Staying Open to Interpretation: Engaging Multiple Meanings in Design and Evaluation

Phoebe Sengers
Cornell Information Science and
Science & Technology Studies
Ithaca, NY 14850 USA
sengers@cs.cornell.edu

Bill Gaver
Goldsmiths College
University of London
New Cross, London, SE14 6NW UK
w.gaver@goldsmiths.ac.uk

ABSTRACT
Human-Computer Interaction (HCI) often focuses on how designers can develop systems that convey a single, specific, clear interpretation of what they are for and how they should be used and experienced. New domains such as domestic and public environments, new influences from the arts and humanities, and new techniques in HCI itself are converging to suggest that multiple, potentially competing interpretations can fruitfully co-exist. In this paper, we lay out the contours of the new space opened by a focus on multiple interpretations, which may more fully address the complexity, dynamics and interplay of user, system, and designer interpretation. We document how design and evaluation strategies shift when we abandon the presumption that a specific, authoritative interpretation of the systems we build is necessary, possible or desirable.

Author Keywords
Interpretation, hermeneutics, design, evaluation, humanities, arts

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

INTRODUCTION
Interpretation is a central issue for HCI. If we understand interpretation as the process by which users, nonusers, and designers come to assign meaning to the structures and functions of computational systems, whether at the level of what a button press might do or at the level of their relevance for ongoing life, then it is difficult to conceive of interaction without interpretation. Problems and issues around interpretation continually arise in HCI. For example, mental models research aims to analyze and align designer's and users' interpretations of what a system does [e.g. 30, 36], while affective computing aims to enable computers to accurately interpret users' emotional states and users, likewise, to accurately identify emotions computers express [e.g. 33].

While interpretation has always been a key issue in HCI, it has not always been foregrounded as such. In part, this may be because, as disparate as the areas of HCI that handle interpretational issues are, they often share a unified underlying stance on interpretation that circumscribes both the *problem* of interpretation and its presumed *solution*. While different areas disagree on whose interpretation (e.g., the users' or the designers') should be privileged, there is general agreement that there should be a single, correct way to interpret a computer system (e.g. how it works or the emotion it should exhibit or engender), and that the goal of the system's designer should be to convey that interpretation accurately to its users. Interpretation is then understood to be *causing a problem* when users and designers disagree about the meaning to assign to a system's operations, functions, or the role it plays in users' lives. The *solution* to this problem is therefore to adjudicate the disagreement, identify which interpretation is correct, and to design and contextualize systems so that this correct interpretation is agreed upon by all parties.

There is no doubt that clearly conveying specific preferred interpretations is often appropriate and useful. Nevertheless, we will show that HCI can and should systematically recognize, design for, and evaluate with a more nuanced view of interpretation in which multiple, perhaps competing interpretations can co-exist. We will argue that it is not necessarily a problem when users and designers have divergent interpretations of a system. And even when it is a problem, the solution does not necessarily need to be to establish and promote a single correct interpretation.

There are three reasons why we should reconsider whether a single preferred interpretation is necessary or desirable. The first is recent shifts in context of use: the expansion of computing beyond the relatively circumscribed and controlled context of the workplace into most facets of everyday life suggests that the domain of HCI has become broader, more personal, more idiosyncratic, and therefore less accessible to, and appropriate for, designer control. The second is recent interest in HCI in drawing on the arts and humanities [e.g. 11,12,14,16,23,26,28,35,39] whose

DIS 2006, June 26–28, 2006, University Park, Pennsylvania, USA.

perspectives assume a rich field of interpretation going beyond HCI's traditional preoccupations with utility and usability. The third is results from the sociology of technology [25,34,31] which demonstrate that even when a single interpretation of a technology is eventually established, this happens in a messy, complex, and protracted process involving negotiation among a variety of social groups. These three shifts suggest that foregrounding interpretation as a central issue for HCI, and in particular recognizing and taking advantage of multiple simultaneous interpretations, is timely and may open up new and useful design spaces for HCI.

In this paper, we aim to show the difference it makes to take interpretation in this more complex sense as at the heart of HCI. We suggest that interpretation has already been a central concern of HCI, and has been addressed largely by developing HCI's role as an adjudicator of the correct interpretation of a system. We outline a number of alternatives to the "single authoritative interpretation" assumption in HCI and show how they can provide useful new starting points for designing and evaluating systems.

A SINGLE AUTHORITATIVE INTERPRETATION

One of the core insights driving HCI is that for computer systems, form does not follow functionality. It takes work to interpret the behavior of code, and it is therefore a designer's job to support that interpretation to make systems easy to understand. For example, mental models research [e.g. 30,36] aims to support users in acquiring particular mental models that help them usefully interpret system behavior. The central problem from this point of view is the potential divergence of user and designer interpretations, or models, of the system. The solution proposed by mental models is to explicate models that can drive appropriate behaviour, and to design properties of the interface to support acquisition of these models. Users and designers may work from different models – those writing the code may have a different understanding of what is happening "under the hood" – but the designer is nevertheless responsible for identifying and encouraging a preferred interpretation among users. In Norman's formulation, "The designer should want the User's Model to be compatible with the underlying Conceptual Model, the Design Model. And this can only happen through interacting with the System Image…. If one hopes for the user to understand a system, to use it properly, and to enjoy using it, then it is up to the designer to make the System Image explicit, intelligible, consistent." [30, p. 47]

The notion that a specific preferred interpretation should drive system design choices is not limited to mental models research, of course. Many areas in HCI work from the assumption that a system should be *designed* to support a single interpretation. Ambiguity is generally seen as something to be coped with or resolved, not supported (for alternatives see [16,1]). Similarly, the *evaluation* of a system should measure whether the preferred interpretation the designer had in mind is actually taken up by users. A system may "work" if it is interpreted in ways different from those the designer had intended, but the difference will generally count against the design.

