Push Me, Shove Me and I Show You How You Feel

RECOGNISING MOOD FROM EMOTIONALLY RICH INTERACTION

The mood or emotional state you are in colours the way you interact with people and systems. Future interactive systems need to recognise emotional aspects in order to be truly adaptive. We designed an alarm clock, which elicits rich expressive behaviour and demonstrated that it is able to read your mood from the way you set it. We validated film clips, used them to induce moods after which participants had to set the alarm clock. From the dynamic setting behaviour we inferred parameters from which we calculated equations to identify the mood. The results illustrate the importance of a tight coupling between action and appearance in interaction design, through freedom of interaction and matching inherent feedback.

KEYWORDS
Interaction design, emotional state, inherent feedback, research through design

INTRODUCTION

As products become more intelligent and adaptive, communication with them becomes more complex. For human-human communication expressing emotion is essential. People express emotion through behaviour. In human-product communication people express their emotion as well, e.g., by shouting at the monitor, shoving a chair away, or encouraging the printer. Yet this behaviour does not enhance communication between user and product at all. On the contrary, if we actually express our negative emotion we might wreck the product. The central question of this paper is: "Can products recognise emotion from a person's behaviour in order to improve the person-product interaction)?"

There are several ways to attack the question of recognising emotions. The affective computing group at MIT MediaLab [8] focuses on physiological information, e.g., blood pressure, skin conductivity and heart rate, to recognise emotion. Although this information is very reliable [1, 8], it does not allow a person to express her emotion to the system. After all, you do not express your emotions through physiology. So while the system does recognise emotions through physiology, it denies the user's urge to express them.

Another way to detect emotion is by voice and face recognition [8]. One of the problems with this approach is that it often seamlessly leads to, what Verplank refers to as "anthropomorphic distraction" (cited in [6], p. 46). In human-human interaction an important advantage for communicating emotions is the fact that we all have the 'same' body, and similar skills and senses. Consequently we expect the other person to have the same basic abilities to recognise and express emotions. However, as products have different bodies, we hesitate which skills and senses to attribute to these products. So when a product can recognise a person's emotions from voice or facial expression the 'body' of the product should express this ability. The solution, however, is not to make this 'body' more anthropomorphic. As people already attribute humanlike characteristics to non-humanlike products [9], making them even more anthropomorphic would elicit expectations that can not be met. Products (like Microsoft Word's paperclip)

Stephan Wensveen s.a.g.wensveen@io.tudelft.nl
Kees Overbeeke c.j.overbeeke@io.tudelft.nl
Tom Djajadiningrat j.p.djajadiningrat@io.tudelft.nl

ID-StudioLab
Delft University of Technology
Landbergstraat 15, NL2628 CE
Delft, the Netherlands

PROTOTYPE DESCRIPTION

The prototype of the clock consists of two displays and twelve sliders. The front display shows the current time, the central display shows the alarm time.
When the sliders are slid from the starting situation (a) towards the central display the alarm time appears in this display (b). With the first displacement of a slider, time is added to the current time to make up the alarm time. With each successive displace-ment, more time is added to (moving towards the centre: c) or subtracted from (moving towards the edge: d) the alarm time. Each slider ranges from 0 to 60 minutes. Upon reaching the preferred wake up time the central display is pressed (e) and the alarm is set (f).

Figure 1: The alarm clock prototype

should not pretend to have the same abilities as humans when they clearly don't.

We advocate a third way that both allows for expression of emotion and avoids anthropomorphic design. This product design-driven approach takes the interaction with the product as the starting point for the detection of emotion. While you interact with the product to communicate 'factual' information, the product senses your emotions from the way you handle it.

To make this possible the product should meet the following three conditions:

1. Elicit rich emotional behaviour while the user communicates 'factual' information.

2. Have the ability to recognise this emotional behaviour.

3. Reflect and understand the expressed emotion.

The consequence of this stance is that research on emotion recognition starts with the design of product interaction. This is what we call 'research through design'. We thus start with a design: an alarm clock. An alarm clock is a perfect example of a product that should adapt to a diversity of emotional experiences. It is also a product with a simple functionality yet it has all the features of the of the current interface malaise [2].

