
Evaluating Experience-focused HCI

Joseph 'Jofish' Kaye
Cornell University
Information Science
301 College Ave.
Ithaca, NY 14850 USA
jofish at cornell dot edu

Abstract

There is growing interest in experience-focused, rather than task-focused, HCI. Task-focused HCI has developed methods for creating and validating knowledge, but those methods may not be applicable or sufficient for experience-focused technology. In particular, new evaluation techniques to validate knowledge need to be created, discussed, and understood. I address this in three ways. First, it is important to understand the historical, technical and social factors that impact the evaluation criteria the community consider valid today. Second, I propose an ethnomethodological approach to evaluation that emphasizes the ways users use and make sense of technologies. And third, I demonstrate the validity of my approaches by means of several case studies.

Keywords

Experience, evaluation.

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Introduction

Human-computer interaction, like all other research disciplines, has certain ways it creates knowledge and certain ways it validates that knowledge. In HCI, the process of showing what we have done is valuable –

the process of validating the knowledge we create – is called evaluation, and a key concept behind much of the practice of HCI is the iterative design cycle: design, build, evaluate, and repeat. As HCI began to grow as a distinct field in the eighties, it developed a variety of techniques for evaluation: Nielsen's heuristics, GOMS, cognitive walkthrough, and so on. [13] These usability evaluation methods are excellent ways of evaluating whether a particular system, device, or approach is useful for accomplishing a given task.

However, a growing trend in contemporary HCI, involves the use of computer technology for more experiential and less task-focused needs: the use of technology outside the workplace and in everyday life, in an ongoing and ad-hoc manner. HCI continues to seek to design, build and evaluate technologies for this emerging focus of use, but in so doing must take into account other aspects of users' experience of the process. It is not sufficient to evaluate, say, a blogging site like LiveJournal on just the process of posting a blog entry. While the usability of the software remains important, it is also necessary to evaluate if the end result looks cutting-edge or professional or trendy, if it fulfills the users' needs for impression management, for privacy, for moving through different life stages. These experiential qualities may be hard to measure, but are a fundamental part of what the experience of the technology means to the user. This is why it is necessary to look at the evaluation of experience-focused HCI.

Background

This focus on experience in HCI is one that has been growing over time: McCarthy & Wright have suggested foundations based in pragmatist philosophy [16], and

Dourish has proposed a phenomenological approach [3]. The design of user experience (generally although not necessarily of websites) has received particular attention, such as Wiberg [18] and Hassenzahl, et. al. [10] – and, of course, the conference of the same name (DUX). But the question of how we as a discipline can recognize and validate work in experience-focused HCI is one that remains open.

Not a great deal of work specifically on the evaluation of experience-focused HCI has been published. Gaver and colleagues have shown the utility of cultural commentators from outside disciplines – filmmakers, journalists, ethnographers – for evaluating experience-focused systems [6][7], while Sundstrom et. al. found that having users' friends film them using an evocative technology was a useful way to characterize the experience of that technology [17]. Boehner et. al. have emphasized the role of 'dynamic feedback', where any information about user or system performance is not only collected for the evaluators' analysis but for the users' reflection, analysis and use as well. [2] Isbister et. al. have shown the potential for evocatively shaped ceramic objects to allow users to express emotions during interaction with a technology without reference to a standardized set of emotional categories [12]. Höök et. al. organized a workshop on evaluating affective technologies in January 2005, with a number of relevant approaches to experience-focused HCI. My own work on technologies for couples in long distance relationships to communicate intimacy involved the use of cultural probe-like diaries to gather thick descriptions of users' experiences with a technology [14][15].

However, the result of many of these approaches to evaluation does not look like the kind of knowledge that

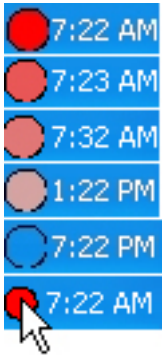


Figure 1. Clippings of the VIO in use, demonstrating the color change over time.

much of HCI is familiar with. N – the number of subjects – may seem laughably small to those with a background in experimental psychology, and a documentary film or a collection of postcards may not look like an evaluation to those who equate scientific approaches with rigor. And it's not clear how an evaluation method developed for evaluating a computer game can be applied to the evaluation of a mobile phone or a multiplayer mixed-reality game. How, then, can HCI incorporate these new ways of generating and validating knowledge while maintaining a consistently high standard of research?

