

### Designing for Emotion (Among Other Things)

Journal:	<i>Philosophical Transactions B</i>
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 <a href="http://rstb.royalsocietypublishing.org/site/misc/issue-codes.xhtml">here</a> 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<sup>i</sup> appears extremely difficult. Systems based on facial recognition

Gaver - Designing for Emotion

2

1  
2  
3 may have had some success in the laboratory, but encounter severe  
4 difficulties in the field (e.g. Cowie, this volume). The problem  
5 is that our emotions are fleeting, mixed and masked. For example,  
6 consider a scenario in which you take a wrong turn while driving  
7 a friend home. You might experience a surge of frustration,  
8 perhaps even a small twinge of fear as you realise you're  
9 disoriented. Still, you have to laugh when your friend makes a  
10 joke at your expense. Feigning outrage, you put on a mask of  
11 confidence as you assure her that you know where you are going.  
12 All the time you are happy to be in her company, even if a bit  
13 tired after a long day. Let's say your in-car navigation system  
14 is trying to track your emotions to guide how it presents  
15 information. What should it do? Is it really going to be able to  
16 make sense of such a fluid amalgamation of emotional states, much  
17 less respond appropriately?  
18  
19  
20  
21  
22  
23  
24

25  
26 Even if systems can somehow disambiguate a mixture of basic  
27 emotions based on some combination of facial recognition and  
28 physiological monitoring, this is only a small part of  
29 comprehending the meaning of an emotional experience. For  
30 instance, the other day I was happy and excited to catch a large  
31 tench while fishing in the countryside. Later that same day, I  
32 played a chaotic mixture of hide and seek, catch, and tickling  
33 with my two young daughters. Again I was happy and excited. Even  
34 if we grant that my emotions in these two situations were somehow  
35 the same, it is clear that the meaning of these situations, the  
36 relationships I had with the fish and with my little girls, were  
37 quite different. This suggests that for systems to do a good job  
38 of reasoning about emotional experiences, they will have to  
39 understand a great deal about the context in which the emotions  
40 arise. If we believe, as some do, that emotions are the product  
41 of arousal and cognitive assessment (Mandler, 1984), then each  
42 emotion is specific to its circumstances, with the linguistic  
43 terms we use to identify them pointing merely to family  
44 resemblances. In other words, emotions are situated (Suchman,  
45 1987). Similar reasoning based has led some researchers (e.g.  
46 Boehner et al., 2005) to consider accurately sensing emotion to  
47 be impossible in principle.  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Gaver - Designing for Emotion

3

1  
2  
3 But let us assume that we can accurately sense emotions in a way  
4 that is either independent of context or includes it  
5 sufficiently. Do we want to? The social and cultural implications  
6 of emotional sensing are also troublesome.  
7  
8

9  
10 To begin with, it is easy to imagine that automatic sensing of  
11 emotions would intrude on our privacy in significant ways. What  
12 if your closest friend bought a gadget that allowed them to track  
13 your affective state accurately over time. Would you want them  
14 to have access to your unmediated emotions? Would you be able to  
15 decline if they exerted emotional pressure? Maybe it would be  
16 fun, at least for a while. This is only the tip of the iceberg,  
17 however. Supermarkets might track your emotional responses as you  
18 navigate the aisles, using their existing security cameras to  
19 augment market research. Web companies might be interested in  
20 keystroke patterns and mouse pressure as a measure of your  
21 emotional reactions to various search results. They might sell  
22 this information on to your insurance company or bank. Clearing  
23 immigration at the airport would involve officers tracking your  
24 emotional responses as they ask leading questions about your  
25 political and religious beliefs, passing the results to the  
26 police if necessary. If automatic tracking of emotions is  
27 successful, the ways this information is used will be as  
28 difficult to trace, much less control, as any of the personal  
29 information that is currently tracked and trafficked on the web  
30 (Greenfield, 2006).  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40

41 Even more insidious, perhaps, is the potential for emotional  
42 tracking to alter our relationships to our emotions and to each  
43 other. Consider systems proposed to support older people living  
44 at home ('aging in place'). Automatic sensing of emotion could  
45 augment systems that track whether older people have fallen, or  
46 failed to get out of bed, or to eat adequately. By tracking the  
47 emotions of older people in their homes, the argument goes, we  
48 could reassure ourselves of their continued wellbeing. The danger  
49 is that such systems might automate caring relationships  
50 traditionally maintained through personal social contact.  
51 Naturally we wouldn't consciously abandon the aged to automatic  
52 care. But imagine being in a hurry to get home, and wondering  
53  
54  
55  
56  
57  
58  
59  
60

Gaver - Designing for Emotion

4

1  
2  
3 whether to visit an older friend on the way. Wouldn't this be  
4 less likely if you had a device to reassure you not only that  
5 they were active and safe, but showing all the physiological and  
6 expressive signs of happiness as well?  
7  
8

9  
10 The problem is that such tools are designed inherently to stand  
11 in for our own personal assessments of emotion. Charles Arthur  
12 (2009) tells a story that illustrates this point vividly. US  
13 police sometimes question subjects with a 'lie-detector', asking  
14 them to place their hands on the machine, which then produces  
15 sheets of paper reading 'true' and 'false' as they answer  
16 questions. The trick is that the machine is an office photocopier  
17 and both the questions and judgements have been previously  
18 scripted. Apparently many suspects are fooled by this, leading  
19 Arthur to comment: 'our tendency to believe what machines tell us  
20 - even if we don't understand them - still baffles me'. This  
21 tendency may lead us to trust devices that purport to sense  
22 emotions more than our own perceptions, whether they are seen as  
23 tools for monitoring emotion at a distance, or in people who are  
24 physically co-present, or even in ourselves.  
25  
26  
27  
28  
29  
30  
31  
32