We first describe the prototype of this alarm clock in Figure 1. Then we report experiments on the recognition of emotion through interaction with the product.

Design
How did we design this alarm clock to meet the conditions?

Condition 1: Elicit rich emotional behaviour while the user communicates 'factual' information.
Expressing emotions presupposes freedom of expression, and thus freedom in interaction. Our alarm clock design invites the user to express herself by offering a myriad of ways of setting the "factual" information, i.e., the wake up time. People can choose to set it by displacing as many sliders as they can grasp (Figure 1b) or by sliding one slider at a time (Figure 1c/1d). This behavioural freedom affords emotions to influence the behaviour. This freedom is

further enhanced by the fact that sliders can go back and forth and actions are easily reversed.

Condition 2: Have the ability to recognise this behaviour.
This is what this paper is essentially about and will thus be explained in the experiments.

Condition 3a: Reflect the expressed emotion.
The basic assumption of this research is that different emotions have to leave different behavioural traces. A slap or a caress leaves a different trace on a face. Likewise, setting the time in a different mood leaves a different trace on the alarm clock. We defined mood as a temporary state of mind. A mood can be positive or negative (valence) and calm or excited (arousal).

Whereas the central display shows the wake up time (factual information), the successive patterns of the sliders reflect the influence the emotion had on the behaviour leading to this wake up time. We call this trace inherent feedback. It is information provided as a natural consequence of making an action. It arises from the movement itself [5]. This trace is essentially dynamic. In Step 3 (see below) we define two groups of parameters: action and pattern. The action parameters describe how the participants actually move the sliders. The pattern parameters describe the result of the action. So, over the actions (i.e., an uninterrupted displacement of one slider) we build up a history of the interaction that is reflected in the successive patterns.

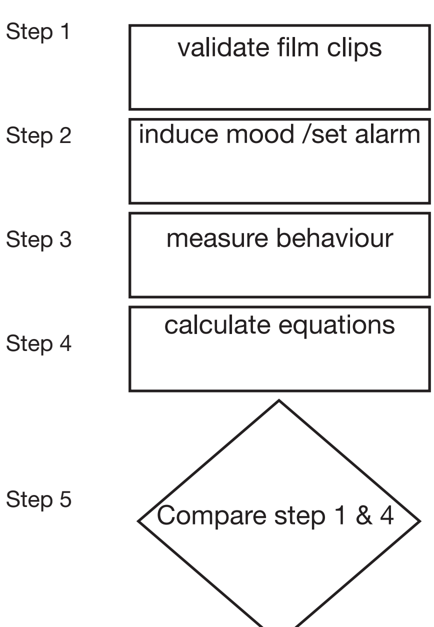


Figure 2: Five steps to validation

Condition 3b: Understand the expressed emotion.
The next morning the alarm wakes you with a sound. The choice of this sound makes clear it understood you.

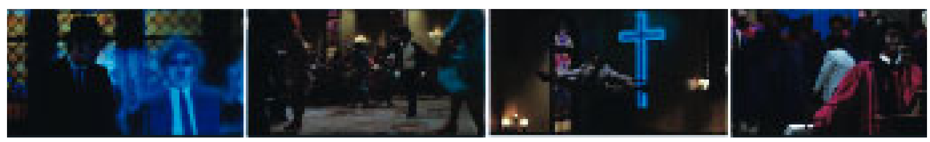
THE ENVISAGED SYSTEM

We envisage the final system to operate as follows. In the evening the wake-up time is set (factual information). This is done differently when in a different mood (mood information). Based on both factual and mood information the alarm clock makes a decision for an alarm sound. The next morning the person wakes up to this sound and silences it by gently touching or hitting the snooze button. This behaviour expresses the person's emotions about the appropriateness of the decision. From this behaviour the system gets feedback on its decisions and can learn and adapt accordingly (see [10] for a full description of the system).