What Interpretation?

Users' interpretation occurs at a variety of levels and in a variety of ways, however, and these vary in their appropriateness for designer control [cf. 26]. At the lowest levels of interpretation, users need to interpret a system's interface and actions to use it at all: from "Is this a button?" to "What does this button do?" to "How do I do this task?". At middle levels of interpretation, users need to unpack how that system might relate to their everyday lives – from an understanding of "What is this system intended to be used for?" to "What activities is it appropriate for?" to "What role can it play in *my* life?" At the highest levels of interpretation, users interpret the values embodied in and the social and cultural meaning of systems – "What does it mean about me, my social group, my society, my culture?"

Lower levels of interpretation tend to involve issues traditionally associated with 'usability', and at these levels it would appear reasonable to assume a single preferred interpretation that accurately accounts for how the system works. Higher levels in the interpretation chain, in contrast, involve personal decisions about values and meaning for a specific user that appear less amenable to, and appropriate for, designer control. But the story is more complex than this. Levels of interpretation are not independent and sequential – for example, users' high-level interpretations of what the system might mean for their relationships may affect their lower-level interpretations of what the system is for. Interpretation at all levels is strongly dependent on context and the resources that users' social and cultural situations provide for interpretation. In addition, there is an interaction between *user* interpretation of a system and *system* interpretation of a user. Users interpret not only the system's interface, but also the ways in which the system is representing themselves and other users. The story about interpretation that emerges is complex, suggesting that the assumption that user interpretation can and should be controlled by designers may need to be rethought.

Whose Interpretation?

One set of alternatives to the 'classic' model of interpretation in HCI is provided by approaches such as user-centered design and ethnographically-inspired work that highlight differences between users' interpretations of systems and those intended by their designers. Suchman's *Plans and Situated Actions*, for example, identifies mismatches that occur in real-world situations between users' interpretational strategies and the assumptions designers hold about how users will approach systems [41].

The classic model of interpretation and user-centered approaches share an understanding of the *problem* of interpretation: they both see interpretation as a problem if users and designers disagree about the interpretation of a

system. They differ, however, on the *solution*. Approaches such as mental models generally focus on solving this mismatch by having designers alter their system's presentation so that users will come to see it as originally intended. User-centered design suggests that rather than changing the *user's* interpretation, we ought to change the *designer's*, i.e. he or she should redesign the system to better support users' preferred interpretations.

More generally, a set of approaches have arisen in HCI that is still based on the notion that there is a single – or at least deservedly dominant – interpretation for a system, *but suggests that this may not be the one explicitly advocated by its designer*. Approaches such as ethnomethodology or participatory design suggest alternative interpretations of "what's really going on" or what *should* really be going on around a system, other than the designer's intentions for it.

These new interpretations suggest in turn the need for new design and evaluation methods. For example, participatory design draws on Marxism to argue that designers are unwittingly in collusion with management against low-level workers. Workers on the shop floor are considered to have a more correct interpretation of what is really going on, and must therefore be drawn into the design process as equal partners [e.g. 4,38] Ludic design draws on critiques of utility **t**o argue that designers unconsciously design systems for work-related values such as efficiency — even when those systems are intended for home or leisure activities. Alternative values, such as curiosity, play, exploration, and reflection are also important from this point of view, and new design strategies and methods are needed to design for them [e.g. 20, 10; similar arguments motivate e.g. 3].

MULTIPLE, HETEROGENEOUS INTERPRETATIONS

While approaches that question designers' interpretations sound quite different from approaches that focus on interpretation as the responsibility of the designer, they share a common assumption: that there is a single, preferred interpretation which it is the job of HCI to adjudicate and support. In one case, that interpretation is the designer's; in others, it may be users or a third-party analyst whose interpretation should be primarily supported in the system design and evaluation process. But the ensuing proliferation of possible interpretations suggests that it may be not only possible but useful to address multiple interpretations simultaneously. No single one of these perspectives may necessarily be "correct;" instead, all may be useful in highlighting aspects of how systems will be understood, be used, and find roles in individual's and community's lives.

Science & Technology Studies, for example, has documented the many ways that technologies are *interpretively flexible*, i.e. lend themselves to many different interpretations besides those intended by their makers [2, 34]. People appropriate and reinterpret systems to produce their own uses and meanings, and these are frequently incompatible with design expectations and inconsistent within and across groups. Sociologists and historians have demonstrated that technologies' meanings become stable only through a protracted process of negotiation, interaction, and, sometimes, outright battle between groups of users, designers, manufacturers, policy makers and other mediators [e.g. 25,31].

While these historical studies show that many technologies are interpreted in multiple ways as they develop, allowing multiple interpretations to co-exist offers advantages beyond the initial "settling-in" stages of technological development. For example, the ability for different stakeholders to hold different interpretations of the same system can provide a kind of conceptual lubricant, allowing different perspectives and motivations to be applied to the same technologies without conflict. For instance, SMS messaging may be seen as a tool for coordination by businessmen, a social glue by teenagers, and a method for passing wireless control signals by hackers. These involve very different conceptions of the nature and purpose of SMS by different stakeholders that simply don't need to be agreed by consensus—indeed, specializing SMS around agreed purposes would risk reducing the possibilities it offers to niche users.

