of the system that may contain some basic functionality [3]. The prototype is evaluated during the design phase and further cycles of redesign and testing take place until the system meets user expectations and final implementation can take place (see Fig. 2). In this model of system development evaluation is central. Evaluation spurs the development and continual redesign of the system.

3.1. Usability testing

Fig. 2 illustrates the role of evaluation in systems prototyping and the need for usability engineering. Perhaps the most powerful of these methods is usability testing. Usability testing refers to the evaluation of information systems through the analysis of typical end users interacting with the system. The phrase "discount usability engineering" was coined to describe the application of low cost methods for conducting such tests [8]. More recently, a trend towards "virtual" usability engineering, involving remote testing, has emerged and will be described below [9]. The basic strategy entails video recording all human–computer interaction (i.e., video recording all computer screens) and audio-recording all subject verbalizations as they interact with system prototypes. Such studies are being conducted using methods from experimental psychology, where subjects are asked to "think aloud" while interacting with the system. Based on data collected from a representative sample of subjects, typically involving as few as 8–10 participants per study, the majority of usability problems and issues can be identified and summarized [10]. For example, subjects consisting of representative users of a system (e.g., physicians or nurses) are asked to interact with a system or prototype to perform a task, such as entering patient data. Subjects are also typically asked to verbalize their thoughts while doing the task. The complete audio and video recording of the interaction can then be analyzed using methods involving the coding and classification of user problems (see [10] for details). The resulting information from such evaluation is summarized and presented to the designers, allowing for iterative modifications to the system, as shown in Fig. 2. In the development of a complex system, there may be several cycles of design and testing involving usability engineering before the system is completed. It has been reported that full usability testing can be efficiently inserted into the design cycle of a patient record system, with a single initial cycle of design-evaluation-redesign leading to as much as a ten-fold reduction in usability problems [10].

In health care, an increasing number of laboratories are applying usability engineering for formative evaluation of information systems. A range of methodologies involving video-based analysis of user interactions with prototype information systems have been developed [11–14]. Recent work has included the application of remote usability testing for distance evaluation of Web-based health care information systems. For example, tools that allow for remote access of end users' screens over the Web, can be used to make recordings of user interactions with a system under study [9]. Information collected from such process-centered evaluation can be used to redesign a system. This information can include: (1) suggestions by users for improvements to both the user interface and system functionality, (2) identification of usability problems such as lack of consistency in interface operations, and (3) quantitative measurements including time to task completion and system response times [15]. Such methods were initially employed in the analysis of user interfaces. They can now be used more broadly and can include experimental testing focusing on users' satisfaction with basic system functions or quality of information provided [16]. For example, in testing Web-based clinical guidelines physician users may be encouraged (during usability testing) to comment on guideline content and applicability to medical practice [9]. The results of such evaluation can then be fed back into the redesign of the guideline format in order to increase uptake and satisfaction with the information provided [9].

A usability engineering approach to system development is particularly appropriate when developing systems where user requirements may be difficult to obtain using standard methods (such as interviews and questionnaires) and technical feasibility for some system functions may be unknown or

uncertain. Many of the more advanced applications in health informatics, including decision support tools, patient record systems and educational tools are amenable to development using an iterative approach, involving prototyping, as described above [10]. In addition, the availability of prototyping tools, in particular Web-based development software, are making this approach increasingly attractive. In contrast, iterative methods may be less useful in developing systems characterized by a low degree of interactivity (therefore requiring less input from end users) and more standard applications, which can be based on design of previous systems (requiring less fine-tuning or adapting to particular end user needs). However, as health care applications inevitably become more complex and the demand for customizing systems to particular user need increases, usability engineering methods are likely to become critically important.