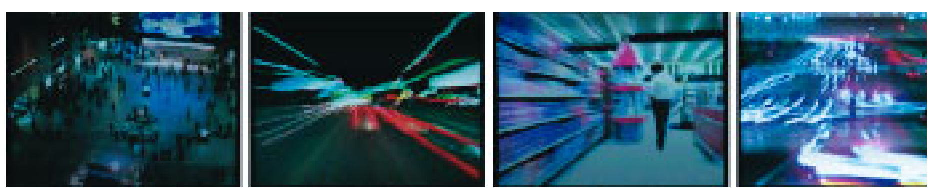
THE VALIDATION

We now turn our attention to the recognition of emotion by the alarm clock. The general reasoning is depicted in Figure 2.
In order to test if the alarm clock can do the trick we took the following steps:

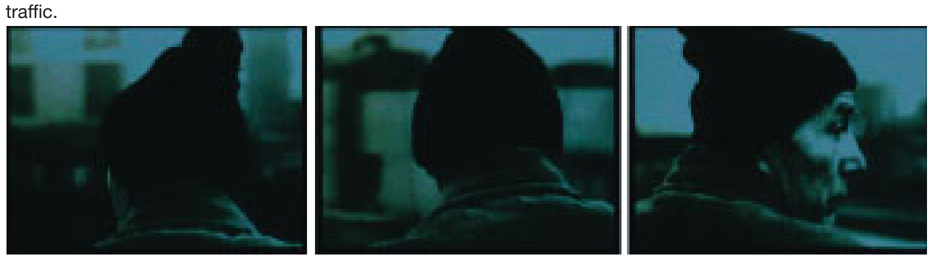
1. Validate film clips to induce mood (valence and arousal) and urgency.
2. Induce a mood and urgency while participants set the alarm clock.
3. Measuring the behaviour of participants using behavioural parameters.
4. Calculate equations for mood and urgency from these parameters.



High arousal and positive valence: A clip (duration: 2:55) from the Blues Brothers. It depicts the two main characters attending a church service where the reverend, played by James Brown, and the Rev. James Cleveland Choir perform 'the Old Landmark'.



High arousal and negative valence: Film clips taken from the movie Koyaanisqatsi . These fragments, with the music composed by Philip Glass, were edited into a climaxing sequence by the first author of this article. The total clip (duration: 2:31) consists of several high-speed video fragments of people in train stations and cars in traffic.



Low arousal and negative valence: A black and white clip from Stalker (duration: 2:06) featuring close up shots from behind two persons sitting on a train wagon. The repeating thumping sound of the train over the railway sleepers accompanied the clip.

Low arousal and positive valence: A scene from Easy Rider (duration: 2:38) was used. It depicts the three main characters on a North American desert highway riding their motorcycles into the sunset. The underlying score is the song "The weight" played by 'The Band'.
Figure 3: The film clips targeting the valence/arousal space

5. Identify the mood by comparing the calculated values (Step 4) with the validation values (Step 1). We now describe these five steps in detail.

