HUMAINE

D9b

Preliminary plans for exemplars:

Usability

Kristina Höök, Katherine Isbister, Fiorella de Rosis, Jarmo Laaksolahti, Brigitte Krenn and WP9 members

Version 1.0

Date: 28 May 2004

IST project contract no.	507422
Project title	**HUMAINE** **Human-Machine Interaction Network on Emotions**
Contractual date of delivery	*month 5*
Actual date of delivery	*28 May 2004*
Deliverable number	*D9b*
Deliverable title	Preliminary plans for exemplars: Interaction
Type	Public Report
Number of pages	40
WP contributing to the deliverable	WP9
Task responsible	KTH
Author(s)	Kristina Höök, Katherine Isbister, Fiorella de Rosis, Jarmo Laaksolahti, Brigitte Krenn
EC Project Officer	Philippe Gelin

Address of author:

Kristina Höök

Department of Computer and Systems Sciences, Forum 100, 164 40 Kista, Sweden

Table of Contents

1. The status of this report

This report is the review that defines the starting point of the process for work package WP9, whose brief title is ‘Usability’.

2. Thematic definition of this workpackage

While the rest of HUMAINE will produce knowledge and theories of many different aspects of emotional systems (ECA behaviours, emotion recognition, interaction principles, exemplar databases, etc.) the goal of WP9 is to take all of that knowledge as input to projects that aim to produce functioning emotional systems – end-user applications. A functioning emotional system is one where end-users get involved, when their experience during the interaction is meaningful and hopefully close to the one intended by the designer/developer.

Our strategy for working within an end-user application framework to evolve the usability of emotional systems can be divided into two focus areas:

1. *Ends*: Determining qualities and criteria that demark emotional systems as usable, evoking the desired experiences for users.

2. *Means*: Forging process—design methods and project goals—that will steer a project towards producing a 'successful' application.

Regarding *Ends*, we do not believe there are simple, replicable correlations between system properties and all end users' experiences. The *experience* of using an emotional system is not a property of the system itself, but rather is something that arises in the interaction between user and system. To quote Sengers and colleagues (2004):

> "Rather than experience as something to be poured into passive users, we argue that users *actively* and *individually* construct meaningful human experiences around technology. They do so through a complex process of interpretation, in which users make sense of the system in the full context of their everyday experience."

Along the same lines, Suchman has criticized the cognitistic reductionist basis for the field of affective computing (Suchman, 2002), reducing human emotional responses to discrete, universal component parts, arising from some underlying state, and offers examples of alternative designs that put emotional interaction at the core but where: "we might differently conceptualise affective encounters at the interface: as irreducibly contingent meetings of particularly situated persons with equally particular, dynamic, and culturally inflected things." As we discuss below, usability of emotional systems and the process of arriving at functioning emotional systems need to be sensitive to how end-users are part in co-constructing the experience and be part of this dynamic and cultural process.

Usability traditionally focuses on goals such as effectiveness, efficiency, safety, utility, learnability, and memorability. These objective usability goals contrast with user experience goals, which cover subjective qualities such as being fun, rewarding, motivating, satisfying, enjoyable, and helpful. Usability goals and user experience goals often stand in complex relationships, involving tradeoffs such as safety vs. fun or efficiency vs. enjoyability (Preece et al., 2002). Introducing emotion thus raises many new dimensions for research on usability to address. It becomes, for instance, a serious issue whether users feel a system is 'sympathique' or morally acceptable, whether it engages them emotionally. Furthermore, emotions address directly inherently adaptive faculties of humans, posing challenges for methods of user studies and artefact design.

Work Package 9 will take a primarily qualitative, situated/contextual approach to measuring *Ends*. A sub-group of Work Package 9 may also focus on evaluating and developing measures aimed at isolating meaningful evaluative variables and testing various components of systems against these variables.