Research

My Ph.D research addresses the question that underlies these problems. In many ways, the problem is one of epistemology: how can we know what we know? How can we know that the knowledge generated by these approaches is valid? What does it mean to have a good experience-focused HCI evaluation method? What are the standards by which we can evaluate a work of experience-focused HCI?

The first component of this work involves understanding the history of HCI with respect to evaluation, and the corresponding social and technical factors that have resulted in evaluation methods, methodologies, and criteria that the HCI community consider valid today. Needless to say, it is not the first time that HCI has addressed questions of the validity of the knowledge it creates. For example, the *Damaged Merchandise* controversy can be seen as a clash between those who believed in knowledge validated by the experimental method and those who believed in expertise and ecological validity [9]. Dourish's controversial paper at CHI 2006 emphasized the

problem of applying standards of validity that might be appropriate for, say, software engineering, to ethnographic reports [4]. Such work emphasizes the very real ways in which epistemological issues impact our everyday work in HCI.

The second component of this work involves understanding a theoretical approach to this work to inform and inspire the development of a methodology and set of criteria for the evaluation of experience-focused HCI. I posit that evaluation may require a different set of approaches and understandings than those afforded by design (c.f. [3][16]). In particular, even in cases of participatory or user-centered design, the agency or power in the relationship is by definition primarily in the hands of the designer. Evaluating a device or system requires that the actions of primary importance are those performed by the users and other stakeholders. What matters is not the intent of the designer, except as a point of comparison: rather, the focus needs to be on the ways the users make use and sense of the technology. For this reason, my work has concentrated on ethnomethodological approaches grounded in the work of Garfinkel, which emphasize the methods that everyday users use to make sense of the world [5]. I intend to use this approach to generate both a coherent methodology and a set of concrete methods for the experience-focused evaluation of HCI.

The third component of this work, appropriately enough, involves the validation of this work by means of case studies. My aim is to show that application of the methodologies I develop successfully characterize users' experience of using a technology, and provide ways for evaluators to engage with the product of that characterization to decide what to do next. To date,

I've been involved in case studies of the virtual intimate object [14][15], the Home Health Horoscope [7], and, at Microsoft Research Cambridge, of the Whereabouts Clock, the Ambient Ink Display [8] and a preliminary study of users' experiences of simple haptic devices [1]. I expect to perform one or two further studies before finishing my Ph.D.

Potential Contribution

My potential contribution with this work is twofold. First, to provide a methodology for the evaluation of experience-focused HCI. And second, to contribute to HCI's ongoing discourse about the nature of our intellectual endeavor, by emphasizing the constructed nature of the ways in which we create and validate knowledge.

References

- [1] Brown, L., and Kaye, J. 'J'. Eggs-ploring the influence of material properties on haptic experience. *Work-in-Progress; under consideration for CHI'07*.
- [2] Boehner, K. DePaula, R., Dourish, P. & Sengers, P. (2005). Affect: from information to interaction. In Bertelsen, O. et al. (eds). *Critical Computing –Between sense and sensibility*, ACM, pp. 59-68.
- [3] Dourish, P. *Where the Action is: The Foundations of Embodied Experience*. MIT Press, Cambridge, MA 2001.
- [4] Dourish, P. Implications for Design. *Proc. CHI'06*. ACM Press, NY. 541-550
- [5] Garfinkel, H. *Studies in Ethnomethodology*. Prentice-Hall, NJ. 1967.
- [6] Gaver, W. In press. Cultural Commentators: Non-Native Interpretations as Resources for Polyphonic Assessment. *IJHCI*, to appear.
- [7] Gaver, W., P. Sengers, T. Kerridge, J. "J." Kaye, and J.s Bowers (2007). "Enhancing Ubiquitous Computing with User Interpretation: Field Testing the Home Health Horoscope." In *Proc. CHI 2007*, to appear.
- [8] Gaw, S., Kaye, J. 'J.', Wood, K., Smyth, G. Impression Management in Tablet PC Users: Learning from Negative Feedback. *Work-in-Progress; under consideration for CHI'07*.
- [9] Gray, W. D., & Salzman, M. C. 1998. Damaged Merchandise? A Review of Experiments that Compare Usability. *Human-Computer Interaction* 13(3) 231-261.
- [10] Hassenzahl, Beu, and Burmester. Engineering Joy. *IEEE Software*, January/February (2001).
- [11] Höök, K., Laaksolahti, J., Isbister, K. (2006) Proceedings of the *WP9 Workshop on Innovative Approaches for Evaluating Affective Systems*. <http://www.sics.se/interaction/wp9ws/programme.php>
- [12] Isbister, K., Höök, K., Sharp, M., and Laaksolahti, J. 2006. The sensual evaluation instrument: developing an affective evaluation tool. In *Proc. CHI '06*. ACM Press, New York, NY, 1163-1172.
- [13] Jacko, J. & Sears, A., Eds. *The HCI Handbook*, Chapters 56-58. Lawrence Erlbaum 2003.
- [14] Kaye, J.'J.', Levitt, M.K., Nevins, J., Golden, J., and Schmidt, V. Communicating Intimacy One Bit At A Time. *Ext. Abs. CHI '05* 1529-1532.
- [15] Kaye, J. 'J.' (2006) I Just Clicked To Say I Love You: Rich Evaluations of Minimal Communication. *alt.chi, Ext. Abs. of CHI'06*.
- [16] McCarthy, J. & Wright, P. (2004). *Technology as experience*. Cambridge, MA: MIT Press.
- [17] Sundström, P, Ståhl, A. and Höök, K. In Situ Informants Exploring an Emotional Mobile Messaging System in Their Everyday Practice, *under consideration*.
- [18] Wiberg, C. (2003). A Measure of Fun. Extending the scope of web usability. Ph. D. thesis. Department of Informatics. Umeå University.