Systems that can be interpreted in multiple ways allow individual users to define their own meanings for them, rather than merely accepting those imposed by designers. This may be particularly important for systems intended for use in domains more open than the workplace, where peoples' relative freedom to choose their own experiences with and through technology may be undermined by technologies that convey strong narratives about their preferred uses. Systems that are open to interpretation don't need to be tailored to fit every possible niche audience; instead, the same system may support many ways of experiencing and acting in the world.

Surprisingly, systems that explicitly allow multiple interpretations may be safer than those implying a single preferred one, since they highlight users' responsibilities in interaction. For instance, some traffic engineers advocate removing road-signs, traffic signals, and the like, to force people to attend to and judge traffic conditions for themselves [27]. More generally, if people are enabled to play a substantial role in determining the meaning of systems, this implies that they will be actively engaged in the process of understanding both the system and its situation of use. An active engagement in sense-making may not only be pleasurable or liberating, it may also be useful in safety-critical applications.

Given the already-existing proliferation of meanings around technologies, as well as the potential advantages of designing for multiple meanings, the challenge for HCI becomes, not to decide upon and support a specific, correct interpretation of a system, but to incorporate and balance multiple, perhaps conflicting interpretations and *processes* of interpretation in design and evaluation. One way to support re-interpretations is through the use of skins or end-

user programming, in which the look or functionality of a system, and indirectly its eventual meaning, are left explicitly for users to tailor. But a much greater range of approaches is possible. Approaches to design are emerging in HCI that, without asking users to alter software's look or functionality, actively support multiple possible interpretations of a system's functionality, interface, and relationship to user life. In the rest of this paper, we present a taxonomy of practical design and evaluation strategies that leverage complex, heterogeneous interpretation.

DESIGNING FOR MULTIPLE INTERPRETATION

Our goal in this section is to lay out the contours of the emerging design space addressing multiple interpretations. The strategies we describe have begun to emerge piecemeal from our own and others' practice-based research. We here identify common patterns and principles that underlie these emerging strategies, suggesting directions for the development of new design possibilities.

If we take supporting multiple interpretations as a central goal, *design shifts from deciding on and communicating an interpretation to supporting and intervening in the processes of designer, system, user, and community meaning-making*. There are several ways to do so:

1. Designs can clearly specify *usability*, while leaving interpretation of *use* open.
2. Designs can support a space of interpretation around a given topic.
3. Designs can stimulate new interpretations by purposefully blocking expected ones.
4. Designs can gradually unfold new opportunities for interpretation over the course of interaction.
5. Designs can make space for user re-interpretation by downplaying the system's authority.
6. Designs can thwart any consistent interpretation.

Here, we will describe each of these strategies in turn.

Clearly specifying usability without constraining use

One way to support many different interpretations is to design systems as a blank canvas which can be interpreted by users in many possible ways. Of course, any system can be re-interpreted in unexpected ways, and this may be particularly true of unusable systems that confuse users. The goal of this design strategy, in contrast, is to have clear *usability* – what the system does and how it can be controlled is obvious – but the ultimate purpose, meaning, and usefulness of the device is left open for users to decide.

One example of such a system is the Key Table, developed as part of the Equator IRC (Figure 1). The Key Table was designed to support a simple interaction: load sensors supporting the tabletop measure the force with which things are placed on it, and a wirelessly linked picture frame swings out of kilter proportionally to this force. While the behavior of the Key Table is clear – force equals angle of picture frame - how users should make sense of this

Figure 1: The Key Table

behavior is not. To see how people would make sense of the Key Table, a volunteer household was recruited to live with it in their home for a month. The target family was mistakenly not told about the designers' intended meaning of the device. Instead, the research team captured the interpretations that emerged. To their surprise, the family's imaginations were captured by the portrait of a dog that had been placed in the Key Table's picture frame. Based on this picture, the users anthropomorphized the table as an animated presence in their home, seeing it as having moods of its own. This ended with them rechristening the table as 'Terrence the Table', playing games with it "just as we do our cats," and dressing the table in unusual materials.

Systems that are so open to interpretation shift the focus in technology design from instantiating a particular vision to exploring the ways in which users take up an artifact – as Gay and Hembrooke suggest, technology design becomes simultaneously social science research [21]. Although the focus shifts from design to situated use, the result can still be useful understandings of how future technologies should be designed. MacKay and her colleagues on the InterLiving project, for example, tested the uptake of "Technology Probes," small-scale communication technologies such as an electronic post-it system. The technology probes revealed how families take up, use, and misuse these technologies, inspiring further design [24].

Supporting a space of interpretations around a topic

In the previous example, the functionality of a system is specified precisely, while users are left open to decide how the system should relate to their lives. It is also possible for a system's design to open a space for interpretation in a more targeted way, by suggesting a subject or topic that the system is intended to be about, while not specifying how users should relate to that topic. For example, the electronic History Tablecloth, developed by the Equator IRC, encourages reflection on the flow of objects through the home – what moves around, what stays in place – by

Figure 2: The History Tablecloth

highlighting how long objects have stood on it [19]. If an object remains in place, over time the tablecloth slowly begins to glow underneath it (Figure 2). The Tablecloth suggests that whether objects are stationary or moving may be interesting to think about. It does not, however, suggest what the implications of this might be: whether it is good for objects to remain in place or move, or what the flow of objects in a home say about its inhabitants' lifestyle or values. These implications are left to its users to decide.