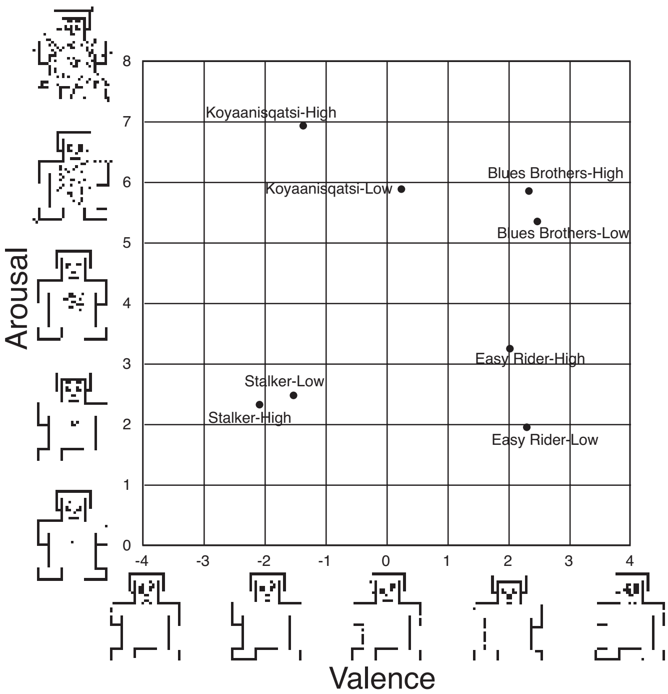


Figure 4: Plot of mean SAM scores (n=13). High/low means high/low urgency.

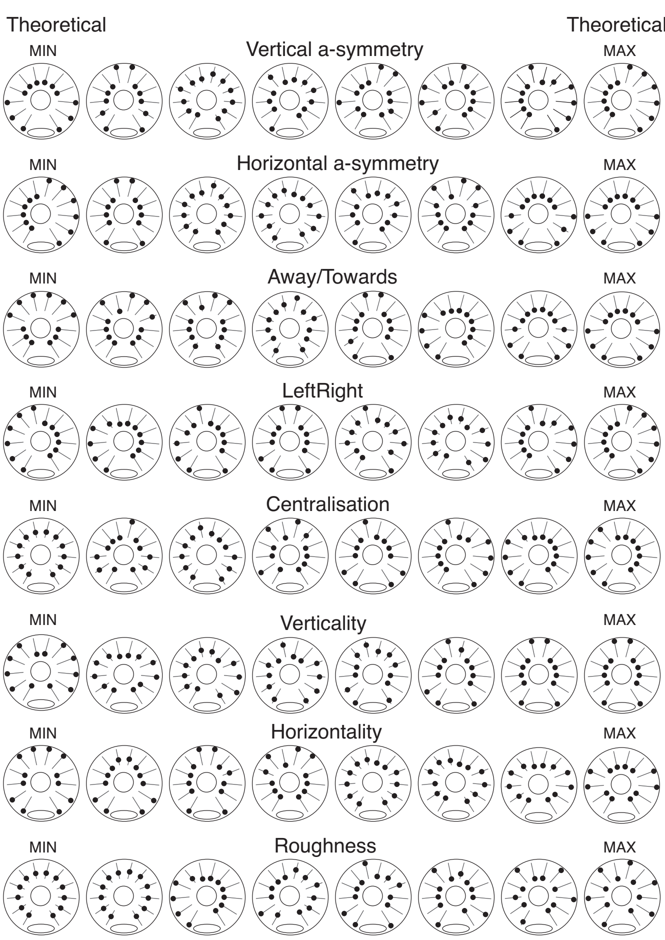


Figure 5: A visual description of the pattern parameters.

STEP 1: VALIDATING MOOD INDUCTION PROCEDURE

Aim: In order to control for the user's mood we used a Mood Induction Procedure (MIP). Westermann, Spies, Stahl and Hesse [11] showed that an MIP using film clips and appropriate instructions is the most effective way to induce both positive and negative mood states. In the instruction participants are explicitly asked to empathize with the situation shown. This step has a dual aim. The first aim is to find film clips that target the dimensional quadrants of the circumplex model. This model organizes emotions in a two dimensional valence/arousal space. Larsen and Diener [4] state that the circumplex model of emotion… "suggests a clear structure for the effects emotion will have on behaviour (…) and thus has large heuristic value. Similarly, a circumplex model of emotion, by accounting for a majority of the variance in affect measures, suggests a simple yet powerful way to organise facts about emotion." We validated beforehand the MIP's power according to the procedure below.

The second aim of Step 1 is to find the individual values for valence, arousal and urgency as attributed by each participant. These scores are needed in Step 4 in order to be able to correlate an individual's mood with his or her setting behaviour. Therefore we report here the results of the 13 participants of the actual experiment (Step 2).