Biographical Sketch

After a childhood living in London, Paris, Singapore, and Tokyo, I received a B.S. in Brain & Cognitive Science from MIT and spent most of a year as a management consultant for an Italian home appliances company. This was followed by an M.S. from the MIT Media Lab, during which I was awarded a McDonalds Fellowship. I worked primarily on smart kitchens and homes, which culminated in a thesis, *Symbolic Olfactory Display*, on the use of scent to display information. I co-founded one startup and worked for another before entering Cornell University to work with Professor Phoebe Sengers in Fall 2003. I am in the Information Science program; I also maintain close ties with Science & Technology studies department.

My first project was *Intimate Objects*, a series of prototype devices for couples in long distance relationships, which culminated in the *Virtual Intimate Object*; I received a research grant from the Center for the Study of Long Distance Relationships to partially support this work. I also worked on the Home Health Horoscopes project, studied ways in which academics keep, display, maintain and discard archival materials, and am active in organizing the Culturally Embedded Computing group under the guidance of Phoebe Sengers. I spent the summer of 2005 as an intern with Genevieve Bell in the Domestic Design & Technology Research group at Intel, and the last six months of 2006 as a Visiting Researcher at Microsoft Research Cambridge.

While at Cornell, I've been a TA and occasional lecturer for the Critical Computing class and for Introduction to Web Design & Programming. I co-chair the Cornell SIGCHI Chapter, now in its fourth year of the Invited

Lecture Series (formerly the Distinguished Lecturer Series), a unique lecture series that concentrates on featuring talks from around eight advanced graduate students, junior faculty and young researchers in HCI a year.

My current work is focused on issues around experience-focused HCI. I co-organized a workshop at CHI'06 titled "Sexual interactions: why we should talk about sex in HCI", about one aspect of experience-focused HCI; one result from this is currently in the process of becoming a special issue of *interactions*.

Expected benefits

Like any other attendee, I would look forward to engaging with my peers and the sense of community to develop from such a discussion. However, in particular, I feel that discussions with an interdisciplinary group who are engaged with current work in the field such as the CHI DC, would provide a particularly rich opportunity to understand how to engage better with the discipline. For example, experience-focused HCI and particularly experience-focused evaluation are traditionally at the fringes of HCI practice. I have had difficulties getting my work published on the historical background of evaluation in HCI, which forms a significant part of my work: reviewers question whether historical discourse is legitimately part of HCI. I posit that such an approach is an entirely valid part of a reflective practice. How can I make clear the relevance of these topics to those in the field who do not feel they are working on any of them? I feel that I am at a particularly appropriate stage to benefit from the DC: with an anticipated eighteen months to go on my Ph.D, I feel I can contribute my experience to the consortium, and yet my work is still at a point where it can benefit from discussion.

Summary of estimated expenses

Flight ITH-SJC ¹ :	\$614
Hotel: Sharing double room, CHI rates: \$119/night/2 x 5 nights = \$298	
Meals, 6 days:	\$200
Total:	\$1112

¹ While ITH-SFO is cheaper than ITH-SJC, the ground transportation costs preclude savings. Flights from BGH and ELM, an hour away, are no cheaper; flights from SYR, an hour away, are barely cheaper after ground transportation.