By leaving the more personal aspects of interpretation – what does this mean for me, as a person? – explicitly open, systems such as the History Tablecloth may support richer, more personally meaningful, and even more correct interpretations than they might explicitly be able to model or present. Some systems in affective communication directly leverage open interpretation to allow people to express personal, perhaps idiosyncratic emotions without being limited by categories internal to the system [6]. Höök, Ståhl, and Sundström's eMoto [42], for example, allows users to select affectively evocative background animations for SMS messages by shaking and squeezing a special pen. Neither the gestures nor the backgrounds are explicitly labeled with their emotional meaning; instead, users manipulate gestures to get backgrounds that correspond to their own sense of the emotional tone of the message. eMoto provides resources for communicating emotion without embodying a preferred interpretation of what that emotion might be.

Stimulating reinterpretations by blocking expected ones

One way in which designs can suggest new interpretations is by explicitly blocking interpretations that may be obvious or expectable. One example of such a system is the Drift Table (also developed by the Equator IRC) (Figure 3), a coffee table with built-in porthole that allows people to slowly 'drift' over the English countryside [10]. The Drift Table is intended to open new design spaces for technologies as supporting exploration, curiosity, and contemplation, rather than tasks. The Drift Table, in

Figure 3: Side (left) and top (right) views of the Drift Table

opening new design spaces, is faced with the challenge that users are not likely to come to the system ready to understand it. With their background cultural understanding of technology, users are almost guaranteed to initially interpret the Drift Table as a gadget to be used to accomplish a task. The goal for the Drift Table is not to communicate a single *correct* interpretation but to *avoid* communicating an *incorrect one.* The Drift Table is intended to suggest that technology does not need to be simply task-oriented, in ways users may expect. What it then is good for and how it can be taken up in a person's everyday life is left up to users to decide.

In order to block the obvious interpretation of being for a task, the Drift Table was designed explicitly to *not* support task-oriented use as a travel device. For example, there is no way to type in coordinates to go to a particular point. The only way users can move across the landscape is to place objects on the surface of the coffee table; the table 'drifts' in the direction of their weight, at a purposefully low rate. The view can be reset (typically to the table's own physical location), but this requires pressing a small and unobtrusive button, designed to emphasize that resetting the view is not a primary feature of the device.

In observing people's long-term interactions with the Drift Table over several weeks in their homes, the project team saw that users went through several stages of interpretation. Initially, users tended to see the Drift Table as a trendy gadget or design object. After initial use, they often became annoyed by what they perceived as limitations of the design to support tasks they wanted to engage in with it. While for some users, the failure of the device to behave as a task-oriented technology led them to abandon interest, others gradually came to meet the device on its own terms, so much so that previously suggested 'improvements' to support task-oriented behavior were now considered inappropriate. As one user eloquently described it: "Initially, I thought fantastic, another hi-tech toy in town. Then I became annoyed after the first day by the porthole. I couldn't show it to people as it is too small…. But that's worn off now. I thought about having a switch for double speed. Now that's worn off too. You should take a look around on the way like on a train journey. One should accept it and use it as it is. Another thing I thought was that it would be great to have a keypad so as to type in a coordinate. Then I thought no, it's for drifting around. I like

it for what it does. It's extremely sophisticated but without the arsing about. It has one use. It drifts. I like that understatedness about it. After a couple of days I was about to get bored with it because of its weaknesses but those are strengths. From shiny new object, to where's the buttons, to this is what it does."

Unfolding new opportunities for interpretation

Our experiences with the Drift Table suggest the power of strategies that explicitly *channel users' interpretive processes*, rather than supporting one or more static interpretations. One design strategy in this vein is Penny's notion of "autopedagogic interfaces" as a way to bootstrap user understanding of interactive artworks that may react to their behavior in complex ways [32]. Because, as Penny points out, no one wants to read a manual before they engage with an artwork, interactive artworks are constrained in the novelty of their interaction: either they need to follow a well-known interaction paradigm, or they risk confusing users with complex, often apparently random behavior. Penny suggests gradually ramping up the complexity of the system's reactions over time, leading the user through a series of interpretations of the systems' behavior which gradually increase in their richness.

In Penny et al.'s VR artwork Traces, for example, users' physical movement through the CAVE leave behind 3-dimensional traces that gradually 'come alive' and interact with the users' movements. The behavior of the traces follows three stages (Figure 4<): the first, passive, trace simply follows the user's movements directly; the second, active, trace uses cellular automata algorithms to sparkle like flames behind the user; the third, behaving trace is flicked from the user's body like particles of mud, then flocks and moves towards and away from the user's body. By building up interaction complexity in stages, users can gradually develop more complex interpretations: from "the system follows my movements" to "the system responds to my movements" to "the system interacts with my movements." While Traces does not explicitly support multiple interpretations at any particular time, it does channel a user's process of interpretation so that multiple interpretations unfold over the temporal course of interaction.

Making space by downplaying system authority

In our culture, technology often carries connotations of precision, correctness, and authority which can make users feel that the system's apparent interpretation (e.g., the data it collects and presents) must be more correct than users' own understandings. For this reason, in building systems that allow for re-interpretation, it is not enough for the system to suggest a variety of interpretations. It is also essential for any such system that users feel they have a *license* to reinterpret the system's behavior and its relationship to them, or users may feel simply frustrated or confused [23,8,39]. One way to support such a license is to make clear the limits of the system's own interpretation. In this section, we describe two design strategies that explicitly limit the authority of a system's interpretation to open a space for user re-interpretation.

