PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Designing for Emotion (Among Other Things)

Journal:	*Philosophical Transactions B*
Manuscript ID:	RSTB-2009-0153
Article Type:	Research
Date Submitted by the Author:	07-Jul-2009
Complete List of Authors:	Gaver, William; Goldsmiths, University of London, Design Department
Issue Code: Click here to find the code for your issue.:	COMPUTATION
Subject:	Behaviour < BIOLOGY, Cognition < BIOLOGY
Keywords:	design, emotion, interaction

Designing for Emotion (Among Other Things)

William Gaver

Interaction Research Studio
Goldsmiths, University of London

ABSTRACT

Using computational approaches to emotion in design appears problematic for a range of technical, cultural and aesthetic reasons. After introducing some of reasons I am sceptical of such approaches, I describe a prototype we built that tried to address some of these problems, using sensor-based inferencing to comment upon domestic 'well-being' in ways that encouraged users to take authority over the emotional judgements offered by the system. Unfortunately, over two iterations we concluded that the prototype we built was a failure. I discuss the possible reasons for this and conclude that many of the problems we found are relevant more generally for designs based on computational approaches to emotion. As an alternative, I advocate a broader view of interaction design in which open-ended designs serve as resources for individual appropriation, and suggest that emotional experiences become one of several outcomes of engaging with them.

QUESTIONING COMPUTATION OF EMOTION

Let me start by laying my cards on the table: I am not a fan of computational approaches to emotion, at least not when applied to interaction design in the form of systems that either try to sense and respond to user emotions, or simulate emotional expressions themselves. My reasons for this are multiple, reflecting doubts about the technical feasibility of such approaches, their personal, social and cultural implications, and the overall aesthetics of their approach to human computer interaction.

First, automatically sensing and representing people's emotions[i] appears extremely difficult. Systems based on facial recognition

may have had some success in the laboratory, but encounter severe difficulties in the field (e.g. Cowie, this volume). The problem is that our emotions are fleeting, mixed and masked. For example, consider a scenario in which you take a wrong turn while driving a friend home. You might experience a surge of frustration, perhaps even a small twinge of fear as you realise you're disoriented. Still, you have to laugh when your friend makes a joke at your expense. Feigning outrage, you put on a mask of confidence as you assure her that you know where you are going. All the time you are happy to be in her company, even if a bit tired after a long day. Let's say your in-car navigation system is trying to track your emotions to guide how it presents information. What should it do? Is it really going to be able to make sense of such a fluid amalgamation of emotional states, much less respond appropriately?

Even if systems can somehow disambiguate a mixture of basic emotions based on some combination of facial recognition and physiological monitoring, this is only a small part of comprehending the meaning of an emotional experience. For instance, the other day I was happy and excited to catch a large tench while fishing in the countryside. Later that same day, I played a chaotic mixture of hide and seek, catch, and tickling with my two young daughters. Again I was happy and excited. Even if we grant that my emotions in these two situations were somehow the same, it is clear that the meaning of these situations, the relationships I had with the fish and with my little girls, were quite different. This suggests that for systems to do a good job of reasoning about emotional experiences, they will have to understand a great deal about the context in which the emotions arise. If we believe, as some do, that emotions are the product of arousal and cognitive assessment (Mandler, 1984), then each emotion is specific to its circumstances, with the linguistic terms we use to identify them pointing merely to family resemblances. In other words, emotions are situated (Suchman, 1987). Similar reasoning based has led some researchers (e.g. Boehner et al., 2005) to consider accurately sensing emotion to be impossible in principle.

But let us assume that we can accurately sense emotions in a way that is either independent of context or includes it sufficiently. Do we want to? The social and cultural implications of emotional sensing are also troublesome.

To begin with, it is easy to imagine that automatic sensing of emotions would intrude on our privacy in significant ways. What if your closest friend bought a gadget that allowed them to track your affective state accurately over time. Would you want them to have access to your unmediated emotions? Would you be able to decline if they exerted emotional pressure? Maybe it would be fun, at least for a while. This is only the tip of the iceberg, however. Supermarkets might track your emotional responses as you navigate the aisles, using their existing security cameras to augment market research. Web companies might be interested in keystroke patterns and mouse pressure as a measure of your emotional reactions to various search results. They might sell this information on to your insurance company or bank. Clearing immigration at the airport would involve officers tracking your emotional responses as they ask leading questions about your political and religious beliefs, passing the results to the police if necessary. If automatic tracking of emotions is successful, the ways this information is used will be as difficult to trace, much less control, as any of the personal information that is currently tracked and trafficked on the web (Greenfield, 2006).