Regarding *Means*, we believe that evaluation of affective systems is vital not just at the end of the design process, but as an integral part of the design process *from the beginning*. Having the ability to bounce early intuitions and design sketches off of real users can make key contributions to the evolution of a truly engaging end application, and may even inform the affective theory that led to the application itself.

Work in the field of HCI has made great strides toward merging design and evaluation of productivity-oriented systems (see the goals section of this document for some relevant citations); we seek to extend and adapt participatory design methods and approaches to the needs and desired outcomes of affective system projects. For example, encouraging nonverbal participation and evaluation, and using physical prototypes and fake systems operated by humans. We believe this holistic and integrated approach will not only lead to better system designs, but may also provide important 'in situ' insights back to affective theory makers.

Thus, it is not the primary intent of WP9 to test computational emotion theories or evaluate how particular parameters for raising an eyebrow in an ECA should work. Instead, the task of WP9 will be to produce evaluation theory, methods, and measures needed to take all that knowledge into the design process and through it, to produce emotional systems that will engage with end-users to create compelling experiences. However, as with Ends, a sub-group of Work Package 9 may decide to pursue a more component-based analysis of affective systems.

3. Review of key concepts in the thematic area

Terms that might need explanations from the human-computer interaction and usability area:

Usability: The process of creating and evaluating a system which delivers a positive user experience.

Evaluation: The process of submitting a design (whether concept, prototype, or finished system) to examination through some form of observed use, towards gathering information that will be helpful for further iteration.

User-centred design: Design of systems which integrally incorporates the needs, context, and insights of future users.

Participatory design: Design which incorporates members of the target user group as part of the design team.

Phenomenology: a 20th-century philosophical movement dedicated to describing the structures of experience as they present themselves to consciousness, without recourse to theory, deduction, or assumptions from other disciplines such as the natural sciences (Microsoft Encarta)

User Experience: The overall set of perceptions and reactions during a system's use, which are co-emergent from the person and the system with which s/he is engaged.

Quantitative Measures: Instruments to collect data in controlled settings, which will be analyzed statistically. (See Ebling and John, 2000, for an interesting discussion of the relative merits of quantitative and qualitative data in usability evaluation).

Qualitative Measures: Methods for collecting data which seek to preserve context and subjectivity, and which lean toward thick description rather than statistical summaries. (See Ebling and John, 2000, for an interesting discussion of the relative merits of quantitative and qualitative data in usability evaluation).

Validity: The extent to which an instrument adequately and accurately captures the theoretical construct it seeks to measure (e.g. does a Likert scale asking how 'usable' a system was capture the theoretical construct of successful user experience).

4. Review of key achievements in the thematic area

Ends:

- Existing quantitative measures of user responses include: Likert scale response sets given in questionnaire form (e.g. those used in the research of Reeves and Nass, 1996), quantitative coding of video-taped data with inter-coder reliability checks (e.g. work like that referenced in Work Package 5), and the use of biometric sensor data (e.g. brain waves, skin conductance, facial muscle contraction, heart rate) to measure affective response to systems (e.g. Hazlett, 2003, and work like that referenced in Work Package 4). These methods usually involve brief, single-episode interactions with prototypes or systems. The use of video coding and bio-sensors reflects the understanding that affective states may be difficult for users to self-evaluate using traditional questionnaire-style measures.

- Existing qualitative measures include: qualitative/ethnographic summaries of use contexts (e.g. Beyer and Holzblatt, 1999); summaries of user preferences for product qualities based on observational settings such as focus groups, and/or user engagement with prototype systems (e.g. Jordan, 2001); thick description of user engagement with prototypes and systems, based upon observation of interactions (e.g. Höök, Sengers, and Andersson, 2003); collations of verbal comments made during open-ended question sessions (Ebling and John, 2000).