33 Arguments about privacy and automation can be seen as luddite  
34 reactions to new technologies, of course. Similar reasoning might  
35 lead to the conclusion that thermometers are immoral because they  
36 dissuade us from attending to our own experience of temperature  
37 (Cowie, personal communication). From this perspective, if we can  
38 learn to manage our privacy and relationships over the web, then  
39 we can do the same in the face of emotional sensing. These sorts  
40 of debates about the cultural desirability of new technologies  
41 are difficult to decide, though we can all make choices based on  
42 their consideration.  
43  
44  
45  
46  
47

48 In any case, my doubts about computational approaches to emotion  
49 as a basis for human computer interaction go beyond these broad  
50 technical and cultural concerns to include more personal beliefs  
51 about fruitful approaches to interaction design. The basic  
52 question is whether we want to model interfaces after agents or  
53 environments. Do we want to pretend that computers are sentient  
54 beings with whom we can converse and form relationships, or is it  
55  
56  
57  
58  
59  
60

1  
2  
3 more fruitful to think of them as virtual environments, furnished  
4 with a variety of tools, machines and objects, in which we can  
5 explore and pursue our aims? My view on this dates back to the  
6 early 1980's, when direct manipulation (a.k.a. GUI) superseded  
7 command-line interfaces. The new interfaces presented an  
8 environment with its own blend of 'physics and magic' (Smith,  
9 1986) in which users have the illusion of acting directly on a  
10 computational world (Hutchins, Hollan & Norman, 1985). This  
11 paradigm shift seemed to make irrelevant both the rigid command  
12 languages of the time and the hype of natural language processing  
13 to come<sup>ii</sup>. Simultaneously, it opened a new design space in which  
14 designers could shape the affordances of virtual environments  
15 (Gaver, 1991) and, in the long run, of computational products.  
16 This seemed, and seems, a tremendously exciting prospect to me,  
17 while the notion of interfaces that promote the illusion of  
18 agency, encouraging people to project trust and affection upon  
19 them, appears not only difficult to achieve in likeable ways, but  
20 prone to produce systems that are inauthentic, patronising and  
21 manipulative.  
22  
23  
24  
25  
26  
27  
28  
29

30  
31 No doubt, artificial agents have been, and will be, deployed  
32 quite widely in computer games and interactive narratives, sales  
33 kiosks and toys such as the Pleio and Sony Aibo,. But I would  
34 argue that such applications are enjoyed as 'clever machines',  
35 digital versions of mechanical automata that most people  
36 recognise as non-agents even while enjoying their quasi-  
37 intelligence. Moreover, I suspect such applications will be  
38 accepted and enjoyed as long as playing along with the simulated  
39 emotional engagement they offer doesn't present serious hazards.  
40 From this point of view, recognisably artificial agents,  
41 especially used in relatively inconsequential domains, may offer  
42 interesting possibilities for designers (as I explore in the next  
43 section). On the whole, however, I remain committed to an  
44 approach in which computational devices are seen as environments  
45 complementing and supporting our abilities rather than seeking to  
46 emulate them.  
47  
48  
49  
50  
51  
52  
53

54  
55 My final doubt about pursuing computational approaches hinges on  
56 the view that emotion, however pursued, is seldom if ever an  
57  
58  
59  
60

1  
2  
3 appropriate focus for design. Clearly emotion is a crucial facet  
4 of experience. But saying that it is a 'facet of experience'  
5 suggests both that it is only one part of a more complex whole  
6 (the experience) and that it pertains to something beyond itself  
7 (an experience of *something*). It is that *something* - a chair, the  
8 home, the challenges of growing older - which is an appropriate  
9 object for design, and emotion is only one of many concerns that  
10 must be considered in addressing it. From this point of view,  
11 designing for emotion is like designing for blue: it makes a  
12 modifier a noun. Imagine being told to design something blue.  
13 Blue what? Whale? Sky? Suede shoes? The request seems  
14 nonsensical. Similarly, focusing design on emotion without a  
15 grounded sense of the situation in which emotions are meant to  
16 gain meaning appears to be a category error. Instead, we need to  
17 understand how to design for engaging experiences more generally.  
18  
19  
20  
21  
22  
23  
24  
25

#### 26 **A SAMPLE OF THE DIFFICULTIES: THE HOME HEALTH MONITOR**

27 My comments about the problems of using computational approaches  
28 to emotion in design are based not only on observation or  
29 reasoning, but on the bitter experience of developing and  
30 deploying two iterations of a system that, while serving both of  
31 as a gentle parody of and suggestion for improvement to  
32 traditional emotional computing approaches, still followed their  
33 essential logic.  
34  
35  
36  
37

38 The basic idea of the Home Health Monitor was to use sensors in  
39 the home to track household conditions that might be symptomatic  
40 of the overall emotional state, or mood, of its inhabitants. Data  
41 returned by these sensors are processed and used to build a  
42 representation of the household's 'wellbeing', defined broadly  
43 and relative to the particular household, and the outcome  
44 displayed to support reflection (see Figure 1). For instance, we  
45 might design a sensor device to measure when a given door is open  
46 or shut because the home's occupants have informed us that it is  
47 only closed when household members want to avoid each other. The  
48 raw sensor data is processed to uncover attributes such as the  
49 total time the door is open or closed during the day, how often  
50 it moves, or how early its state first changes. Rules compare the  
51 day's readings with trends found over the preceding days to  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 determine whether they are unusually high or low, and map this to  
4 an increment or decrement of, e.g., the 'sociality' metric  
5 accordingly. The pattern of metric scores provides a  
6 representation of the home's wellbeing that is mapped to an  
7 output for users. In the first iteration (Gaver, Sengers et al.,  
8 2007), the system constructed 'horoscopes' from sentences culled  
9 from online examples and categorised according to the wellbeing  
10 metrics; in the second iteration, we tried three different forms  
11 of output.  
12  
13  
14  
15  
16

