

College of Computing

HCI area - Quantitative and Qualitative Modeling and Evaluation

Introduction

Two activities go hand-in-hand in a majority of HCI research: modeling and evaluation. Modeling addresses what you know about the user, and often their surrounding social and physical environment. A variety of existing models, such as the Human-Model Processor, and modeling techniques, such as Contextual Inquiry, address differing domains and levels of specificity. Models may be used to predict performance, organize field data, and describe potential interactions with a computer interface. As you read, examine the various models and modeling techniques that provide the foundation for the research. When will these models be useful in other research settings? What do you need to know to complete a model? How can you gather that information?

One use of models is to inform the evaluation of an interface. These activities are linked as the specificity and domain of the models constrains the questions that can be addressed in an evaluation. You will notice that specific, quantitative models are used to inform specific, quantitative evaluations. Likewise, more general, qualitative models are often the basis for various qualitative studies. The feasibility of combining various evaluation techniques is influenced by the compatibility of the underlying models. If the models make conflicting assumptions about the user, perhaps even disagreeing on what can or cannot be known, then the validity of combining evaluation techniques is in question.

One of the distinguishing characteristics of the HCI area in Computer Science is the importance of evaluation of how any computer-assisted system impacts its intended user population. Evaluation in HCI (and other human-centered disciplines) is quite different from evaluation in other areas of Computer Science, mainly because it is sometimes hard to construct experiments or observations that give definitive quantitative answers regarding the merit of one system over another. Instead, evaluation in HCI consists of demonstrating a scientific approach to answer questions about a systems relative merit in its context of use. This approach can consist of a myriad of techniques. Sometimes, a very reliable quantitative result is derivable, as is the case in narrowly-focused human motor observations such as a Fitts' Law experiment or a Keystroke-Level Model analysis. Other times, when the impact on work practices is sought, it is nearly impossible to control all influences in a natural setting. A student of HCI should become familiar with the variety of evaluation techniques and develop a sense of suitability of these techniques.

One of the best ways to achieve the ability to critique evaluation approaches is to read examples of evaluation work in the literature. As you read, critique the research based on the repeatability of the experimentation (Could a competent researcher reproduce the findings following the procedures described by the authors?) and the strength of the analysis and conclusions (Did the authors do enough to convince you of their evaluation results?) This is a particularly good way to assess quantitative results, and although this criteria can be used in the assessment of qualitative research, another useful criteria is to ask about the depth of explanation of the particular phenomena being reported.

One way to organize the information that you gather is to fill in the simple, 2x2 matrix:

	Modeling	Evaluation
Quantitative		
Qualitative		

You should pay attention to the horizontal connections between modeling and evaluation techniques. Likewise, notice the connections and disconnections between quantitative qualitative techniques.

General Resources

Surveys and detailed coverage of many modeling and evaluation techniques are covered in **CS 6750: Introduction to HCI**. Follow on courses such as **CS 6455: User Interface Design and Evaluation** address **Qualitative Methods**, while courses such as **PSYC 6018 Principles of Research Design** and **PSYC 7101 Engineering Psychology I: Methods** address **Quantitative Methods**. Being familiar with both, as well as being able to **pick** which methods are suitable for your own research is an essential skill for an HCI researcher.

Many papers in the SIGCHI conference series on ***Human Factors in Computing Systems***, also known as the CHI conference, include significant modeling and evaluation work, both quantitative and qualitative. This is also true of the CSCW conference series (both the ACM **CSCW** conference and the European **ECSCW** series), though CSCW research tends to include more qualitative modeling and evaluation. The ACM **UIST** conference usually does not emphasize modeling and evaluation as much, but there are occasional stellar papers that provide a judicious balance between technology development and evaluation.

Modeling

Fitts' Law, Model-Human Processor and GOMS

Many quantitative models arise from the Human Factors literature. Some models are best suited for describing expert (decision-free), simple motor and cognitive activities. The most well-known examples are Fitts' Law and GOMS. There have been numerous examinations of Fitts' Law in the context of graphical user interface design. Bill Buxton has published several papers on applications and extensions of Fitts' Law. A good example is:

- I. Scott MacKenzie, William Buxton. (1992) Extending Fitts' Law to Two-Dimensional Tasks. *Proceedings of ACM CHI'92 Conference on Human Factors in Computing Systems* pp. 219-226.

GOMS is based on a well-known model of human cognition and behavior, the Model-Human Processor. The following paper describes this model as well as the Keystroke Level Model, first defined by Card, Moran and Newell:

- Card, S.K., Moran, T.P and Newell, A. *The Psychology of Human-Computer Interaction*, Lawrence Erlbaum, 1983.

This work is the foundation for the GOMS family of evaluation techniques. GOMS has been one of the few widely known theoretical concepts in human-computer interaction. Two recent and good survey articles on the history and applications of GOMS are:

- Bonnie E. John and David E. Kieras. (1996) Using GOMS for User Interface Design and Evaluation: Which Technique? *ACM Transactions on Computer-Human Interaction*, v.3 n.4 p.287-319.
- Bonnie E. John and David E. Kieras. (1996) The GOMS Family of User Interface Analysis Techniques: Comparison and Contrast. *Transactions on Computer-Human Interaction* v.3 n.4 p.320-351.