4. A continuum of methodological approaches to system evaluation

In the discussion above, approaches to evaluation in health care were considered within the context of development phases, from project planning to ultimate implementation and support. Considering evaluation from this perspective underlines the importance of evaluation *throughout* the process of software development, and forms the basis for an increasing variety of evaluations in health informatics. In this section, we consider a second dimension for evaluation based on the type of evaluation method, ranging on a continuum from experimental to naturalistic/observational methods, as depicted in Fig. 3. At one end are studies where an attempt is made to control all factors in the evaluation and vary only a single independent variable (e.g., the presence or absence of an information system in evaluations attempting to follow a controlled trials approach). There have also been a number of arguments made that a high degree of variable control may be neither feasible nor desirable when testing systems in real-world contexts, in particular when attempting to achieve a greater degree of generalizability to complex real-world situations [17,18]. In the context of usability studies, experimental approaches might involve laboratory testing of subjects interacting with computer systems with only one or few variables (e.g., display format) manipulated during the testing, with the test being conducted under controlled artificial conditions in a usability laboratory.

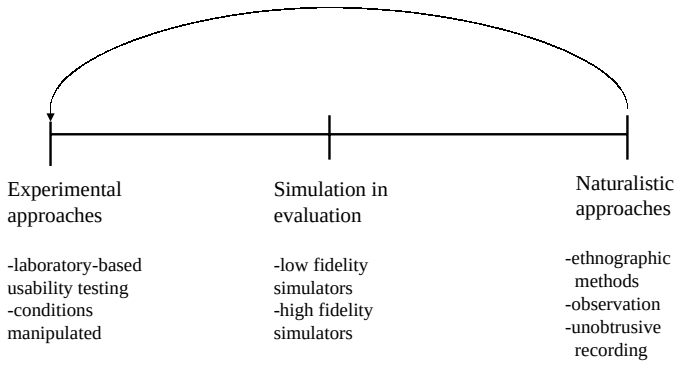

Fig. 3. A continuum of approaches to evaluation of health care information systems.

From Fig. 3, it can be seen that evaluations involving simulation techniques are half-way along the continuum ranging from controlled to naturalistic approaches. Evaluations based on simulations may allow for a high degree of experimental control while also maintaining a high degree of realism in the situations presented to subjects during testing (e.g., high fidelity computer-based simulations). Along these lines, in a number of reported evaluations of computer patient records, physicians' interactions with the system under evaluation were recorded while the physician interviewed a "simulated patient" (i.e., a collaborator playing the part of a patient to maintain a consistent stimulus) in an approach modified from usability testing [19]. Other forms of evaluation involving simulators could involve the use of highly realistic computer-controlled mannequins, which serve as patients and respond to inputs such as injections, for assessing areas such as operating room work activities and team decision making involving advanced information technology [20]. These types of evaluations are similar to those conducted in the aviation industry, where use of computer-controlled flight simulators for evaluating pilot learning and airplane construction have been routinely used to assess training [21].

At the other end of the continuum depicted in Fig. 3 are evaluative studies involving observational and naturalistic approaches. From this perspective, activities and interactions with systems may be monitored or recorded, using audio or video. Here the approach is designed to be as unobtrusive as possible, with little or no experimental control. In areas such as the study of decision-making in complex domains, the use of naturalistic study has been promoted by an increasing number of researchers [22]. Proponents of naturalistic approaches to evaluation have argued that much of the research from "classical" controlled experimental studies has not led to results that are generalizable to real-world situations in complex domains [22].

Within the same set of evaluative studies it may be useful to begin at one end of the continuum, for example conducting initial naturalistic observation of the effects of a patient record system that is under development, and then move to studies that examine more precise variables under different (e.g., more controlled) conditions. In contrast, an evaluation of a health care system might begin with a more obtrusive and controlled artificial approach, for example, usability testing of the system with users in the laboratory. This work might lead to hypotheses as to the effect of system use which could then be evaluated in the larger context of a clinic or hospital using observational or naturalistic methods.

5. Discussion and future work

It is becoming increasingly acknowledged that there are a wide range of possible choices for conducting evaluations in health informatics. This paper has presented a framework for considering evaluation in terms of the SDLC and its variants. From this perspective, research in health informatics evaluation needs to become more closely tied to emerging developments in systems engineering (e.g., advances in methodologies such object-oriented analysis and design [23]). With developments in systems and usability engineering, the interrelation between evaluation and design will continue to become more complex. In general there has been a move from health informatics evaluations focusing exclusively on measuring outcome variables to evaluations involving collection of in-depth process data, for example, data collected on use of a system by subjects performing complex tasks. This has led to a broader perspective on evaluation in health information systems than previously proposed [1]. A challenge for future work will lie in the integration of data collected from multiple

evaluation methods. This may involve conducting evaluations at several points during design and implementation, as well as strategically varying the nature of the evaluation itself within a full cycle of related studies. The complexity of health care environments and decision making will necessitate application of new design methodologies that support an increased degree of user input and evaluation.