17  
18 *=== insert Figure 1 about here ===*  
19

20 This basic sequence of sensing, inference and display is a  
21 familiar one in emotional computing. The trick for our system was  
22 that the outputs were purposely fashioned to be open-ended and  
23 ambiguous, and to undermine the authority of the system's  
24 judgements. For instance, we chose to use automatically generated  
25 horoscopes in the first iteration to take advantage of a  
26 culturally familiar genre in which diagnoses and predictions,  
27 often of an emotional nature and expressed in ambiguous ways, are  
28 usually greeted by readers not as true or false but as ideas to  
29 be entertained. We hoped similarly to encourage people to 'try  
30 on' the interpretations of the Home Health Systems by using  
31 ambiguity and subversion, thus applying a computational approach  
32 to emotion without usurping people's authority. The notion was  
33 that, despite difficulties in accurately inferring emotion  
34 automatically, the interpretations produced by such systems might  
35 encourage and provide resources for people's own more accurate  
36 accounts. In other words, if there is a continuum between  
37 effective randomness and total accuracy in systems' ability to  
38 monitor emotional wellbeing, we believed we could locate a 'sweet  
39 spot' between the two in which systems might spur user  
40 interpretation of events in ways that would be based upon, but be  
41 more accurate than, the interpretations of the technical system  
42 itself (see Figure 2).  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

53  
54 *=== insert Figure 2 about here ===*  
55  
56  
57  
58  
59  
60

1  
2  
3 A number of interesting implications might follow if this *User*  
4 *Appropriated Inference* concept were valid. It would mean that  
5 emotional computing systems would not have to build comprehensive  
6 representations in order to support user understanding. Instead,  
7 much like information visualisation software, the trick would be  
8 for such systems to provide information in a way that would  
9 support people's own pattern recognition abilities. This approach  
10 might help alleviate concerns for intrusiveness and invasion of  
11 privacy, since if system inferences are assumed to be inherently  
12 flawed, the emphasis should be on developing more evocative  
13 sensors and outputs rather than more accurate ones. Moreover,  
14 accurate user inferences would depend on local knowledge,  
15 limiting the ability for outsiders to use system data in  
16 meaningful ways. Finally, an approach based on user-appropriated  
17 inferences might generalise to a great many domains, including,  
18 for instance, systems to support aging in place or energy  
19 efficiency.  
20  
21  
22  
23  
24  
25  
26  
27

28  
29 In sum, we thought we had found a way to build on the appeal of  
30 emotional computing while avoiding many of the attendant hazards.  
31 Demonstrating this potential became a primary motivation for  
32 developing the system. As we shall see, focusing our design  
33 around demonstrating a new approach to emotional computing had  
34 unfortunate consequences for the system's development.  
35  
36  
37  
38

### 39 **Trying Out the Home Health Systems**

40  
41 The first iteration, the Home Health Horoscope, was developed in  
42 participation with a fairly large household in North London  
43 consisting of a nuclear family with children in their late teens  
44 and early twenties as well as a changing cast of partners,  
45 friends and lodgers who stayed with them for varying lengths of  
46 time. We studied how their routines manifested themselves in  
47 sense-able attributes of their household during occasional visits  
48 over more than a year, and developed a series of a dozen sensor  
49 devices and a set of about thirty rules specifically for their  
50 household. These rules determined wellbeing metrics relevant for  
51 their arrangements, and were used to generate 'horoscopes'  
52  
53  
54  
55  
56  
57  
58  
59  
60

Gaver - Designing for Emotion

9

1  
2  
3 automatically that were printed out once a day on a device in  
4 their home (see Gaver, Sengers et al. 2007 for details).  
5  
6

7  
8 The household lived with the resulting system for several months,  
9 and we assessed their experience over this time using a  
10 combination of ethnographic observations and interviews,  
11 documentary film, and informal encounters occasioned by  
12 maintenance visits. Overall, the results were encouraging: we  
13 found that household members, and particularly our lead  
14 informant, engaged with the system continually throughout the  
15 deployment, regularly reading the horoscopes and relating them to  
16 ongoing activities. The horoscopes and overall system were the  
17 subject of many conversations within the household. Crucially,  
18 these discussions often centred not on whether the system  
19 understood the state and activities of the household accurately  
20 (e.g. 'the household is busy today') but whether its  
21 interpretation of their emotional implications (e.g. 'you should  
22 slow down') was appropriate. In agreeing with the former while  
23 taking authority over the latter, the participants demonstrated  
24 the kind of relationship we had hoped to evoke.  
25  
26  
27  
28  
29  
30  
31  
32

33 The deployment was not an unmitigated success, however. The  
34 continual engagement with the system appeared motivated as much  
35 by questions about our research agenda as by interest in what the  
36 system was saying about the household. Moreover, the outputs were  
37 often seen as unequivocally inaccurate, to the extent that at  
38 least some participants speculated that the sensors might simply  
39 be fakes. Nonetheless, while the field trial presented clear  
40 signals of difficulty, we saw enough reason for optimism to  
41 develop a second iteration of the system.  
42  
43  
44  
45  
46

47 The second iteration improved on the first in four major ways.  
48 First, we recruited a household that was physically and socially  
49 less complex than the first, to simplify the task and increase  
50 the accuracy of inferring wellbeing and to make the sensing  
51 infrastructure itself easier to implement. The new household was  
52 comprised of a couple living in a single-story apartment with  
53 their two cats, and as we expected their routines turned out to  
54 be simpler, and the physical space more tractable, than in the  
55  
56  
57  
58  
59  
60