Other Theories of Human Cognition

Other theories of human cognition examine the relationship between information in the head, such as a plan, and information in the world, such as a written to-do list. These theories may be the basis for both qualitative and quantitative models and evaluation techniques. Three contrasting theories are: situated activity, activity theory, and distributed cognition:

- Suchman, L. A. (1987). *Plans and Situated Actions: The problem of human-machine communication*. Cambridge: Cambridge University Press. (Also see <http://jupiter.informatik.umu.se/~mjson/hcipd/suchman.html>) Also available as (2006) ***Human-Machine Reconfigurations: Plans and Situated Actions 2nd edition***.
- Nardi, B. (1996) Activity theory and HCI & Studying Context. In Bonnie Nardi (Ed). *Context and Consciousness: Activity theory and human computer interaction*. Cambridge: MIT press.
- Hutchins, E. (1995). *Cognition in the Wild*. MIT Press. (Also see <http://www.cogs.susx.ac.uk/users/yvonner/dcog.html>)
- Halverson, C. (2002), *Activity Theory and Distributed Cognition or What Does CSCW Need to Do With Theories?*, Journal of CSCW.

Interaction Models

Some useful models explicitly place the user interacting with a computer interface. These models can be used to compare different interface designs. See:

- Hutchins, Hollan, and Norman (1986) Direct Manipulation Interfaces, in Donald Norman and Stephen Draper, *User Centered System Design*, 1986, pp. 87-124.
- Michel Beaudouin-Lafon (2000) Instrumental interaction: an interaction model for designing post-WIMP user interfaces, *Proceedings of CHI'2000*, pages 446-453.

Contextual Inquiry and Design

Contextual Inquiry is a set of methods for gathering qualitative information and human activity in a complex, social setting. A variety of models are used to represent these multivariate environments. Contextual Design is a methodology for using these models to inform an interface design.

- Beyer, H & Holtzblatt, K. (1998) *Contextual design: Defining customer-centered systems*. San Francisco: Morgan Kaufmann.

Gathering Qualitative Data

A common method for gathering qualitative data is interviewing. This short book is an indispensable guide:

- Interviewing as Qualitative Research, by I.E. Seidman

Two texts that cover the collection and analysis of Qualitative data are:

- Analysing Social Settings by Lofland and Lofland
- Strauss A, Corbin J. Basics of Qualitative Research: Grounded Theory Procedures and Techniques now in 3rd edition.

Many researchers now promote using ethnographic techniques to gather data about complex, social settings. As an example, see:

- Hughes, Sommerville, Bentley & Randall. (1993) Designing with ethnography: Making work visible. Interacting with computers. Vol 5:2. Pp. 239-253.

Evaluation

Quantitative vs. Qualitative

The most basic distinction is between a quantitative or qualitative evaluation. In a quantitative evaluation, the purpose is to come up with some objective metric of human performance that can be used to compare interaction phenomena. This can be contrasted with a qualitative evaluation, in which the purpose is to derive deeper understanding of the human interaction experience. A typical example of a quantitative evaluation is the empirical user study, a controlled experiment in which some hypothesis about interaction is tested through direct measurement. A typical example of a qualitative evaluation is an open-ended interview with relevant users. Some resources:

- See [Saul Greenberg's](#) discussion of techniques.

Evaluation Techniques

There are a number of established evaluation techniques that are useful in different situations. When reading about these techniques, focus on understanding when a technique is valid, and the underlying model of human behavior.

Cognitive Walkthrough

The cognitive walkthrough technique is another example of a theory-based evaluation technique.

- Peter Polson, Clayton Lewis, John Rieman, Cathleen Wharton, (1992) Cognitive Walkthroughs: A Method for Theory-Based Evaluation of User Interfaces. *International Journal of Man-Machine Studies* v.36 n.5 p.741-773.

Discount usability

In contrast to the theory-based techniques, there are a range of evaluation techniques that are more practically based, relying less on a foundational theory of human performance or cognition. Two good examples of this class of evaluation techniques are questionnaires and heuristic evaluation.

Summative vs. Formative

An important question to ask when performing evaluation is when to perform the evaluation with respect to the overall life cycle of a system. Formative evaluation occurs prior to much investment in implementation of a design, whereas summative evaluation occurs after a full system has been deployed. Many evaluation techniques can be employed in either a formative or summative mode, but it is important to know what the difference is when applied before or after an artifact has been implemented. You must also take into account the co-evolutionary influence of human tasks and interaction technology.

What is enough evaluation?

It is also important to understand that within the HCI research community, there are different expectations for evaluation. We should not expect the same amount of evaluation efforts in a paper that talks about a toolkit supporting multimodal gesture recognition as we would in a paper concerned with the impact of some existing technology in a domestic environment. When reading a research paper in the HCI area, you need to determine what the appropriate expectations should be for user-centered evaluation and judge accordingly. Remember that all systems have users (a programmer uses a toolkit) and proper consideration for the needs of that user should always be apparent in HCI research.

Georgia Tech Coronavirus Resources

- [Academic Continuity Resources](#)
- [Guidance for Graduate Education](#)
- [Hosting a Virtual Event](#)