Even more insidious, perhaps, is the potential for emotional tracking to alter our relationships to our emotions and to each other. Consider systems proposed to support older people living at home ('aging in place'). Automatic sensing of emotion could augment systems that track whether older people have fallen, or failed to get out of bed, or to eat adequately. By tracking the emotions of older people in their homes, the argument goes, we could reassure ourselves of their continued wellbeing. The danger is that such systems might automate caring relationships traditionally maintained through personal social contact. Naturally we wouldn't consciously abandon the aged to automatic care. But imagine being in a hurry to get home, and wondering

whether to visit an older friend on the way. Wouldn't this be less likely if you had a device to reassure you not only that they were active and safe, but showing all the physiological and expressive signs of happiness as well?

The problem is that such tools are designed inherently to stand in for our own personal assessments of emotion. Charles Arthur (2009) tells a story that illustrates this point vividly. US police sometimes question subjects with a 'lie-detector', asking them to place their hands on the machine, which then produces sheets of paper reading 'true' and 'false' as they answer questions. The trick is that the machine is an office photocopier and both the questions and judgements have been previously scripted. Apparently many suspects are fooled by this, leading Arthur to comment: 'our tendency to believe what machines tell us – even if we don't understand them – still baffles me'. This tendency may lead us to trust devices that purport to sense emotions more than our own perceptions, whether they are seen as tools for monitoring emotion at a distance, or in people who are physically co-present, or even in ourselves.

Arguments about privacy and automation can be seen as luddite reactions to new technologies, of course. Similar reasoning might lead to the conclusion that thermometers are immoral because they dissuade us from attending to our own experience of temperature (Cowie, personal communication). From this perspective, if we can learn to manage our privacy and relationships over the web, then we can do the same in the face of emotional sensing. These sorts of debates about the cultural desirability of new technologies are difficult to decide, though we can all make choices based on their consideration.

In any case, my doubts about computational approaches to emotion as a basis for human computer interaction go beyond these broad technical and cultural concerns to include more personal beliefs about fruitful approaches to interaction design. The basic question is whether we want to model interfaces after agents or environments. Do we want to pretend that computers are sentient beings with whom we can converse and form relationships, or is it

more fruitful to think of them as virtual environments, furnished with a variety of tools, machines and objects, in which we can explore and pursue our aims? My view on this dates back to the early 1980's, when direct manipulation (a.k.a. GUI) superseded command-line interfaces. The new interfaces presented an environment with its own blend of 'physics and magic' (Smith, 1986) in which users have the illusion of acting directly on a computational world (Hutchins, Hollan & Norman, 1985). This paradigm shift seemed to make irrelevant both the rigid command languages of the time and the hype of natural language processing to come[ii]. Simultaneously, it opened a new design space in which designers could shape the affordances of virtual environments (Gaver, 1991) and, in the long run, of computational products. This seemed, and seems, a tremendously exciting prospect to me, while the notion of interfaces that promote the illusion of agency, encouraging people to project trust and affection upon them, appears not only difficult to achieve in likeable ways, but prone to produce systems that are inauthentic, patronising and manipulative.

No doubt, artificial agents have been, and will be, deployed quite widely in computer games and interactive narratives, sales kiosks and toys such as the Pleio and Sony Aibo,. But I would argue that such applications are enjoyed as 'clever machines', digital versions of mechanical automata that most people recognise as non-agents even while enjoying their quasi-intelligence. Moreover, I suspect such applications will be accepted and enjoyed as long as playing along with the simulated emotional engagement they offer doesn't present serious hazards. From this point of view, recognisably artificial agents, especially used in relatively inconsequential domains, may offer interesting possibilities for designers (as I explore in the next section). On the whole, however, I remain committed to an approach in which computational devices are seen as environments complementing and supporting our abilities rather than seeking to emulate them.

My final doubt about pursuing computational approaches hinges on the view that emotion, however pursued, is seldom if ever an

appropriate focus for design. Clearly emotion is a crucial facet of experience. But saying that it is a 'facet of experience' suggests both that it is only one part of a more complex whole (the experience) and that it pertains to something beyond itself (an experience of *something*). It is that *something* – a chair, the home, the challenges of growing older – which is an appropriate object for design, and emotion is only one of many concerns that must be considered in addressing it. From this point of view, designing for emotion is like designing for blue: it makes a modifier a noun. Imagine being told to design something blue. Blue what? Whale? Sky? Suede shoes? The request seems nonsensical. Similarly, focusing design on emotion without a grounded sense of the situation in which emotions are meant to gain meaning appears to be a category error. Instead, we need to understand how to design for engaging experiences more generally.

A SAMPLE OF THE DIFFICULITIES: THE HOME HEALTH MONITOR

My comments about the problems of using computational approaches to emotion in design are based not only on observation or reasoning, but on the bitter experience of developing and deploying two iterations of a system that, while serving both of as a gentle parody of and suggestion for improvement to traditional emotional computing approaches, still followed their essential logic.