Means:

- Classical controlled experiment design has been used with fully functioning systems to isolate and analyze specific system variables which impact the user's affective response, such as the use of flattery or humor (e.g. Reeves and Nass, 1996; Morkes, Kernal, and Nass, 1998); or to test the application of emotion theory to system design (Zhou and Conati, 2003).

- The HCI community has engaged in quantitative evaluation of fully functioning systems through controlled observation of user interaction, typically using questionnaires and occasionally video-coding, to assess user responses (see Helander, 1988 and Preece et al., 1994 for an overview of current techniques). There is also application of post-use surveys (e.g. Mentis and Gay, 2003).

- There is an emerging movement within the HCI community toward more qualitiative evaluation of user affect and engagement (e.g. Höök, Sengers, and Andersson, 2003).

- Participatory design can help shape affectively successful systems through engagement of end users early in the process (see Muller and Kuhn, 1993, for a description of this process and types of interventions).

General Context and Background:

From previous studies on usability of emotional systems, some lessons have been learnt. While we do not want to claim to have an exhaustive literature survey here, we are providing some insights into the kinds of problems that have been encountered. We divide our survey into one part concerning ECAs, since most work so far has focused on ECA evaluation (ECA

Research Summary) and one part on other kinds of interactive emotional systems and strategies (General Affective System Research Summary), for example: development and assessment of affective objects, music, gestures; recognising affect from biosensors or other tangible interaction devices; recognition of affect from speech input; and affective interactive narratives.

ECA Research Summary

The following usage of interface agents, all well represented in literature and existing systems, are possible domains for agents that possess emotions.

- *Agents for help and learning.* Agents can 'sit at the side' of an application, providing help and guidance on its usage (e.g. Horvitz et al. 1998). Agents can also act as tutors or co-learners in a learning application (e.g. Lester et al. 1997). Possessing emotions or exhibiting emotions might make these agents act more timely (not disturb) and motivating.

- *Delegation.* Users can delegate tasks to agents, such as activities that should be done when the user is away from the computer, monitoring events from other sources, or performing activities at distant locations in the network (e.g. Lieberman 1995). In these situations, the agents may use emotions to direct their efforts to the most important goals/paths to follow.

- *The subjective focus*. Agents provide an anchor for subjective evaluations that follow a presentation (e.g. who it is that gives a helpful suggestion) (e.g. Schank, 1991). An agent that possesses emotions and personality will be able to provide a more interesting, compelling, subjective focus.

- *The dialogue partner.* Agents provide a counterpart in natural language dialogue (e.g. Bretan 1995, Beskow et al. 1997). Again, expressing emotions, personality, and attitude, for example through affective speech, might contribute to the naturalness of the dialogue, but it might also raise too high expectations of anthropomorphism – thus a careful balance between expression and expectation needs to be employed.

- *Emotional behaviour.* Agents can both show emotions and arouse emotions in the user for general entertainment purposes (e.g Elliott, 1997). It can be in MUDs, interactive drama, games, or in advertising and branding. It can also be used for the user to explore emotions.

- Agents as user *representatives*. Agents can be instructed to behave in certain ways and then go out into a world and represent us – avatars are simple examples of this (e.g. Vilhjálmsson and Cassell, 1998). Emotion modelling will contribute to creating avatar behaviours that need not be consciously controlled by the user. Instead the agent can take on the responsibility for generating the right facial expression and body behaviour at the right time.

Studies of interactive characters

Interface characters have been much criticised and debated in the HCI community (Shneiderman, 1997, Lanier, 1996, Suchman, 1997). They are said to violate good usability principles, to obscure the line of responsibility between human and machine, and to confuse both designers' and users' understanding of the computer's abilities and inner models of events. The proponents, on the other hand, regard these parameters as opportunities rather than reasons to avoid characters in the interface (Höök, 2000, Waern and Höök, 2000).