The following procedure was used:

Participants: A total of 13 undergraduates (6 men, 7 women) of the School of Industrial Design Engineering participated in the experiment.

Stimuli: Film experts suggested film clips, described in Figure 2, targeting the dimensional quadrants of the circumplex model.

Procedure: In each session four participants sat in a darkened room four meters away from a four-meter wide projection screen. Participants were instructed to try to empathise with the depicted situation after which they had to fill in an evaluation form. The evaluation form consisted of two questions relating to the arousal and the valence of their mood using the 'Self Assessment Model' (SAM) [3] (see also Figure 4). Arousal and valence were evaluated using a 9-point scale from 0 to 8 and from -4 to +4, respectively.

To induce urgency each clip was preceded with a text in the form of a 30-second motion graphic. This read either an urgent message: "Tomorrow you have to go to the airport. Catch a plane. You cannot oversleep! You have to get up at eight o'clock. 8:00" or a non-urgent message: "*Tomorrow… no obligations. A day off. You still want to do something nice. Set your alarm around… 8 o'clock.*" The messages were also superimposed during the last 30 seconds of each film clip. The urgency was evaluated on the same form using a 9-point scale ranging between "*In this situation it is… totally unimportant that an alarm clock wakes me up.*" to "*…very important that an alarm clock wakes me up*". A practice situation familiarized the participants with the evaluation form. Then the eight situations (each of the four clips had an urgent and non-urgent condition) were shown in a random order. To neutralize the effect of the preceding situation a one-minute relaxation clip was shown, depicting moving clouds with the sound of crashing waves.

The mean values for each situation are placed in a valence-arousal plot (Figure 4). The eight situations discriminate well between valence and arousal level. The mean urgency level ranges from 2.46 to 7.54, as expected.

STEP 2: INDUCING A MOOD & SETTING THE ALARM
Aim: The aim of this step is to have the participants set the alarm clock while being influenced by different film clips.
Participants: The same participants as in Step 1.

Stimuli: The same eight situations of Step 1 were used to induce mood and urgency after which participants were asked to set the alarm clock prototype.

Procedure: This experiment again was conducted in a darkened room where the individual participant sat behind a table with the alarm clock at a distance of 4 meters from a 4-meter wide projection screen. The experimenter shortly introduced the operation of the alarm clock after which the participant was allowed to try out the possibilities of the sliders and get a feel of the alarm clock.

At the end of each movie clip an on-screen message asked the participant to set the alarm clock for eight o'clock. After the setting the 1-minute relaxation clip was shown.

To be able to measure the setting behaviour the alarm clock produced electronic read-outs from each slider with a sample frequency of 100 Hertz. These readouts are the identification of a slider, the starting time of a slider action, the duration of this action, the value in minutes (0-60) at the start of this action and the value at the end of that action. We use 'minutes' as unit of measure because the distance a slider can travel is expressed on a scale ranging between 0 to 60 minutes. One action is defined here as the uninterrupted displacement of one slider.

STEP 3: BEHAVIOURAL PARAMETERS
To be able to categorise behaviour we defined two groups of parameters. The parameters that define the action itself and those that define the end result of an action, the slider pattern. All parameters are computed from the electronic readouts of each slider.

Table 1: Individual and mean success percentages for identifying valence, arousal, urgency and situations.

	Valence	Arousal	Urgency	Situations
Chance level	50.0	50.0	50.0	12.5
pp 1	100	87.5	75.0	62.5
pp 2	75.0	62.5	87.5	37.5
pp 3	87.5	100	87.5	75.0
pp 4	75.0	75.0	75.0	25.0
pp 5	62.5	50.0	100	25.0
pp 6	100	87.5	75.0	75.0
pp 7	75.0	100	100	75.0
pp 8	87.5	100	75.0	62.5
pp 9	100	100	87.5	87.5
pp 10	87.5	100	75.0	62.5
pp 11	62.5	75.0	100	50.0
pp 12	75.0	75.0	100	50.0
pp 13	75.0	100	87.5	75.0
mean	81.8	85.6	86.5	58.7

Action parameters:
Action no is the serial number of an action.