The basic idea of the Home Health Monitor was to use sensors in the home to track household conditions that might be symptomatic of the overall emotional state, or mood, of its inhabitants. Data returned by these sensors are processed and used to build a representation of the household's 'wellbeing', defined broadly and relative to the particular household, and the outcome displayed to support reflection (see Figure 1). For instance, we might design a sensor device to measure when a given door is open or shut because the home's occupants have informed us that it is only closed when household members want to avoid each other. The raw sensor data is processed to uncover attributes such as the total time the door is open or closed during the day, how often it moves, or how early its state first changes. Rules compare the day's readings with trends found over the preceding days to

determine whether they are unusually high or low, and map this to an increment or decrement of, e.g., the 'sociality' metric accordingly. The pattern of metric scores provides a representation of the home's wellbeing that is mapped to an output for users. In the first iteration (Gaver, Sengers et al., 2007), the system constructed 'horoscopes' from sentences culled from online examples and categorised according to the wellbeing metrics; in the second iteration, we tried three different forms of output.

=== insert Figure 1 about here ===

This basic sequence of sensing, inference and display is a familiar one in emotional computing. The trick for our system was that the outputs were purposely fashioned to be open-ended and ambiguous, and to undermine the authority of the system's judgements. For instance, we chose to use automatically generated horoscopes in the first iteration to take advantage of a culturally familiar genre in which diagnoses and predictions, often of an emotional nature and expressed in ambiguous ways, are usually greeted by readers not as true or false but as ideas to be entertained. We hoped similarly to encourage people to 'try on' the interpretations of the Home Health Systems by using ambiguity and subversion, thus applying a computational approach to emotion without usurping people's authority. The notion was that, despite difficulties in accurately inferring emotion automatically, the interpretations produced by such systems might encourage and provide resources for people's own more accurate accounts. In other words, if there is a continuum between effective randomness and total accuracy in systems' ability to monitor emotional wellbeing, we believed we could locate a 'sweet spot' between the two in which systems might spur user interpretation of events in ways that would be based upon, but be more accurate than, the interpretations of the technical system itself (see Figure 2).

=== insert Figure 2 about here ===

A number of interesting implications might follow if this *User Appropriated Inference* concept were valid. It would mean that emotional computing systems would not have to build comprehensive representations in order to support user understanding. Instead, much like information visualisation software, the trick would be for such systems to provide information in a way that would support people's own pattern recognition abilities. This approach might help alleviate concerns for intrusiveness and invasion of privacy, since if system inferences are assumed to be inherently flawed, the emphasis should be on developing more evocative sensors and outputs rather than more accurate ones. Moreover, accurate user inferences would depend on local knowledge, limiting the ability for outsiders to use system data in meaningful ways. Finally, an approach based on user-appropriated inferences might generalise to a great many domains, including, for instance, systems to support aging in place or energy efficiency.

In sum, we thought we had found a way to build on the appeal of emotional computing while avoiding many of the attendant hazards. Demonstrating this potential became a primary motivation for developing the system. As we shall see, focusing our design around demonstrating a new approach to emotional computing had unfortunate consequences for the system's development.

Trying Out the Home Health Systems

The first iteration, the Home Health Horoscope, was developed in participation with a fairly large household in North London consisting of a nuclear family with children in their late teens and early twenties as well as a changing cast of partners, friends and lodgers who stayed with them for varying lengths of time. We studied how their routines manifested themselves in sense-able attributes of their household during occasional visits over more than a year, and developed a series of a dozen sensor devices and a set of about thirty rules specifically for their household. These rules determined wellbeing metrics relevant for their arrangements, and were used to generate 'horoscopes'

automatically that were printed out once a day on a device in their home (see Gaver, Sengers et al. 2007 for details).

The household lived with the resulting system for several months, and we assessed their experience over this time using a combination of ethnographic observations and interviews, documentary film, and informal encounters occasioned by maintenance visits. Overall, the results were encouraging: we found that household members, and particularly our lead informant, engaged with the system continually throughout the deployment, regularly reading the horoscopes and relating them to ongoing activities. The horoscopes and overall system were the subject of many conversations within the household. Crucially, these discussions often centred not on whether the system understood the state and activities of the household accurately (e.g. 'the household is busy today)' but whether its interpretation of their emotional implications (e.g. 'you should slow down') was appropriate. In agreeing with the former while taking authority over the latter, the participants demonstrated the kind of relationship we had hoped to evoke.

The deployment was not an unmitigated success, however. The continual engagement with the system appeared motivated as much by questions about our research agenda as by interest in what the system was saying about the household. Moreover, the outputs were often seen as unequivocally inaccurate, to the extent that at least some participants speculated that the sensors might simply be fakes. Nonetheless, while the field trial presented clear signals of difficulty, we saw enough reason for optimism to develop a second iteration of the system.