1  
2  
3 first household. Second, we abandoned horoscopes as an output  
4 style because our first trial indicated that individual sentences  
5 tended to imply particular contexts in inappropriate ways, and  
6 because they could have undesirable cultural connotations either  
7 in the styles used to write them or as a genre. Instead we used  
8 'readings' in the form of short sentences, often taking the form  
9 of mildly judgemental aphorisms, that we wrote ourselves; we  
10 later replaced these with photographs, as well as pie charts of  
11 the metric values themselves. Third, we increased the legibility  
12 of the sensors because we found that far from becoming  
13 'invisible' (Weiser, 1991), the volunteers continually speculated  
14 about what they might actually be sensing. Thus we designed new  
15 sensors with physical extensions indicating what they might be  
16 sensing and their orientation, as well as small displays showing  
17 the number of events they picked up during the day (see Figure  
18 3). Finally, we used a new approach in deploying the system  
19 because we found our first one, in which we told participants as  
20 little as possible about what to expect, excited intrigue and  
21 suspicion rather than the openness for which we had hoped. Thus  
22 we explained the system as we developed and installed it, and  
23 indicated exactly what each sensor was measuring as we put it in  
24 place. Overall, we hoped these changes would make the system more  
25 accurate and easier to understand than the first iteration.  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35

36  
37 *=== insert Figure 3 about here ===*  
38  
39

40 At first, the second Home Health Monitor deployment appeared  
41 promising. The sensor units and printer were easy to deploy, and  
42 the volunteers admired the way their finish and aesthetics fit  
43 the home. Moreover, the volunteers attended to the sensor units  
44 and their displays in order to make sure the system was working  
45 as they imagined was intended. For instance, early in the  
46 deployment they repositioned a set of pressure pads used to track  
47 sofa usage after noticing that an entire evening spent lying on  
48 the sofa went undetected. We were encouraged by these initial  
49 signs of engagement.  
50  
51  
52  
53  
54

55 Over the next several weeks, however, it became increasingly  
56 apparent that the volunteers were not happy with the system. They  
57  
58  
59  
60

1  
2  
3 showed none of the distracting speculation or suspicion of the  
4 first household, but little of the excitement either. When asked  
5 how they liked the system, they would shrug and make mildly  
6 positive comments, often with a ruefully apologetic smile. They  
7 seldom elaborated on their impressions spontaneously, seemingly  
8 reluctant to dishearten us. Over time, however, it became clear  
9 that they were disappointed. We had recruited them on the  
10 recommendation of friends of theirs who had tried out a different  
11 prototype we had produced. Having seen that prototype, they were  
12 excited about the prospect of this deployment. As they became  
13 familiar with the Home Horoscope, however, they felt let down by  
14 the experience it provided, once even remarking they had hoped  
15 for something more like their friends had gotten.  
16  
17  
18  
19  
20  
21  
22

23 One of the most obvious problems with the Home Health Monitor, to  
24 the volunteers, was that experience it offered was very thin. As  
25 far as they were concerned, the system did little more than print  
26 out a single card every day. Though the sensor displays might be  
27 checked from time to time, once the system was 'tuned' these  
28 offered few surprises and in any case were understood not to be  
29 the main emphasis of the system. Moreover, the complexity of the  
30 underlying technology used to implement the system - the visible  
31 sensor units scattered through the house - and the evident care  
32 put into their design only exacerbated this impression. As one of  
33 the volunteers put it: 'You would never imagine that it would  
34 require this much work to get so little out'.  
35  
36  
37  
38  
39  
40  
41

42 Of course, we had hoped that the emotional diagnoses offered by  
43 the cards might provide opportunity for ongoing engagement, for  
44 instance in the form of periodic discussions during the day.  
45 After all, that was intended to be where the 'action' of the  
46 system would be: in sensing, interpreting, and commenting on the  
47 household's emotional wellbeing in evocative ways. Usually,  
48 however, the system's output was seen as simply redundant when  
49 correct, and annoying when wrong. The ambiguous outputs were not  
50 usually found intriguing, but merely irksome. This was  
51 exacerbated by their perceived inconsistency. The system did not  
52 use the history of its previous outputs in choosing new ones, so  
53 its apparent judgements from day to day could seem contradictory,  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 and this undermined occasions when it did seem insightful. For  
4 instance, after D was sick and lying on the sofa all day, the  
5 system's 'we are closer to ants than butterflies' captured how  
6 she felt. However, D was quick to put this insight into  
7 perspective: 'The day before I got 'beware the barrenness of an  
8 easy life' so it could just think I am lying around being lazy'.  
9  
10  
11

12  
13 As it became clear to us that the system was not working, we  
14 rather desperately sought to improve it. We thought that using  
15 pictures rather than text might help matters by allowing greater  
16 ambiguity, reducing the problem of inconsistent or inappropriate  
17 judgements, and richer grounds for engagement and curiosity. But  
18 pictures imply their own, often over-definite contexts, and have  
19 connotations that may be even more difficult to control than  
20 those of text. For instance, on seeing a photograph intended to  
21 convey a stable home life, one of the volunteers remarked 'I  
22 don't like ironing, so I am not sure what it is saying to me'.  
23 The use of personal photos as well as those sourced from the web  
24 caused further problems, when there were tensions between the  
25 meanings the volunteers had invested in them and the reasons they  
26 imagined the system had for selecting them. Occasionally, the  
27 volunteers would juxtapose the pie-chart depiction of wellbeing  
28 metrics with the photographs, but this was usually done to  
29 diagnose problems rather than out of a sense of pleasurable  
30 insight. Overall, the new outputs did not change their  
31 problematic relationship with the system. Far from spurring the  
32 kind of critical reappropriation of emotional interpretation we  
33 had anticipated, the volunteers' relationship with the system was  
34 characterised more by a kind of frustrated irritation, and  
35 eventually by withdrawal and indifference.  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45