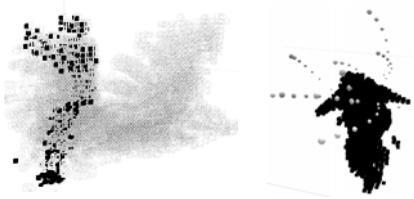

Figure 4: Traces' passive (left) and behaving (right) traces; user in black for reference

Seamfulness is a design strategy for ubiquitous computing developed from Weiser's initial vision by researchers at Glasgow and SICS [10,9]. In contrast to visions of ubiquitous computing as seamlessly hiding underlying technical glitches, uncertainties, and breaks, seamful designs explicitly represent the limitations and uncertainties in data, allowing users to make up their own minds about how to interpret it. As Chalmers et al. argue, "*Seamful designs go beyond mere accommodation of seams; they let users find ways to take advantage of seams and appropriate them for their own ends*" ([9], p. 7). GPS data in mobile systems, for example, is often inaccurate. Several researchers in mobile games have found that representing the inaccuracies of GPS data directly to users allows them to adapt their game playing strategies so that taking advantage of the seams becomes part of the fun [e.g. 14,10,9]. In the process, such devices require users to actively re-interpret potentially noisy, inaccurate, and conflicting data.

While seamful design undermines the vision of ubicomp as "all-knowing" by directly presenting inaccuracies in data for interpretation by the user, it is less clear how such a strategy would apply to ubiquitous devices that not only sense and report data but also make powerful inferences from and react to that data in ways too complex to directly represent. A design strategy to support a license to reinterpret such systems is *alien presence*, developed by Mateas as an alternative approach to ambient intelligence as invisible, all-knowing devices [37]. Alien presence, as a form of ambient intelligence, uses Artificial Intelligence techniques to actively interpret patterns of human activity and generate responses as a function of these interpretations. Unlike other ambient intelligences, alien presences are actively designed to make clear that their interpretation is only one, idiosyncratic interpretation of shared context. By presenting a system as 'alien' rather than 'intelligent', alien presence signals to users that the system has one interpretation, but not necessarily the only or right one.

One example of alien presence is Böhlen and Mateas's Office Plant #1 [5] (Figure 5), a robotic sculpture in the

Figure 5: Mateas and Böhlen's Office Plant #1

form of a high-tech plant that responds to the emotional and social tenor of its owners' incoming email stream. The device filters the users' email into social and emotional categories such as "chatty" or "FYI," with the resulting categories driving subtle and very slow changes in the plant's shape: the petals move in and out, the fronds wave, etc. Office Plant #1 clearly has an interpretation of the user's email, but this interpretation is not a straightforward status monitor of the presence or emotional quality of that email. It is presented indirectly and in an alien form that requires interpretation. By presenting their interpretations in defamiliarized ways, such systems provide opportunities for reflection on and reinterpretation of the context which users and systems share, but see in different ways.

Thwarting any consistent interpretation

The previous 5 design strategies present several different ways to open up and balance the processes of user, designer, and system interpretation. But, as Gaver, Beaver, and Benford [16] argue, it is also possible and sometimes even desirable to design in a way that does not easily support *any* single interpretation. This is a common strategy in the art world; many art pieces amalgamate incomplete references to other works and genres without allowing the result to "fit" known categories completely. This ambiguity opens up a space for interpretation that cannot easily be settled or resolved to a single interpretation.

Gaver, Beaver, & Benford argue that ambiguity can be a desirable property not only for artwork but also in interface design, for example to express uncertainty in a systems' precision, to support users in rethinking the roles systems play in their lives, or to raise questions about people's moral relationships towards the ways of living suggested by technologies [16]. Drawing on this argument, Aoki & Woodruff argue that communication technologies should be explicitly designed to support multiple potential interpretations of users' behavior. Technologies that are unclear about why a user is unresponsive allow users to save face in social situations where one user is more interested in social contact than another [1; see also 7].

EVALUATION STRATEGIES

In the previous section, we catalogued new design strategies that emerge when explicitly designing for the possibility of heterogeneous interpretations. But explicitly recognizing the legitimacy of multiple interpretations leads to challenges and opportunities not only for *design* but also for *evaluation.* Common approaches to evaluation in HCI are based on developing and testing against a priori evaluation criteria corresponding to the designers' anticipated interpretation of a system. But in taking multiple interpretations into account, systems can no longer be effectively evaluated in terms of criteria generated from a single, authoritative interpretation.

If we focus on reinterpretation of systems instead as an evaluation metric, we can be in danger of declaring every system a success, since every system, perhaps especially unusable ones, can trigger new interpretations and be used in ways we do not intend. It is important, therefore, to highlight that *designing systems to support a rich range of interpretations does not abdicate the designer from responsibility for the eventual success of the system* [cf. 22]. Instead, designers might develop new kinds of evaluation criteria that focus on their design goals: not "did the preferred interpretation take hold with users?" but "How many different interpretations does a particular 'blank canvas' generate, and why?" or "Do users feel both stimulated and empowered to develop their own interpretation of an alien presence system?"