The second iteration improved on the first in four major ways. First, we recruited a household that was physically and socially less complex than the first, to simplify the task and increase the accuracy of inferring wellbeing and to make the sensing infrastructure itself easier to implement. The new household was comprised of a couple living in a single-story apartment with their two cats, and as we expected their routines turned out to be simpler, and the physical space more tractable, than in the

first household. Second, we abandoned horoscopes as an output style because our first trial indicated that individual sentences tended to imply particular contexts in inappropriate ways, and because they could have undesirable cultural connotations either in the styles used to write them or as a genre. Instead we used 'readings' in the form of short sentences, often taking the form of mildly judgemental aphorisms, that we wrote ourselves; we later replaced these with photographs, as well as pie charts of the metric values themselves. Third, we increased the legibility of the sensors because we found that far from becoming 'invisible' (Weiser, 1991), the volunteers continually speculated about what they might actually be sensing. Thus we designed new sensors with physical extensions indicating what they might be sensing and their orientation, as well as small displays showing the number of events they picked up during the day (see Figure 3). Finally, we used a new approach in deploying the system because we found our first one, in which we told participants as little as possible about what to expect, excited intrigue and suspicion rather than the openness for which we had hoped. Thus we explained the system as we developed and installed it, and indicated exactly what each sensor was measuring as we put it in place. Overall, we hoped these changes would make the system more accurate and easier to understand than the first iteration.

=== insert Figure 3 about here ===

At first, the second Home Health Monitor deployment appeared promising. The sensor units and printer were easy to deploy, and the volunteers admired the way their finish and aesthetics fit the home. Moreover, the volunteers attended to the sensor units and their displays in order to make sure the system was working as they imagined was intended. For instance, early in the deployment they repositioned a set of pressure pads used to track sofa usage after noticing that an entire evening spent lying on the sofa went undetected. We were encouraged by these initial signs of engagement.

Over the next several weeks, however, it became increasingly apparent that the volunteers were not happy with the system. They

showed none of the distracting speculation or suspicion of the first household, but little of the excitement either. When asked how they liked the system, they would shrug and make mildly positive comments, often with a ruefully apologetic smile. They seldom elaborated on their impressions spontaneously, seemingly reluctant to dishearten us. Over time, however, it became clear that they were disappointed. We had recruited them on the recommendation of friends of theirs who had tried out a different prototype we had produced. Having seen that prototype, they were excited about the prospect of this deployment. As they became familiar with the Home Horoscope, however, they felt let down by the experience it provided, once even remarking they had hoped for something more like their friends had gotten.

One of the most obvious problems with the Home Health Monitor, to the volunteers, was that experience it offered was very thin. As far as they were concerned, the system did little more than print out a single card every day. Though the sensor displays might be checked from time to time, once the system was 'tuned' these offered few surprises and in any case were understood not to be the main emphasis of the system. Moreover, the complexity of the underlying technology used to implement the system – the visible sensor units scattered through the house – and the evident care put into their design only exacerbated this impression. As one of the volunteers put it: 'You would never imagine that it would require this much work to get so little out'.

Of course, we had hoped that the emotional diagnoses offered by the cards might provide opportunity for ongoing engagement, for instance in the form of periodic discussions during the day. After all, that was intended to be where the 'action' of the system would be: in sensing, interpreting, and commenting on the household's emotional wellbeing in evocative ways. Usually, however, the system's output was seen as simply redundant when correct, and annoying when wrong. The ambiguous outputs were not usually found intriguing, but merely irksome. This was exacerbated by their perceived inconsistency. The system did not use the history of its previous outputs in choosing new ones, so its apparent judgements from day to day could seem contradictory,

and this undermined occasions when it did seem insightful. For instance, after D was sick and lying on the sofa all day, the system's 'we are closer to ants than butterflies' captured how she felt. However, D was quick to put this insight into perspective: 'The day before I got 'beware the barrenness of an easy life' so it could just think I am lying around being lazy'.

As it became clear to us that the system was not working, we rather desperately sought to improve it. We thought that using pictures rather than text might help matters by allowing greater ambiguity, reducing the problem of inconsistent or inappropriate judgements, and richer grounds for engagement and curiosity. But pictures imply their own, often over-definite contexts, and have connotations that may be even more difficult to control than those of text. For instance, on seeing a photograph intended to convey a stable home life, one of the volunteers remarked 'I don't like ironing, so I am not sure what it is saying to me'. The use of personal photos as well as those sourced from the web caused further problems, when there were tensions between the meanings the volunteers had invested in them and the reasons they imagined the system had for selecting them. Occasionally, the volunteers would juxtapose the pie-chart depiction of wellbeing metrics with the photographs, but this was usually done to diagnose problems rather than out of a sense of pleasurable insight. Overall, the new outputs did not change their problematic relationship with the system. Far from spurring the kind of critical reappropriation of emotional interpretation we had anticipated, the volunteers' relationship with the system was characterised more by a kind of frustrated irritation, and eventually by withdrawal and indifference.