46  
47 As we admitted that the system had not succeeded, both amongst  
48 the design team and with the volunteers, our conversations became  
49 easier if no less disappointing.  
50

51 Often they would drift to comparisons to other technologies,  
52 though these further reinforced an unfavourable assessment. As  
53 one of the volunteers put it: 'I just don't see how I could  
54 benefit from it. I don't see the point of many of these  
55 technologies. Other than being a gadget what's the point? I don't  
56  
57  
58  
59  
60

1  
2  
3 like the idea of a system knowing whether you are home or not,  
4 unless you were vulnerable and needed some system looking after  
5 you.' Many of our discussions centred on surveillance cameras,  
6 location-tracking of children and the Big Brother society, and it  
7 became apparent that they saw the Home Health Monitor as an  
8 instance of this kind of objectionable use of technology.  
9  
10  
11

### 12 **What Went Wrong**

14 In the first iteration, we focused on signs of success and hoped  
15 the difficulties we observed could be overcome. In the second,  
16 the problems were too evident and too fundamental to be  
17 dismissed. As we reluctantly admitted to ourselves, over the  
18 first six weeks or two months of the deployment, that the Home  
19 Health Monitor had failed to engage the volunteers as we hoped,  
20 we started to reflect on the causes of this failure rather than  
21 seeking to find some evidence for success. Here we discuss some  
22 of our more salient speculations and their relevance for designs  
23 based on computational approaches to emotion more generally.  
24  
25  
26  
27  
28  
29

30 An obvious reason that the system might have failed is simply  
31 because it was poorly implemented. We may have used too few  
32 sensors, chosen the wrong sites for their deployment, or  
33 positioned them badly. The rules we used for mapping sensor data  
34 to wellbeing metrics may have been inappropriate, biased, too  
35 complicated or too simple. The metrics themselves could have been  
36 poorly chosen. The outputs may have been badly designed. And so  
37 on. In short, we may simply have constructed a bad instantiation  
38 of the Home Health idea, and a more expert group may have done a  
39 better job.  
40  
41  
42  
43  
44

45 Remember, however, that the system was designed to find a 'sweet  
46 spot' between randomness and accuracy that would be more  
47 interesting than either extreme. This implies that the system  
48 should be forgiving of implementation problems leading to  
49 inaccuracy. But we clearly failed to demonstrate an engaging  
50 level of partial inaccuracy. This leads to two possible  
51 conclusions. First, it might be that developing a system to infer  
52 wellbeing accurately enough to be distinguishable from chance is  
53 far more difficult than we had thought. Alternatively, the  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 boundary between randomness and accuracy may be more of a  
4 knife's-edge than a sweet spot. Either way, the implication for  
5 future designs based on the computation of emotion seems to be  
6 that satisfactory performance may depend on a degree of accuracy  
7 that is extremely difficult to achieve.  
8  
9

10  
11 If our volunteers failed to appreciate a degree of evocative  
12 ambiguity in the system, instead perceiving its outputs either as  
13 annoyingly inaccurate or tediously redundant, this might be  
14 because the application was misconceived. Given how simple the  
15 household was, the system was unlikely to provide new information  
16 to its members. From this point of view, we might have had more  
17 success if we had used the system in a more complicated  
18 household, or to communicate between two households. In either  
19 case, the system's judgements would be more likely to convey a  
20 perspective novel at least to some of the inhabitants. The more  
21 promising results of the first iteration, which involved  
22 relatively complex domestic arrangements, give some backing to  
23 this conjecture. The reason we had chosen a simpler household for  
24 the second iteration was to make the automatic inferencing  
25 requirements more tractable, however, implying a tradeoff between  
26 the potential interest of the system and the requirements of its  
27 development. Moreover, our volunteers did not seem to find a  
28 system that would comment on emotionally relevant aspects of  
29 their home life appealing in the first place. Instead, they  
30 perceived the system as related to a variety of surveillance and  
31 monitoring systems that they dislike both personally and for  
32 their cultural and political implications. For these people, at  
33 least, it is questionable that any system depending on automatic  
34 tracking of emotions would be appealing enough to overcome  
35 concerns for privacy.  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47

48 Another perspective on the Home Health Monitor's failure emphasizes  
49 the interaction style it embodied. It used a strategy of  
50 *information narrowing*, in which, each day, hundreds of data  
51 points from multiple sensors were distilled to a single, one-  
52 sentence reading. The danger of this strategy is that a single  
53 output appears meagre and prone to error. If it is wrong there is  
54 no fallback position. This contrasts with a strategy we have used  
55  
56  
57  
58  
59  
60

1  
2  
3 in other designs (such as the Drift Table, described later),  
4 which we might think of as *information widening*. In these  
5 systems, the output from one or a few sensors is used to provide  
6 access to a much richer output dataset. Such a strategy seems  
7 more promising as a way of creating engaging experiences from  
8 mundane domestic activity. Most applications based on computation  
9 of emotion use an information narrowing approach, however,  
10 seeking to draw high-level inferences from masses of lower-level  
11 data. This may make them vulnerable to the sort of brittleness we  
12 witnessed with the Home Health systems.  
13  
14  
15  
16  
17  
18