Evaluation is also a form of interpretation, however, and can itself be single or multiple. Thus an emphasis on multiple interpretations might suggest gathering together a rich mélange of interpretative accounts that might be inconsistent or contradictory, instead of or in addition to focusing on meta-level criteria for success. In this case, *evaluation shifts from determining whether an authoritative interpretation was successfully communicated to identifying, coordinating, stimulating, and analyzing processes of (evaluative) interpretation in practice.* While the results may conflict, the responsibility remains with the designer to weigh the results and to justify his or her eventual conclusions. Several strategies present themselves that leverage multiple interpretations for evaluation.

Incorporating user interpretation into evaluation

First, we can systematically incorporate users' own interpretations of their experiences into the evaluation process. Ethnographers, of course, have long argued in favor of using people's accounts of their own meaning to understand systems, and ethnography is useful in capturing rich and multi-layered accounts of people's experiences with new designs. Although this is most commonly used in HCI for design, it is equally valuable in evaluation (see [18] for an example).

Another approach to incorporate user interpretation into evaluation is Boehner and Gay's notion of dynamic feedback [21,8]. When using dynamic feedback, whatever information is collected about or from users is also given back to users to interpret. While dynamic feedback can be incorporated into system design [see e.g. 8, 39], it can also be used as part of evaluation. For example, in our on-going

evaluation of Affector [6], a system for communicating mood between friends' offices, we are collecting statistics of how many times per day its users interact face-to-face. Superficially, if patterns of face-to-face interaction change during the course of using Affector, it would suggest some type of effect, although it is difficult to say whether more or fewer interactions is an improvement. To get a richer understanding of what is happening with Affector, shifts in baseline measures are reported back to users. Their narratives of the reasons for changes potentially reveal richer understandings of Affector than our analysis of the numbers alone. Presenting users' information back to them not only gives additional insight for evaluation, it also can stimulate new reflection and interpretation among users and gives them a license to participate in the evaluation of a system as well as its interpretation (see also [10]).

As we described in the case study of the Drift Table, users' interpretations of systems can shift substantially over time. Longitudinal studies may therefore also be invaluable in understanding such changes in interpretation. For example, Friedman et al. studied the use of a virtual window over a period of 16 weeks [15; personal communication]. They found that users' interpretation of the display window, particularly the higher-level interpretations such as how they felt about it, incorporated it into their workplace practices and everyday routines, and their sense of values embodied in using the display, could change dramatically over the course of the study. Results like these suggest that short-term studies are only catching a single snapshot of the many interpretations users may develop across time in using the system. Of course, for low-level interpretation issues, such as "what does this button do?", problems and opportunities can probably be identified on much shorter time-scales. But for high-level interpretation issues, such as "what implications does this system have for how I want to lead my life?", it is likely that long-term studies will be necessary in order to formulate comprehensive accounts.

Multiple, potentially inconsistent assessments

Incorporating user interpretations and running long-term studies are understood techniques within HCI, compatible with the assumption that evaluation should compare actual understandings against preferred interpretations. Acknowledging the potential value of multiple interpretations, however, may lead to a more radical reformulation of what it means to assess the success of new systems. In this view, systems might best be evaluated by gathering and presenting a variety of assessments from a diverse population of interpreters, allowing outsiders to get a rich and layered view of how the system is used, the roles it plays, and the cultural implications it suggests.

After all, users are not the only potential source of interpretations that may inform the evaluation process. In evaluating systems explicitly designed to support multiple interpretations at several levels, we have found it useful to draw on notions of evaluation from the humanities and arts, which underscore the value of expert critics. Traditional designers, for example routinely use critiques, and these can involve expert commentators from a variety of areas including the humanities. We believe it might also be useful to involve commentators from outside traditional academia such as journalists and art critics, as well as those from fields traditionally thought to be unrelated to HCI such as psychoanalysis or forensic anthropology, as long as they can offer new perspectives on systems and the ways people interact with them. External evaluators have certainly been proposed in HCI as a way to provide a more objective viewpoint, i.e. to supply an interpretation which is likely more correct than that of the perhaps biased designer. Our goal is somewhat different; not to find a more correct interpretation of the system, but to play multiple interpretations off of each other, including those of the designers, the users, and of one or more external experts.

For example, in both the Drift Table and Key Table projects we hired a filmmaker to create a documentary about how users were taking up the devices we had made. In both cases, we told the filmmaker as little as possible about our intentions for the device before sending him to users' homes. We compared the results to other interpretations: our own, as well as that of the ethnographer who had studied the Drift Table. What we found was that the filmmaker added his own interpretations to that of the users. In the case of the Drift Table, this was relatively subtle, amounting to occasional side comments about expecting the Drift Table to be "some kind of techno-gadget" but finding that he wanted to own one. In the case of the Key Table, however, the filmmaker's interpretation was much more influential, apparently amplifying the family's inclination to view the table as a kind of virtual pet. In each case, however, the filmmaker's interpretation was not clearly differentiated from that of the users he filmed, but instead served as a sort of filter through which the users' interpretations were viewed.

Thus expert perspectives such as we used in the documentary approach overlay multiple interpretation as a resource for understanding how our systems are taken up, but they do not necessarily make such an understanding easier. Instead, they subvert any single perspective on evaluation and encourage audiences of our work to find their own. In fact, any form of evaluation in the end relies on our expert peers in the HCI community to judge its effectiveness and success – or lack thereof. This suggests opportunities for evaluation which are aimed, not at finding a final answer of what worked and didn't work, but at supplying data in a form which expert readers can interpret for themselves.