As we admitted that the system had not succeeded, both amongst the design team and with the volunteers, our conversations became easier if no less disappointing.
Often they would drift to comparisons to other technologies, though these further reinforced an unfavourable assessment. As one of the volunteers put it: 'I just don't see how I could benefit from it. I don't see the point of many of these technologies. Other than being a gadget what's the point? I don't

like the idea of a system knowing whether you are home or not, unless you were vulnerable and needed some system looking after you.' Many of our discussions centred on surveillance cameras, location-tracking of children and the Big Brother society, and it became apparent that they saw the Home Health Monitor as an instance of this kind of objectionable use of technology.

What Went Wrong

In the first iteration, we focused on signs of success and hoped the difficulties we observed could be overcome. In the second, the problems were too evident and too fundamental to be dismissed. As we reluctantly admitted to ourselves, over the first six weeks or two months of the deployment, that the Home Health Monitor had failed to engage the volunteers as we hoped, we started to reflect on the causes of this failure rather than seeking to find some evidence for success. Here we discuss some of our more salient speculations and their relevance for designs based on computational approaches to emotion more generally.

An obvious reason that the system might have failed is simply because it was poorly implemented. We may have used too few sensors, chosen the wrong sites for their deployment, or positioned them badly. The rules we used for mapping sensor data to wellbeing metrics may have been inappropriate, biased, too complicated or too simple. The metrics themselves could have been poorly chosen. The outputs may have been badly designed. And so on. In short, we may simply have constructed a bad instantiation of the Home Health idea, and a more expert group may have done a better job.

Remember, however, that the system was designed to find a 'sweet spot' between randomness and accuracy that would be more interesting than either extreme. This implies that the system should be forgiving of implementation problems leading to inaccuracy. But we clearly failed to demonstrate an engaging level of partial inaccuracy. This leads to two possible conclusions. First, it might be that developing a system to infer wellbeing accurately enough to be distinguishable from chance is far more difficult than we had thought. Alternatively, the

boundary between randomness and accuracy may be more of a knife's-edge than a sweet spot. Either way, the implication for future designs based on the computation of emotion seems to be that satisfactory performance may depend on a degree of accuracy that is extremely difficult to achieve.

If our volunteers failed to appreciate a degree of evocative ambiguity in the system, instead perceiving its outputs either as annoyingly inaccurate or tediously redundant, this might be because the application was misconceived. Given how simple the household was, the system was unlikely to provide new information to its members. From this point of view, we might have had more success if we had used the system in a more complicated household, or to communicate between two households. In either case, the system's judgements would be more likely to convey a perspective novel at least to some of the inhabitants. The more promising results of the first iteration, which involved relatively complex domestic arrangements, give some backing to this conjecture. The reason we had chosen a simpler household for the second iteration was to make the automatic inferencing requirements more tractable, however, implying a tradeoff between the potential interest of the system and the requirements of its development. Moreover, our volunteers did not seem to find a system that would comment on emotionally relevant aspects of their home life appealing in the first place. Instead, they perceived the system as related to a variety of surveillance and monitoring systems that they dislike both personally and for their cultural and political implications. For these people, at least, it is questionable that any system depending on automatic tracking of emotions would be appealing enough to overcome concerns for privacy.

Another perspective on the Home Health Monitor's failure emphases the interaction style it embodied. It used a strategy of *information narrowing*, in which, each day, hundreds of data points from multiple sensors were distilled to a single, one-sentence reading. The danger of this strategy is that a single output appears meagre and prone to error. If it is wrong there is no fallback position. This contrasts with a strategy we have used

in other designs (such as the Drift Table, described later), which we might think of as *information widening*. In these systems, the output from one or a few sensors is used to provide access to a much richer output dataset. Such a strategy seems more promising as a way of creating engaging experiences from mundane domestic activity. Most applications based on computation of emotion use an information narrowing approach, however, seeking to draw high-level inferences from masses of lower-level data. This may make them vulnerable to the sort of brittleness we witnessed with the Home Health systems.

A final speculation about the reasons for the Home Health systems' failure has to do with the process we used to develop it. Usually we pursue design as research, in which we focus on developing systems that are compelling and finished in their own terms, and with respect their aesthetic, emotional, social and cultural implications. Methodological and conceptual innovations usually emerge as a result of this practice. The development of the Home Health systems turned our normal practice on its head. Our interest in the user-appropriated inference concept overwhelmed the designs themselves, soon after the initial concept emerged. Instead of designing the systems in their own right, their development became an exercise in illustrating a conceptual point. We ended up pursuing design *for* research rather than *as* research. Pragmatically, this seemed to distract our attention from such basic questions as whether people would actually be interested in reflecting on wellbeing in the home and how a system might support this successfully.

This last point may seem the most subtle, but it is also fundamental to my perspective: design works best when grounded in the details of a rich and complex situation, rather than one or a few abstract concepts.