19 A final speculation about the reasons for the Home Health  
20 systems' failure has to do with the process we used to develop  
21 it. Usually we pursue design as research, in which we focus on  
22 developing systems that are compelling and finished in their own  
23 terms, and with respect their aesthetic, emotional, social and  
24 cultural implications. Methodological and conceptual innovations  
25 usually emerge as a result of this practice. The development of  
26 the Home Health systems turned our normal practice on its head.  
27 Our interest in the user-appropriated inference concept  
28 overwhelmed the designs themselves, soon after the initial  
29 concept emerged. Instead of designing the systems in their own  
30 right, their development became an exercise in illustrating a  
31 conceptual point. We ended up pursuing design *for* research rather  
32 than as research. Pragmatically, this seemed to distract our  
33 attention from such basic questions as whether people would  
34 actually be interested in reflecting on wellbeing in the home and  
35 how a system might support this successfully.  
36  
37  
38  
39  
40  
41  
42  
43

44 This last point may seem the most subtle, but it is also  
45 fundamental to my perspective: design works best when grounded in  
46 the details of a rich and complex situation, rather than one or a  
47 few abstract concepts.  
48  
49  
50

### 51 **Learning From Failure**

52 Our experiences with the Home Health systems underscored the  
53 pitfalls of computational approaches to emotion discussed in the  
54 first part of this essay. Developing systems that volunteers  
55  
56  
57  
58  
59  
60

1  
2  
3 perceived as accurate turned out to be very difficult, and we had  
4 difficulty demonstrating a 'sweet spot' between outputs perceived  
5 as redundant or wrong. The distillation of data from multiple  
6 sensors to simple inferences about wellbeing did not seem  
7 compelling to our volunteers, and whatever interest they had in  
8 reflecting on the emotional wellbeing of their households was  
9 undercut by concerns about surveillance and privacy. To be sure,  
10 there were differences between the two volunteer households,  
11 which highlights the fact that people are half the equation for  
12 systems such as the Home Health Monitor. Nonetheless, it is  
13 impossible to avoid the conclusion that overall, the systems  
14 simply failed to realise the promise of emotional computing.

15  
16  
17  
18  
19  
20  
21  
22 None of this serves as strong evidence against a programme of  
23 design based on computational approaches to emotion. Nor do the  
24 issues about such an approach that I raise in the first part of  
25 this essay seem resolvable by reason alone. As a designer, I  
26 believe that the potential of developing systems based on  
27 computational approaches to emotion will be proven by example,  
28 not argument, and admit that others may succeed where we have  
29 failed. Nonetheless, both the arguments I raise and the  
30 experiences I discuss may serve as resources in personal,  
31 pragmatic judgments about the likely fruitfulness of a programme  
32 of design based on using computational approaches to emotion.  
33 Based on both, I suggest that designing to reflect emotions as  
34 part of the complexity of lived experience is both more  
35 tractable, and leads to richer and more engaging results, than a  
36 computational approach to emotion in particular or a focus on  
37 emotion, per se, in general.

#### 38 39 40 41 42 43 44 45 **THE DRIFT TABLE**

46  
47 Let me end this discussion with an example from my studio's  
48 practice that may serve to illustrate what it means to design for  
49 experiences that are emotional without putting emotion at the  
50 centre of design, and also give a sense of the practice that  
51 serves as context for these remarks. Several years ago, as part  
52 of a project on designing devices for the home, we developed a  
53 prototype called the Drift Table (Gaver, Bowers et al. 2007). The  
54 Drift Table is a fairly small (1m<sup>2</sup>) wheeled table with a circular  
55  
56  
57  
58  
59  
60

1  
2  
3 viewport in the centre of its top surface that shows slowly  
4 scrolling aerial photography (Figure 4). The impression is of  
5 looking through a window as one slowly floats several hundred  
6 metres above the ground. Four load sensors measure the centre of  
7 gravity of weights distributed on the table's surface, so that  
8 shifting weights towards one side of the table causes it to drift  
9 in that direction, adding weights causes it to go lower and  
10 faster, and removing them causes it to drift slowly and randomly  
11 at a relatively high altitude. A small display shows the  
12 current location, and a tiny reset button allows the view to be  
13 switched back to the home location, but other than that, and  
14 apart from a hidden on/off switch, no other controls are given.  
15 One can simply shift weights around the table to explore the vast  
16 landscape - about a terabyte of high-resolution photography,  
17 covering all of England and Wales - to which it gives access.  
18  
19  
20  
21  
22  
23  
24

25  
26 *=== insert Figure 4 about here ===*  
27

28  
29 The Drift Table was designed to explore ludic engagement within  
30 the home, where 'ludic' refers to playful, self-motivated  
31 exploration based on curiosity and whim (Gaver, 2009). We were  
32 interested in this as an alternative to more typical utilitarian  
33 applications of technology within the home. Thus the Drift Table  
34 was not conceived to solve any problems or pursue any tasks, but  
35 simply to offer an engaging situation that people could explore  
36 for themselves. We had some ideas of how people might use it, of  
37 course, but we didn't really know. So, as with many of our  
38 prototypes, we loaned it to a volunteer household to live with  
39 over several months, and observed what they did with it using a  
40 combination of ethnographic observation, unstructured interviews,  
41 and documentary film.  
42  
43  
44  
45  
46  
47

48  
49 What we found, briefly, is that some of the volunteers engaged  
50 with the Drift Table intensely throughout their tenure, while  
51 others lost interest either because they did not find electronic  
52 devices appealing in principle, or because it was insufficiently  
53 interactive to maintain their interest. Those who did engage with  
54 the table worked with it far more intently than we had ever  
55 imagined. They routinely set off on journeys of several hundreds  
56  
57  
58  
59  
60

1  
2  
3 of miles, despite the fact that these would take hours and  
4 involve difficult challenges of navigation. They flew to their  
5 old hometowns to view remembered landmarks, visited friends'  
6 neighbourhoods to spot features they could drop into  
7 conversation, explored areas they'd heard about on the news, and  
8 so on. One of the household members worked at home, and reported  
9 taking breaks to readjust the Table's course as a routine  
10 activity during the day. Several of the enthusiasts reported  
11 gathering around the table in the evening, when it would serve as  
12 an alternative to television as a focus of activity and  
13 conversation. They compared it to other things they enjoyed,  
14 ranging from a late-night transmission of satellite imagery to a  
15 particularly interesting airplane journey. In short, the Table  
16 engaged them over time and in many ways.