An important area that will need to be addressed in future work is the potential relationship and integration of methods focusing on examining process variables (e.g., usability testing) with methods involving measurement of outcome variables, involving the use of more traditional approaches to summative evaluation of health care systems. Furthermore, as the information technology we assess inevitably becomes more complex, evaluation methodologies will need to be continually refined in order to keep pace, making design and evaluation of effective health information systems a challenging and continually evolving process. However, the greatest challenge to information system development in health care, may well be improving our understanding of decision making in complex health care environments.

Summary

This paper has considered the role of evaluation in the design of health care information systems. It has been argued that evaluation must be considered throughout the entire systems development life cycle (SDLC) in creating health care applications. Modern design methodologies based on rapid development and iterative prototyping rely on formative evaluation in order to provide designers with the input needed to improve systems. Methods emerging from the field of usability engineering, in particular usability testing, are essential for conducting such evaluations. The paper concludes with a framework for considering evaluation methods ranging from controlled experimental approaches to naturalistic observational approaches. Suggestions for further research in advancing the state of evaluation in health informatics are discussed.

Acknowledgements

I wish to thank Dr. Vimla Patel and my colleagues from the Centre for Medical Education at McGill University. Thanks for their support and encouragement during my years at McGill. I would also like to thank James Cimino at Columbia University for his contributions and valuable advice regarding many aspects of health informatics. In addition, I would like to acknowledge HEALNet for its continued support. Finally, I would like to thank Elizabeth Borycki for editing drafts of this paper.

References

[1] C.P. Friedman, J.C. Wyatt, Evaluation Methods in Medical Informatics, Springer, New York, 1997.
[2] J.A. Hoffer, J.F. George, J.S. Valacich, Modern Systems Analysis and Design, Second Edition, Addison-Wesley, Don Mills, Ont., 1999.
[3] S. McConnell, Rapid Development: Taming Wild Software Schedules, Microsoft Press, Redmond, Washington, 1996.
[4] A.W. Kushniruk, V.L. Patel, Cognitive evaluation of decision making processes and assessment of information technology in medicine, Int. J. Med. Inform. 51 (1998) 83–90.
[5] V. Mehta, A. Kushniruk, S. Gauthier, Y. Richard, E. Deland, M. Veilleux, A. Grant, Use of evidence in the process of practice change in a clinical team: a study forming part of the Autocontrol project. Int. J. Med. Inform. 51 (1998) 169–180.