Learning From Failure

Our experiences with the Home Health systems underscored the pitfalls of computational approaches to emotion discussed in the first part of this essay. Developing systems that volunteers

perceived as accurate turned out to be very difficult, and we had difficulty demonstrating a 'sweet spot' between outputs perceived as redundant or wrong. The distillation of data from multiple sensors to simple inferences about wellbeing did not seem compelling to our volunteers, and whatever interest they had in reflecting on the emotional wellbeing of their households was undercut by concerns about surveillance and privacy. To be sure, there were differences between the two volunteer households, which highlights the fact that people are half the equation for systems such as the Home Health Monitor. Nonetheless, it is impossible to avoid the conclusion that overall, the systems simply failed to realise the promise of emotional computing.

None of this serves as strong evidence against a programme of design based on computational approaches to emotion. Nor do the issues about such an approach that I raise in the first part of this essay seem resolvable by reason alone. As a designer, I believe that the potential of developing systems based on computational approaches to emotion will be proven by example, not argument, and admit that others may succeed where we have failed. Nonetheless, both the arguments I raise and the experiences I discuss may serve as resources in personal, pragmatic judgments about the likely fruitfulness of a programme of design based on using computational approaches to emotion. Based on both, I suggest that designing to reflect emotions as part of the complexity of lived experience is both more tractable, and leads to richer and more engaging results, than a computational approach to emotion in particular or a focus on emotion, per se, in general.

THE DRIFT TABLE

Let me end this discussion with an example from my studio's practice that may serve to illustrate what it means to design for experiences that are emotional without putting emotion at the centre of design, and also give a sense of the practice that serves as context for these remarks. Several years ago, as part of a project on designing devices for the home, we developed a prototype called the Drift Table (Gaver, Bowers et al. 2007). The Drift Table is a fairly small ($1m^2$) wheeled table with a circular

viewport in the centre of its top surface that shows slowly scrolling aerial photography (Figure 4). The impression is of looking through a window as one slowly floats several hundred metres above the ground. Four load sensors measure the centre of gravity of weights distributed on the table's surface, so that shifting weights towards one side of the table causes it to drift in that direction, adding weights causes it to go lower and faster, and removing them causes it to drift slowly and randomly at a relatively high altitude. A small display shows the current location, and a tiny reset button allows the view to be switched back to the home location, but other than that, and apart from a hidden on/off switch, no other controls are given. One can simply shift weights around the table to explore the vast landscape – about a terabyte of high-resolution photography, covering all of England and Wales – to which it gives access.

=== insert Figure 4 about here ===

The Drift Table was designed to explore ludic engagement within the home, where 'ludic' refers to playful, self-motivated exploration based on curiosity and whim (Gaver, 2009). We were interested in this as an alternative to more typical utilitarian applications of technology within the home. Thus the Drift Table was not conceived to solve any problems or pursue any tasks, but simply to offer an engaging situation that people could explore for themselves. We had some ideas of how people might use it, of course, but we didn't really know. So, as with many of our prototypes, we loaned it to a volunteer household to live with over several months, and observed what they did with it using a combination of ethnographic observation, unstructured interviews, and documentary film.

What we found, briefly, is that some of the volunteers engaged with the Drift Table intensely throughout their tenure, while others lost interest either because they did not find electronic devices appealing in principle, or because it was insufficiently interactive to maintain their interest. Those who did engage with the table worked with it far more intently than we had ever imagined. They routinely set off on journeys of several hundreds

of miles, despite the fact that these would take hours and involve difficult challenges of navigation. They flew to their old hometowns to view remembered landmarks, visited friends' neighbourhoods to spot features they could drop into conversation, explored areas they'd heard about on the news, and so on. One of the household members worked at home, and reported taking breaks to readjust the Table's course as a routine activity during the day. Several of the enthusiasts reported gathering around the table in the evening, when it would serve as an alternative to television as a focus of activity and conversation. They compared it to other things they enjoyed, ranging from a late-night transmission of satellite imagery to a particularly interesting airplane journey. In short, the Table engaged them over time and in many ways.

Integral to the volunteer's experience with the Drift Table was the emotions that arose during its use. These ranged from excitement and anticipation during its original deployment, to feelings of disappointment and frustration when its limitations became known, to a sense of fascination, intrigue and perhaps a sense of pride as they learned to value it in spite of – even because of – its constraints. Their emotions were aroused not only by the device itself, but also by the landscape to which it gave access. There were moments of delight in seeing hidden features of their local neighbourhood, disgust at miles of urban sprawl, nostalgia for a childhood home. They expressed these emotions during our conversations with them, and in the documentary video we commissioned from an independent filmmaker as a way of finding new perspectives on the volunteers' life with the prototype.

The volunteers' emotional reactions to the Drift Table were only one aspect of their overall experience with it, however. They also appreciated it conceptually as a technological device, as a potential domestic product, and as a device offering certain opportunities and challenges. They valued it aesthetically for its design and interaction (although they felt the wheels let it down a bit). The access it gave to the countryside was both personally compelling and intellectually interesting. It also

served to facilitate sociality within the home (when they gathered around to discuss the view) and occasionally to thwart it (when they found that only a few people could look through the viewport at once, or argued over the inadvertent misplacement of objects on the surface).