Displace is the value of displacement of each slider. A positive value is a displacement towards the center of the alarm clock whereas a negative displacement is a movement away from the center.

Duration is the time a slider slides during one action.

Waiting is the amount of waiting time during which none of the sliders is touched after one action.

Speed is the average speed of a slider during one action.

Sliders is the number of simultaneously used sliders during one action, calculated for every action.

Pattern parameters:
The parameters that define the slider pattern are described here generally as a detailed elaboration is beyond the scope of this paper. Figure 5 offers a visual description of the parameters.

'**Vertical Symmetry**': To determine a value for the vertical symmetry of the pattern we compared the individual value (between 0 and 60 minutes) of each slider with the value of its, in a vertical line mirrored counterpart.

'**Horizontal Symmetry**': The same approach was used to determine the value for the horizontal symmetry. The values of sliders, which mirrored in a horizontal axis were compared.

'**Away/Towards**': To determine if sliders were pushed away from or towards the body we compared the sum of values of the sliders from the lower half of the alarm with that of the upper half.

'**LeftRight**': The horizontal deviance from the centre is a measure for the displacement of sliders either to the right ('leftright' is positive) or to the left side of the alarm clock ('leftright' is negative).

'**Centralisation**': To see to what extent sliders are centralised around the snooze button we determined the distance between the centre of gravity and the actual centre of the alarm clock.

'**Verticality**': To describe the verticality of the pattern we compared the sum of values of the sliders that contribute to a vertical pattern with the sum of values of the other sliders.

'**Horizontal**': For a horizontal dimension we compared the sum of values of the sliders that contribute to a horizontal pattern with the sum of values of the other sliders.

'**Roughness**': To determine the smoothness or the roughness of the pattern we compared the values of successive sliders with each other.

STEP 4: CALCULATING EQUATIONS
To calculate values for valence, arousal and urgency from the parameters we set up regression equations. We first calculated the standardized values of the behavioural parameters for each participant over all actions and situations. These standardized

values were then used as the independent variables in a 'within-subjects' regression analysis with the SAM-scores for arousal and valence, and the rating for urgency as separate dependant variables. We thus calculate the regression for each individual from her own parameters. Per participant we drew up the linear regression equation, separately for valence, arousal and urgency, including only the coefficients of the statistically significant (a < 0.05) behavioural parameters.

STEP 5: IDENTIFYING THE MOOD

We now take the final step. For each action a valence, arousal and urgency value is calculated using the equations of step 4. The means of these values were then calculated over all actions belonging to each one of the eight settings of the alarm time. These means were compared with the values for arousal, valence and urgency as attributed to the situation (combination of film clip and urgency message) by the participant in step 1.

If the calculated value and the attributed value were both higher or both lower than the middle of the scales for arousal, valence and urgency (4, 0 and 4 respectively) we counted this as a correct identification. If all three were correct the entire situation was correctly identified. Table 1 shows the success rates for each participant. The success rate for the correct identification of arousal ranged between 50% and 100%, with a mean of 86% over all participants. For valence this rate ranged between 62.5% and 100% with a mean of 82%. The success rate for urgency ranged between 75 and 100% with a mean of 87% correct predictions. The success rate for the prediction of the total situation ranged between 37.5% and 87.5%. 58.7% of all situations was calculated correctly.

TO CONCLUDE

A success rate of almost 60% corresponds with the success rate of the e-mouse (67%) [1]. The success rate of recognising valence, arousal and urgency (80-90%) separately corresponds well with the success rate of physiological data as reported by Picard (cited in [7], p. 220). Of course we only identify eight categories for the situation, and only two for valence, arousal and urgency.