CONCLUSION

Our aim in this paper is to demonstrate the potential power for HCI of considering multiple, co-existing interpretations. Explicitly recognizing the legitimacy of multiple

interpretations has fundamental implications both for the *process* of design, and its *accountability* (evaluation). Design shifts from deciding on and communicating an interpretation to supporting and intervening in the processes of designer, system, user, and community meaning-making. Evaluation shifts from determining whether an authoritative interpretation was successfully communicated to identifying, coordinating, stimulating, and analyzing processes of interpretation in practice.

One concern readers may have about this approach is that recognizing multiple, conflicting interpretations as potentially legitimate will lead to a situation where the meaning of a system is undecidable or a matter of opinion. As we hope our design and evaluation strategies demonstrate, recognizing multiple, perhaps conflicting interpretations as legitimate does not have to lead to an anything-goes mentality. Designers must still develop approaches that address multiple interpretations in definable and testable ways. These approaches do not replace single-interpretation approaches; rather, they suggest new opportunities for both design and evaluation.

In particular, as McCarthy & Wright have demonstrated [26], recognizing the complex, heterogeneous nature of interpretation can be a powerful tool for building connections between HCI and the humanities. Humanist theories offer new lenses of interpretation useful for exploring what our systems might mean and how they can usefully support heterogeneous interpretations. For example, psychoanalysis suggests that the meaning of a text - or technology - may be rooted in subconscious issues. If we apply a psychoanalytic model to systems evaluation, we may seek to identify, not whether the designer's intended interpretation is taken up by users, but how systems reflect their designers' subconscious concerns and how these unintentionally shape user experience. Or, alternatively, we may look at how users' interpretations of systems are grounded in their own subconscious conflicts and what factors in system design lead users to project these meanings onto the system. Similarly, if we design from a psychoanalytic perspective, the goal of our design may be to create a kind of "Rorschach" system that maximally supports users in projecting their own personal meanings onto it. Evaluation of such a system may examine how wide the range of possible meanings of the system is, rather than whether the author's intended meaning won out [c.f. 10].

Our goal is not to suggest that designing or evaluating for a single, preferred interpretation is wrong; rather, that it is one among a range of possibilities that can be explored. The question of how many interpretations or what kinds of interpretation a specific system should support will, of course, depend on the application under consideration. It will also depend on which level of user interpretation is being addressed. For eample, 'blank canvases', in our formulation, have clear interpretation at low levels of interpretation, but open a space for interpretation at higher levels. Interpretation can also be open in different ways: designers can, for example, through their designs suggest different topics for interpretation. The goal is not to design systems that are completely open to interpretation. It is instead to allow the rhythms of constraint and openness in interpretation to become part of the design language available to us in HCI.

ACKNOWLEDGMENTS

Many thanks to Ken Anderson, Mads Bødker, Kirsten Boehner, Matthew Chalmers, Shay David, Joseph "Jofish" Kaye, Tom Jenkins, Thorsten Joachims, Kristina Höök and anonymous reviewers for comments on versions of this paper. The systems and experiences described in this paper are collaborations with Kirsten Boehner, Andy Boucher, John Bowers, Hans Gellerson, Rev Guron, Tom Jenkins, Andy Law, Sarah Pennington, Eunyong "Elie" Shin, Anthony Steed, Brendan Walker, Simeon Warner, and Nicholas Villar. This work is sponsored by NSF Awards IIS-0238132 and IIS-0534445, an Intel grant for Supporting Well-Being and Reflection in the Digital Home, and the Equator IRC (EPSRC GR/N15986/01).

REFERENCES

1. Aoki, P. & Woodruff, A.(2005). Making space for stories: Ambiguity in the design of personal communication systems. *Proc. CHI '05*. NY: ACM Press, 181-190.
2. Bijker, W. E. (1995). *Of bicycles, bakelites and bulbs: Toward a theory of sociotechnical change*. Cambridge, MA: MIT Press.
3. Blythe, M., Overbeeke, K., Monk, A., & Wright, P. (eds) (2003). *Funology: From usability to enjoyment.* Kluwer Academic Publishers.
4. Bødker, S., Ehn, P., Romberger, S., & Sjögren, D. (Eds.). (1985). *The UTOPIA project: An alternative in text and images (Graffiti 7).* Swedish Center for Working Life, the Royal Institute of Technology, Stockholm, & the University of Aarhus.
5. Böhlen, M. & Mateas, M. (1998). *Office Plant #1*: Intimate Space and Contemplative Entertainment. *Leonardo*. Vol **31**, no. 5, 345-348.
6. Boehner, K., DePaula, R., Dourish, P., & Sengers, P. (2005). Affect: From Information to Interaction. *Proc. Critical Computing '05*.
7. Boehner, K., & Hancock, J (2006). Advancing Ambiguity. *Proc. CHI 2006*, pp 103-106.
8. Boehner, K., Thom-Santelli, J., Zoss, A., Gay, G., Hall, J., & Barrett, T. (2005). Imprints of Place: Creative Expressions of the Museum Experience. *Extended Abstracts of CHI 2005*. NY: ACM Press, 1220-1223.
9. Chalmers, M., Dieberger, A., Höök, K., & Rudström, Å. (2004). Social Navigation and Seamful Design. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society.* 11(3), pp 171-181.