In sum, the Drift Table was compelling to our volunteers. Part of its impact was emotional, but this arose and found its meaning in relation to aesthetic, conceptual, functional and social appreciation. Moreover, these dimensions of appreciation were integrated in our volunteers' lived experiences to the point that distinguishing amongst them is somewhat misleading and unhelpful (McCarthy & Wright, 2004). Mirroring this, we had not distinguished these facets of experience in designing the Drift Table. We did not seek to design an emotionally compelling experience any more (or less) than we set out to design a conceptually resonant one. Or rather, we set out to design for all of these things, but not as analytically articulated desiderata but as integrated and embodied in a design that we tried to make as rich and compelling as possible.

This is the crux of my argument: that rather than singling out emotion as an object of attention, and working to explicitly recognise and represent it, it is more fruitful to recognise emotion as an emergent aspect of experiences that are situated and multi-layered. This leads to a design-led research approach that focuses on crafting the appearance and interactivity of specific designs open to ludic engagement on the part of their users. If done well, the designs will both embody understandings of emotion, aesthetics, sociality and culture and lead to new insights as well[iii]. Emotion may be an important facet of the understandings and insights successful designs rely on and produce, but it is not the only one, and not always the most important.

ACKNOWLEDGEMENTS

The research reported here was pursued in collaboration with Andy Boucher, John Bowers, Nadine Jarvis and Tobie Kerridge of the Interaction Research Studio, and with Phoebe Sengers and Jofish

Kaye from Cornell University, with support from Intel Corporation and the Equator Interdisciplinary Research Collaboration. Thanks to John Bowers for introducing the distinction between information narrowing and widening, to him, Kia Höök, Anne Schlottmann and Phoebe Singers for comments on an earlier draft of this paper, and to the volunteers who opened their homes to our systems and us.

REFERENCES

Arthur, C. (2009), It's always best to keep your own lie detector turned on, *The Guardian*, 21 April 2009),

Boehner, K., DePaula, R., Dourish, P., and Sengers, P. (2005). Affect: from information to interaction. *Proceedings of the Conference on Critical Computing*. ACM, pp. 59–68.

Boehner, K., Sengers, P., and Warner, S. (2008). Interfaces with the ineffable: Meeting aesthetic experience on its own terms. *ACM Trans. Comput.-Hum. Interact*. 15, 3, pp. 1–29.

Cowie, R. (this volume). Perceiving emotion: towards a realistic understanding of the task. To appear in this volume.

Gaver, W. (1991). Technology affordances. In Proceedings of CHI'91, pp. 79–84

Gaver, W., (2009). Designing for Homo Ludens, Still. In *(Re)searching the Digital Bauhaus*. Binder, T., Löwgren, J., and Malmborg, L. (eds.). London: Springer, pp. 163–178.

Gaver, W., Bowers, J., Boucher, A., Law, A., and Pennington, S. (2007). Electronic furniture for the Curious Home: Assessing ludic designs in the field. *International Journal of Human-Computer Interaction,* 22(1-2): 119–152.

Gaver, W., Bowers, J., Kerridge, T., Boucher, A., and Jarvis, N. (2009). Anatomy of a failure: How we knew when our design went wrong, and what we learned from it. *Proceedings of CHI '09*.

Gaver, W., Sengers, P., Kerridge, T., Kaye, J., and Bowers, J. (2007). Enhancing ubiquitous computing with user interpretation: field testing the home health horoscope. *Proceedings of CHI '07*, pp. 537–546.

Greenfield, A. (2006) Everyware : the dawning age of ubiquitous computing. Berkeley, CA : New Riders.

Höök, K. (this volume). Affective loop experiences: Designing for interactional embodiment. To appear in this volume.

Hutchins, E., Hollan, J., and Norman, D. (1985). Direct Manipulation Interfaces. Human-Computer Interaction 1(4). pp. 311—38.

Mandler, G. (1984) Mind and body : psychology of emotion and stress. New York : W.W. Norton.

McCarthy J, and Wright P. (2004). Technology as Experience. MIT Press.

Romero, M., Pousman, Z., and Mateas, M. (2006). Tableau machine: an alien presence in the home. *CHI '06 Extended Abstracts*, pp. 1265–1270.

Sengers, P., Boehner, K., David, S., and Kaye, J. (2005). Reflective design. *Proceedings of the Conference on Critical Computing*. ACM, pp. 49–58

Smith, R. B. (1986) Experiences with the alternate reality kit: an example of the tension between literalism and magic. *SIGCHI Bulletin* 17, SI (May 1986).

Suchman, L. (1987) Plans and situated actions. Cambridge : Cambridge University Press.