Motivational effects

A number of studies have examined the ways in which characters enhance engagement and encourages exploration of a given information space, mostly in relation to learning and creativity. Such *motivational effects* were studied by van Mulken and colleagues (1998), who compared two versions of the presentation system PPP Persona: one with and one without a character. The study showed effects neither on recall of the presentation, nor on how the presentation was understood (objective measurements). However, it revealed a positive effect on the subjective estimation of whether the explanation was difficult or not. Subjects experienced the explanation as simpler with the PPP Persona character than without it. van Mulken and colleagues named this 'the persona effect'.

Another similar study looked at the persona effect for 'Herman the Bug', a pedagogical agent that helps students to create an ecological micro world system with plants, light and air (Lester et al., 1997). Here five different clones of the agent were compared, and the study revealed a persona effect – a strong positive effect on the students' perception of their learning experience. The animated character also had an effect on learning.

In a study by Wright et al. (1998) a plain textual explanation of a medicine was compared to one with the same text but with an animated dragon illustrating the different threats to the blood system. Here a negative effect on how much was remembered afterwards appeared; the dragon disturbed subjects, rather than aided them.

These conflicting results (PPP Persona and Herman the Bug, versus the dragon studies), point to the need for a better understanding of the design of synthetic characters in order to make use of their potential to encourage learning and exploration, and at the same time avoid the scenario in which the character distracts and disturbs the learning process. This involves, we think, a better understanding of the features of and relationship between wayfinding and exploration activities.

As pointed out by Andrew Stern (Hayes-Roth et al. 1998) (designer of the Catz and Dogz system) the artistic design and practical understanding of the creating of synthetic characters is crucial in determining the success of a system. A similar point is made by Elliott and Brzeinski (1998) when they cite Lester et al. (1997):

> "Lester gives the examples of, on the one hand, a humorous, lifelike, joke-cracking, character that ultimately impedes problem solving through his distracting presence; and on the other, a dull assistant that always operates appropriately but yet fails to engage the student. When communications from an agent must be coordinated to be both engaging and purposeful issues in timing, and the multi-layering of actions arise."

Anthropomorphic effects and believability

Another effect of synthetic characters is the ways in which they tend to raise expectations of anthropomorphism of the system (Reeves and Nass, 1996). Such *anthropomorphic effects* seem to have many dimensions. On the one hand the user may expect the system to be intelligent and cognitively potent. Brennan and Ohaeri, (1994) showed that users talked more to the anthropomorphic interface. King and Ohya, (1995) showed that users attributed more intelligence to anthropomorphic interfaces. Koda and Maes, (1996) showed that realistic faces are liked and rated as more intelligent than abstract faces. Opponents of synthetic characters argue that raised anthropomorphic expectations may lead to frustration in the user when the

system cannot meet the expectations (Schneiderman, 1997). For instance, the presence of a talking face might influence the user to expect the system to posses natural language and dialogue competence, which no system of today can live up to. The general conclusion is that the more 'natural' the interface, the higher expectations on intelligence in the system.

Believability and Emotions

Besides anthropomorphism, which refers to any human-like characteristics in an interface that makes the user think that the system has some form of intentionality and human-like capacity for reasoning, the issue of character *believability* is sometimes mentioned in the literature. Exactly what is meant by believability is somewhat unclear. Some use the concept to refer to the facial expressions and body language of the character – the idea is that the more human-like and naturalistic, the more believable. Our view is somewhat broader, including the personality and attitude of the character, perhaps less focused on the bodily expressions. To our knowledge there has been no studies focusing on believability, in this sense, in the field of interface characters.

At its most basic, believability involves the assumption that not all users will approach intelligent agents with the explicit aim to pick them apart to check their intelligence level. Instead they might be willing to indulge in an experience with the interactive character where they are amused, intrigued, and stimulated. It is therefore the designer's task to make the character behave in such a manner that the illusion lasts – at least until the aim of it has been fulfilled. By aim, we mean anything from a task-oriented goal to more entertainment applications.