17  
18  
19  
20  
21  
22  
23  
24 Integral to the volunteer's experience with the Drift Table was  
25 the emotions that arose during its use. These ranged from  
26 excitement and anticipation during its original deployment, to  
27 feelings of disappointment and frustration when its limitations  
28 became known, to a sense of fascination, intrigue and perhaps a  
29 sense of pride as they learned to value it in spite of - even  
30 because of - its constraints. Their emotions were aroused not  
31 only by the device itself, but also by the landscape to which it  
32 gave access. There were moments of delight in seeing hidden  
33 features of their local neighbourhood, disgust at miles of urban  
34 sprawl, nostalgia for a childhood home. They expressed these  
35 emotions during our conversations with them, and in the  
36 documentary video we commissioned from an independent filmmaker  
37 as a way of finding new perspectives on the volunteers' life with  
38 the prototype.

39  
40  
41  
42  
43  
44  
45  
46  
47 The volunteers' emotional reactions to the Drift Table were only  
48 one aspect of their overall experience with it, however. They  
49 also appreciated it conceptually as a technological device, as a  
50 potential domestic product, and as a device offering certain  
51 opportunities and challenges. They valued it aesthetically for  
52 its design and interaction (although they felt the wheels let it  
53 down a bit). The access it gave to the countryside was both  
54 personally compelling and intellectually interesting. It also  
55  
56  
57  
58  
59  
60

1  
2  
3 served to facilitate sociality within the home (when they  
4 gathered around to discuss the view) and occasionally to thwart  
5 it (when they found that only a few people could look through the  
6 viewport at once, or argued over the inadvertent misplacement of  
7 objects on the surface).  
8  
9

10  
11 In sum, the Drift Table was compelling to our volunteers. Part of  
12 its impact was emotional, but this arose and found its meaning in  
13 relation to aesthetic, conceptual, functional and social  
14 appreciation. Moreover, these dimensions of appreciation were  
15 integrated in our volunteers' lived experiences to the point that  
16 distinguishing amongst them is somewhat misleading and unhelpful  
17 (McCarthy & Wright, 2004). Mirroring this, we had not  
18 distinguished these facets of experience in designing the Drift  
19 Table. We did not seek to design an emotionally compelling  
20 experience any more (or less) than we set out to design a  
21 conceptually resonant one. Or rather, we set out to design for  
22 all of these things, but not as analytically articulated  
23 desiderata but as integrated and embodied in a design that we  
24 tried to make as rich and compelling as possible.  
25  
26  
27  
28  
29  
30  
31  
32

33 This is the crux of my argument: that rather than singling out  
34 emotion as an object of attention, and working to explicitly  
35 recognise and represent it, it is more fruitful to recognise  
36 emotion as an emergent aspect of experiences that are situated  
37 and multi-layered. This leads to a design-led research approach  
38 that focuses on crafting the appearance and interactivity of  
39 specific designs open to ludic engagement on the part of their  
40 users. If done well, the designs will both embody understandings  
41 of emotion, aesthetics, sociality and culture and lead to new  
42 insights as well<sup>iii</sup>. Emotion may be an important facet of the  
43 understandings and insights successful designs rely on and  
44 produce, but it is not the only one, and not always the most  
45 important.  
46  
47  
48  
49  
50  
51

## 52 **ACKNOWLEDGEMENTS**

53 The research reported here was pursued in collaboration with Andy  
54 Boucher, John Bowers, Nadine Jarvis and Tobie Kerridge of the  
55 Interaction Research Studio, and with Phoebe Sengers and Jofish  
56  
57  
58  
59  
60

Gaver - Designing for Emotion

20

1  
2  
3 Kaye from Cornell University, with support from Intel Corporation  
4 and the Equator Interdisciplinary Research Collaboration. Thanks  
5 to John Bowers for introducing the distinction between  
6 information narrowing and widening, to him, Kia Höök, Anne  
7 Schlottmann and Phoebe Singers for comments on an earlier draft  
8 of this paper, and to the volunteers who opened their homes to  
9 our systems and us.  
10  
11  
12

#### 13 14 15 **REFERENCES**

16 Arthur, C. (2009), It's always best to keep your own lie detector  
17 turned on, *The Guardian*, 21 April 2009),  
18

19  
20 Boehner, K., DePaula, R., Dourish, P., and Sengers, P. (2005).  
21 Affect: from information to interaction. *Proceedings of the*  
22 *Conference on Critical Computing*. ACM, pp. 59-68.  
23  
24

25  
26 Boehner, K., Sengers, P., and Warner, S. (2008). Interfaces with  
27 the ineffable: Meeting aesthetic experience on its own terms. *ACM*  
28 *Trans. Comput.-Hum. Interact.* 15, 3, pp. 1-29.  
29

30  
31 Cowie, R. (this volume). Perceiving emotion: towards a realistic  
32 understanding of the task. To appear in this volume.  
33  
34

35  
36 Gaver, W. (1991). Technology affordances. In *Proceedings of*  
37 *CHI'91*, pp. 79-84  
38

39  
40 Gaver, W., (2009). Designing for Homo Ludens, Still. In  
41 *(Re)searching the Digital Bauhaus*. Binder, T., Löwgren, J., and  
42 Malmborg, L. (eds.). London: Springer, pp. 163-178.  
43  
44

45  
46 Gaver, W., Bowers, J., Boucher, A., Law, A., and Pennington, S.  
47 (2007). Electronic furniture for the Curious Home: Assessing  
48 ludic designs in the field. *International Journal of Human-*  
49 *Computer Interaction*, 22(1-2): 119-152.  
50  
51

52  
53 Gaver, W., Bowers, J., Kerridge, T., Boucher, A., and Jarvis, N.  
54 (2009). Anatomy of a failure: How we knew when our design went  
55 wrong, and what we learned from it. *Proceedings of CHI '09*.  
56  
57  
58  
59  
60

Gaver - Designing for Emotion

21

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.