M. Weiser, (1991), "The Computer of the 21st Century," *Scientific American*, vol. 265, no. 3, pp. 66–75.

Zimmerman, J., Forlizzi, J., and Evenson, S. (2007). Research through design as a method for interaction design research in HCI. Proceedings of CHI '07, pp. 493–502.

FIGURE CAPTIONS

Figure 1: The basic architecture of the Home Health Monitors.

Figure 2: A 'sweet spot' between accurate and random inferences about emotion might be tractable and engaging.

Figure 3: A sensor designed for legibility.

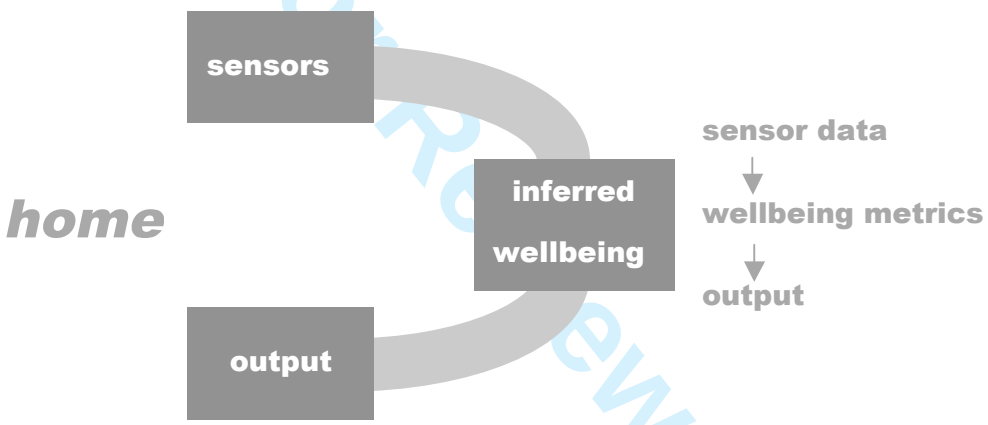
Figure 4: The Drift Table shows slowly scrolling aerial photography through the viewport on its top.

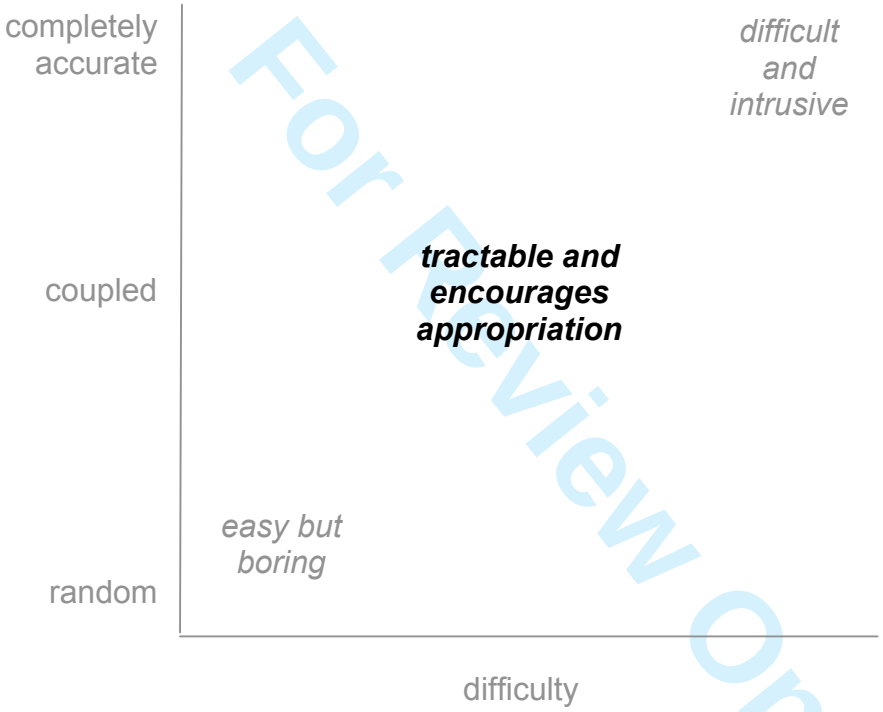
FOOTNOTES

[i] I use terms such as ‘emotion’, ‘mood’, and ‘feelings’ interchangeably in this essay. The reasons for this – that I believe distinguishing among them in design is not often a useful endeavour – should become clear through the course of the discussion.

[ii] The advent of search engines and similar interfaces may be seen as a massive revival of command-line languages, but whether they are to be viewed as agents or as machines is still an open question (they certainly don’t seem emotional).

[iii] Sometimes, such an approach might even be taken to emotional communication (e.g. Boehner, Sengers & Warner, 2008) and reflection (Höök, this volume), not as generically represented and computed but as emergent in specific interactive systems.





A sensor designed for legibility.
162x215mm (600 x 600 DPI)

The Drift Table shows slowly scrolling aerial photography through the viewport on its top.
903x677mm (600 x 600 DPI)