10. Chalmers, M., MacColl, I., & Bell, M. (2003). Seamful design: Showing the seams in wearable computing. *Proc. IEE Eurowearable 2003*, 11-17.
11. Dourish, P. (2001). *Where the Action is.* Cambridge Mass.: MIT Press.
12. Dourish, P., Finlay, J., Sengers, P., & Wright, P. 2004. Reflective HCI: Towards a critical technical practice. In *CHI '04 Extended Abstracts*. NY: ACM Press, 1727-1728.
13. Dunne, A. & Raby, F. (2001). *Design noir: The secret life of electronic objects.* Basel, Switzerland: Birkhäuser.
14. Flintham, M., Benford, S., Anastasi, R., Hemmings, T., Crabtree, A., Greenhalgh, C., Tandavanitj, N., Adams, M., & Row-Farr, J. (2003). Where on-line meets on the streets: experiences with mobile mixed reality games. *Proc. CHI '03.* NY: ACM Press, 569-576.
15. Friedman, B., Freier, N. G., and Kahn, P. H. (2004). Office window of the future?: two case studies of an augmented window. In *CHI '04 Extended Abstracts*. NY: ACM Press, 1559-1559.
16. Gaver, W., Beaver, J., & Benford, S. (2003). Ambiguity as a resource for design. *Proceedings of CHI'03*. NY: ACM Press, 233-240.
17. Gaver, W., Boucher, A., & Martin, H. (2003). Evaluating the Double-Deck Desk. *Proc. HCII '03*.
18. Gaver, W., Bowers, J., Boucher, A., Gellerson, H., Pennington, S., Schmidt, A., Steed, A., Villars, N., & Walker, B. (2004). The Drift Table: Designing for ludic engagement. *Proc. CHI'04 Design Expo*. New York: ACM Press, 885-900.
19. Gaver, W., Bowers, J., Boucher, A., Law, A. Pennington, S. and Villars, N. (submitted). The History Tablecloth: Illuminating Domestic Values. Submitted to *DIS 2006*.
20. Gaver, W.W., & Martin, H. (2000). Alternatives: exploring information appliances through conceptual design proposals. *Proc. CHI'00*, NY: ACM Press, 209-216.
21. Gay, G. & Hembrooke, H. (2004). *Activity-Centered Design*. Cambridge, MA: MIT Press.
22. Höök, K. 2004. Active co-construction of meaningful experiences: but what is the designer's role?. In *Proc. NordiCHI '04*, vol. 82. NY: ACM Press, 1-2.
23. Höök, K., Sengers, P. & Andersson, G. (2003). Sense and sensibility: Evaluation and interactive art. In *Proc. CHI '03*, 241-248.
24. Hutchinson, H., Mackay, W., Westerlund, B., Bederson, B. B., Druin, A., Plaisant, C., Beaudouin-Lafon, M., Conversy, S., Evans, H., Hansen, H., Roussel, N., & Eiderbäck, B. 2003. Technology probes. In *Proc. CHI '03.* NY: ACM Press, 17-24.
25. Kline, R. & Pinch, T. (1996). Users as agents of technological change. *Technology and culture*, Vol. 37, No. 4., 763-795.
26. McCarthy, J. & Wright, P. (2004). *Technology as experience*. Cambridge, MA: MIT Press.
27. McNichol, Tom. Roads gone wild. Wired, Vol 12, No. 12, December 2004. http://www.wired.com/wired/archive/12.12/traffic.html
28. Mitchell, W., Inouye, A. and Blumenthal, M., eds. *Beyond Productivity: Information Technology, Innovation, and Creativity*. National Research Council (2003).
29. Nickerson, R. & Landauer, T. (1997). Human-computer interaction: Background and issues. In M. Helander, T. Landauer, & P. Prabhu (Eds.), *The handbook of human-computer interaction* (pp. 3-31). Amsterdam: Elsevier Science.
30. Norman, D. A. (1986) . Cognitive Engineering. In Norman D. A. & Draper S. W. [Eds.]:*User-Centered System Design*. Hillsdale, NJ: LEA.
31. Oudshoorn, N. & Pinch, T., eds. (2003). *How users matter*. Cambridge, MA: MIT Press.
32. Penny, S., Smith, J., Sengers, P., Bernhardt, A., and Schulte, J. (2001). Traces: Embodied Immersive Interaction with Semi-Autonomous Avatars. *Convergence*. Vol. 7, No. 2.
33. Picard, Rosalind (1997). *Affective Computing*. Cambridge, MA: MIT Press.
34. Pinch, T. & Bijker, W. (1987). The social construction of facts and artifacts. In W. Bijker, T. Hughes, & T. Pinch, eds., *The social construction of technological systems.* Cambridge, MA: MIT Press, pp. 17-50.
35. Reeves, S., Benford, S., O'Malley, C., and Fraser, M. (2005). Designing the spectator experience. In *Proc. CHI '05.* NY: ACM Press, 741-750.
36. Rogers, Y., Rutherford, A., & Bibby, P. (Eds.) (1992). *Models In the Mind*. London: Academic Press.
37. Romero, M. & M. Mateas (2005). A preliminary investigation of Alien Presence. *Proc. HCII '05*.
38. Schuler, D. & Namioka, A. (eds). (1993) *Participatory Design*. Lawrence Erlbaum Ass.
39. Sengers, P., Boehner, K., David, S., & Kaye, J. (2005). Reflective Design. *Proc. Critical Computing*, 49-58.
40. Sengers, P. & Gaver, B. (2005). Design for Interpretation. *Proc. HCII '05.*
41. Suchman, L. (1987). *Plans and situated actions.* Cambridge: Cambridge UP.
42. Sundström, P., Ståhl, A. & Höök, K. (2005) eMoto - Affectively Involving both Body and Mind. *Extended Abstracts of CHI'05*, 2005-2008.