[6] V.L. Patel, A.W. Kushniruk, S. Yang, J.F. Yale, Impact of a computer-based patient record system on data collection, knowledge organization and reasoning, J. Am. Med. Inform. Assoc. 7 (6) (2000) 569–585.
[7] J.J. Cimino, S.A. Socratous, Just tell me what you want!: the promise and perils of rapid prototyping with the World Wide Web, in: J.J. Cimino (Ed.), Proceedings of the 1996 Fall AMIA, Hanley & Belfus Inc., Philadelphia, 1996, pp. 719–723.
[8] J. Nielsen, Usability Engineering, Academic Press, New York, 1993.
[9] A.W. Kushniruk, C. Patel, V.L. Patel, J.J. Cimino, "Televaluation" of information systems: an integrative approach to design and evaluation of Web-based systems. Int. J. Med. Inform. 61 (1) (2001) 45–70.
[10] A.W. Kushniruk, V.L. Patel, J.J. Cimino, Usability testing in medical informatics: cognitive approaches to evaluation of information systems and user interfaces, in: D. Masys (Ed.), Proceedings of the 1997 AMIA Fall Symposium, Hanley & Belfus Inc., Philadelphia, 1997, pp. 218–222.
[11] A. Kushniruk, V. Patel, Cognitive computer-based video analysis: its application in assessing the usability of medical systems, in: R. Greenes, H. Peterson, D. Protti (Eds.), MED-INFO'95—the Eighth World Congress on Medical Informatic*s*, North-Holland, Amsterdam, 1995, pp. 1566–1569.
[12] A.W. Kushniruk, V.L. Patel, J.J. Cimino, R.A. Barrows, Cognitive evaluation of the user interface and vocabulary of an outpatient information system, in: J. Cimino (Ed.), AMIA Fall Symposium Proceedings, Hanley & Belfus Inc., Philadelphia, 1996, pp. 22–26.
[13] P.L. Elkin, D.N. Mohr, M.S. Tuttle, W.G. Cole, G.E. Atkin, K. Keck, T.B. Fisk, B.H. Kaihoi, K.E. Lee, M.C. Higgins, H.J. Suermondt, N. Olson, P.L. Claus, P.C. Carpenter, C.G. Chute, Standardized problem list generation, utilizing the Mayo canonical vocabulary embedded within the Unified Medical Language System, in: D. Masys (Ed.), Proceedings of the 1997 AMIA Fall Symposium, Hanley & Belfus Inc., Philadelphia, 1997, 500–504.
[14] J.M. Coble, J. Karat, M.J. Orland, M.G. Kahn, Iterative Usability Testing: ensuring a Usable Clinical Workstation, in: D. Masys (Ed.), Proceedings of the 1997 AMIA Fall Symposium, Hanley & Belfus Inc., Philadelphia, 1997, 744–748.
[15] J. Rubin, Handbook of Usability Testing, Wiley, New York, 1994.
[16] J.J. Cimino, V.L. Patel, A.W. Kushniruk, Studying the human-computer-terminology interface, J. Am. Med. Informat. Assoc. 8 (2) (2001) 163–173.
[17] J. Orasanu, T. Connolly, The reinvention of decision making, in: G. Klein, J. Orasanu, R. Calderwood, C. Zsambok (Eds.), Decision Making in Action, Ablex Publishing Company, Norwood, NJ, 1993, pp. 3–20.
[18] G.A. Klein, A recognition-primed decision (RPD) model of rapid decision making, in: G. Klein, J. Orasanu, R. Calderwood, C. Zsambok (Eds.), Decision making in action, Ablex Publishing Company, Norwood, NJ, 1993, pp. 138–147.
[19] A.W. Kushniruk, D.R. Kaufman, V.L. Patel, Y. Levesque, P. Lottin, Assessment of a computerized patient record system: a cognitive approach to evaluating an emerging medical technology. M.D. Comput. 13 (5) (1996) 406–415.
[20] D. Gaba, S.K. Howard, Situation awareness in anaesthesiology, Hum. Factors 37 (1) (1995) 20–31.
[21] J. Orasanu, E. Salas, Team decision making in complex environments, in: G. Klein, J. Orasanu, R. Calderwood, C. Zsambok (Eds.), Decision Making in Action, Ablex Publishing Company, Norwood, NJ, 1993, pp. 327–345.
[22] C.E. Zsambok, G. Klein, Naturalistic Decision Making, Lawrence Erlbaum Associates Publishers, Mahwah, NJ, 1997.
[23] G. Booch, J. Rumbaugh, I. Jacobson, The Unified Modeling Language User Guide, Addison-Wesley, Don Mills, Ont., 1999.

Andre Kushniruk is an Associate Professor in Information Technology with expertise in the areas of health informatics, evaluation and human–computer interaction in health care. He has written numerous papers on topics related to improving the usability of health care information systems and has been a key researcher on a variety of collaborative projects in the areas of medical informatics and usability engineering. His work includes the development of novel methods for conducting video analysis of computer users and he is currently extending this research to the remote study of system usability. He has experience consulting on a variety of large projects with hospitals throughout Canada and the United States in developing health care information systems. Dr. Kushniruk has taught at a number of universities in Canada. He holds undergraduate degrees in Psychology and Biology, as well as a M.Sc. in Computer Science and a Ph.D. in Cognitive Psychology. For the past ten years he has worked on health informatics related projects with a number of universities including McGill University, Columbia University, the University of Victoria, the University of Toronto and Mt. Sinai Medical School. Prior to his academic career, Dr. Kushniruk worked in industry developing innovative applications of knowledge-based systems.