Gaver - Designing for Emotion

22

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.

For Review Only

**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.

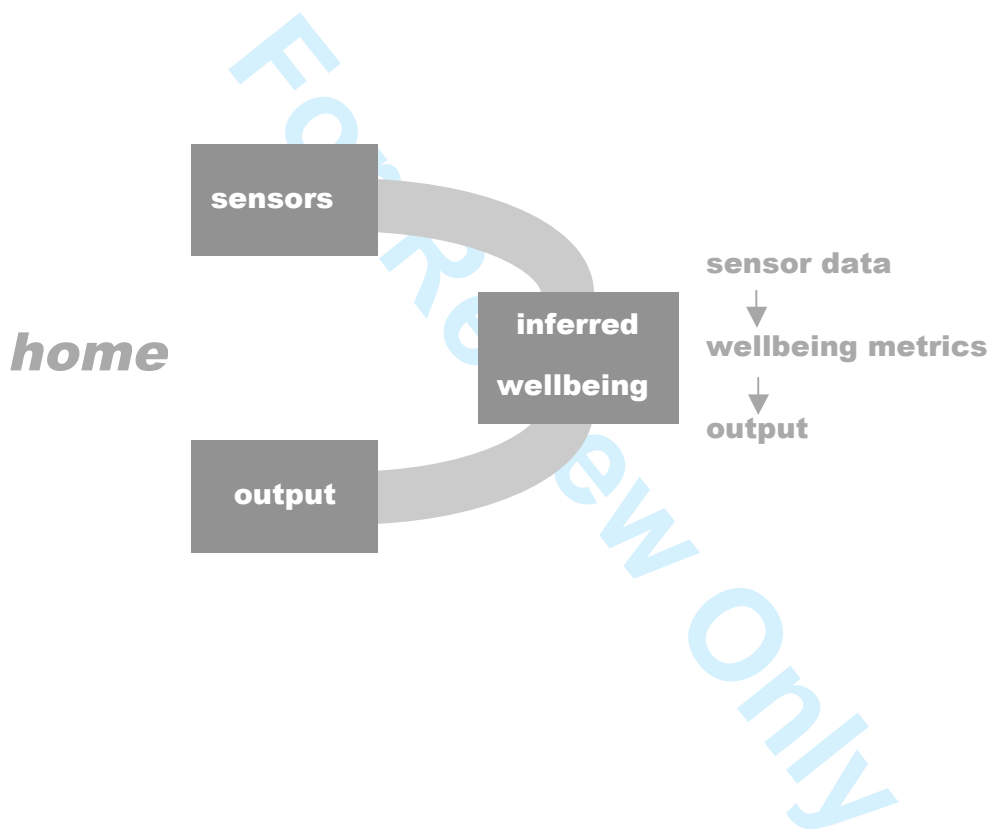
**FOOTNOTES**

<sup>i</sup> 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.

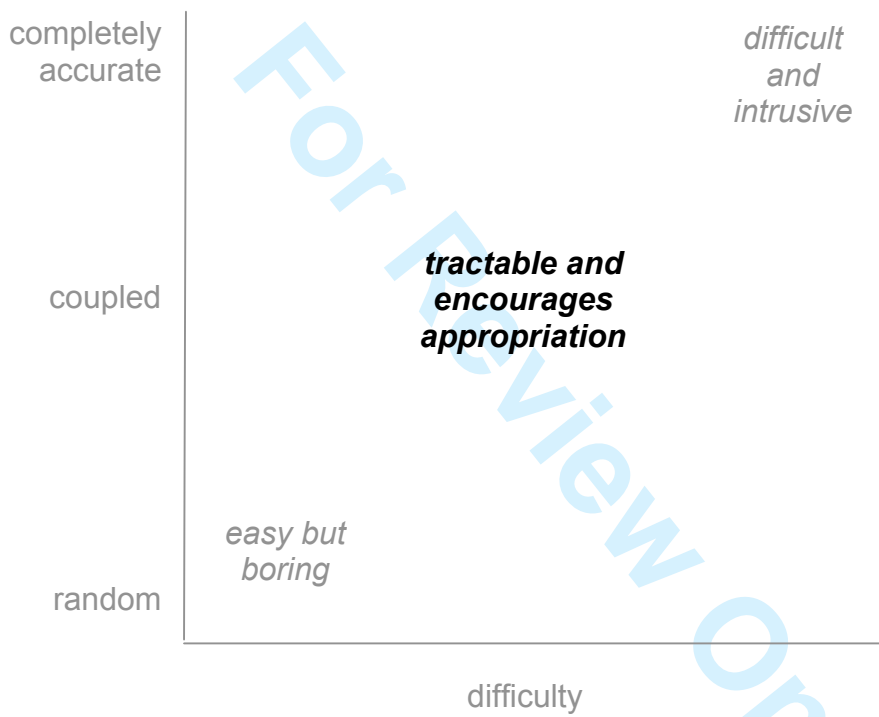
<sup>ii</sup> The advent of search engines and similar interfaces may be seen as a massive revival of command-line languages, but whether they are to be viewed as agents or as machines is still an open question (they certainly don't seem emotional).

<sup>iii</sup> 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.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



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)

Only