Now that we have shown the possibility of recognising emotions from behaviour on an individual level we look at more general principles of how emotional state colours behaviour. Therefore we looked at the significant (a < 0.05) correlations between the attributed values for valence, arousal and urgency and the behavioural parameters. Behind each general principle we give the number of persons out of the 13 participants whose data support the principle and the number of persons that support the opposite. For the remaining persons the parameter has no significant correlation with the mood values.

When we look at valence, we see that when people have been induced by a more positive situation, they 'show' more actions per setting (7, 0), more centrally balanced patterns (7, 0), more horizontally symmetrical patterns (7, 1), more vertically symmetrical patterns (6, 2), slower actions (5, 2), shorter displacements (4, 0) and shorter waiting time between actions (4, 0).

When participants are more aroused, they 'show' more sliding actions towards them (8, 2), less vertically symmetrical patterns (8, 3), less horizontally shaped patterns (6, 3) and faster actions (5, 0).

When participants feel more urgency in a situation, they 'show' smoother patterns (8, 1), less vertically symmetrical patterns (8, 4), more centrally balanced patterns (7, 0), more sliding actions towards them (7, 2), faster actions (4, 0) and shorter actions (4, 0).

The design of the alarm clock thus illustrates the importance of a tight coupling between action and appearance in interaction design. It distinguishes itself from current electronic products through traces and inherent feedback. In current electronic products only the final setting of the controls is taken into account. In this alarm clock the intermediate stages are also considered, that is, the history of the final setting is also used to determine the user's needs.

With inherent feedback we mean that the feedback through the appearance of the product is a natural consequence of the user's actions. Because of the inherent feedback the traces become visible, are made explicit for the user and guide his behaviour.

For example, when using both hands on the sliders in an even and balanced way the resulting pattern is symmetrical and smooth. The way this pattern looks will push the user to either heighten the symmetry and smoothness or disrupt them depending on how she feels. Traces and inherent feedback thus work in synergy. Without inherent feedback using traces is meaningless, as the product cannot guide the user's behaviour: the trace is invisible and cannot invite the user to act in a emotionally rich manner.

From our product design perspective, the appearance of interactive products can no longer be considered as arbitrary. Appearance and interaction need to be designed concurrently.

ACKNOWLEDGEMENTS

The authors like to thank Daniel Brundl and Rob Luxen for the realisation of the working prototype.

REFERENCES

1. Ark, W., Dreyer, D.C. and Lu, D.J. (1999). The emotion mouse. Proceedings of the HCI International Conference.
2. Cooper, A. (1999) The inmates are running the asylum. Indianapolis, SAMS McMillan.
3. Lang P.J. (1985). The Cognitive Psycho-physiology of Emotion: Anxiety and the Anxiety Disorders. Hillsdale NJ: Lawrence Erlbaum.
4. Larsen, R. J., & Diener, E. (1992). Problems and promises with the circumplex model of emotion. Review of Personality and Social Psychology, 13, 25-59.
5. Laurillard D. (1993) Rethinking university teaching: A framework for the effective use of educational technology. Routledge London.
6. McDonald, N. (2001). Action, interaction, reaction. Blueprint, August 2001. pp 44-50.
7. Paiva, A. (ed.) (2000) Affective interactions: towards a new generation of computer interfaces. Berlin, Springer.
8. Picard, R.W. (1997). Affective computing. Cambridge: MIT Press.
9. Reeves, B. and C. Nass (1996). The Media Equation: how people treat computers, televisions and new media like real people and places. Cambridge, Cambridge University Press.
10. Wensveen, S., & Overbeeke, K. (2001) Adapting through behaviour: What my alarm clock should know, do and feel. Proceedings of the International Conference on Affective Human Factors Design. Singapore, June 27-29, pp. 242-247
11. Westermann R. Spies K. Stahl G. K. & Hesse F. W. (1996). Relative effectiveness and validity of mood induction procedures: a meta-analysis. European Journal of Social Psychology 26 pp. 557-580.