In making the illusion last, we can be inspired by many things; animated movie characters is one source of inspiration, literature another, but one important source is of course how human beings behave. Many researchers have therefore turned to the psychological literature on how human cognition and emotion works, as described earlier in this deliverable. But an important issue for believable agents is whether human emotions need to be replicated in this exact way.

As pointed out by Damasio, Descartes made an error when he divided human cognition into body and soul where the soul is some magical way conveys its intentions into the body. Instead, cognition is embedded in the body, in the brain, in the nervous system, in a feedback loop between internal and external events and bodily behaviours, and cannot be separated from it. When we attempt to build systems that either 1) imitate this or 2) tries to "fool" users into believing that the systems has these kinds of properties, we are of course lacking the human body. But it is only in case 1 that we really need it. In case 2 we can equally well make simplistic models that act as if there had been a body, a body language, a social context, an inner life, etc. Whether it is possible to go through with 1 is a matter of philosophical distinctions: if we (spending years and years) build something that in every respect models human behaviour (including emotions, social behaviour, etc.) but that is realised in "hardware" instead of "wetware" can we really claim that it possess intelligence? Or must there be a subjective, conscious, self in order for this to happen? For out purpose here, we can leave this issue aside since we are aiming to establish how we can evaluate and assess interface characters. On the surface we might assume that there is not going to be huge differences between characters that fake emotions and cognition, and characters that do indeed possess complete cognitive and emotion models.

In case 2, when we deceive the user to believe that the created character/agent/system they interact with does indeed have emotions and intelligence, the problem is different. In that case, simple models with partly ready-made stories or simple mechanisms to couple external

events to emotional expressions might be enough. Sometimes people do like to be deceived: they will go to see a movie, they will read a novel or a comic strip, or watch soap operas on TV – just to be deceived for a short while to believe that these characters do exist.

Active users

Even if synthetic personal character assistants are intentional interfaces that are 'deceptive' in some way, the role of the user is not a passive one. The user will interact with the character and infer a lot of the characteristics of it: its personality, intelligence, or emotional state. S/he will also fill in the gaps in the story/dialogue that develops between user and system. Here we can be inspired by theories from the humanities on narrative construction. One of the basic presumptions within narrative theory in general, and discourse psychology in particular, is the notion of a reader/spectator constantly striving for *coherence* in his or her understanding and experience of a given text (e.g. Bordwell, 1985). Coherence is, on this account, *accomplished* or *constructed* by the reader through a huge battery of tacit, and hence non-conscious, everyday assumptions or knowledge about the perceptual, physical and socio-cultural world, influenced of course by pragmatic parameters (e.g. the purpose of the reading).

In the entertainment business when, for example, creating animated characters, it is well-understood that not everything should be explicitly stated but that the spectator must be made involved and curious about the next turn in the story. A basic assumption here is that characters – whether they appear in computers, fiction or cinema – are not only made sense of through the features of face and body, but through the ways in which character act within situations, display humour and share values and moral perspective of the user (Tan, 1996). It is, for instance, not enough to place characters in a film in order to evoke emotions (and thereby commercial success). These characters must move, act, feel, think and talk in concrete situations in ways that raise the spectator's sympathy or antipathy. As pointed out by Porter and Susman (2000) when discussing how to create life-like characters in cartoons ('Toy Story'):

> "They (animators) understand that 'life-like' does not mean 'has movement'; lifelike means 'has a brain'. The underlying notion of Pixar and Disney animation is that action is driven by the character's cognitive processes – that it reflects intelligence, personality and emotions."

These dimensions, however, are much more difficult to design for and evaluate.

Interactive Characters in Computer Games

Current trends in computer games suggest that in the future, they will become populated, not just by people, but also by (possibly large numbers of) autonomous and intelligent agents and characters. In this context, such agent's behaviours and expressions will be of particular importance for inducing to the user a sense of presence and lead to the illusion conveyed by the game. Classical animators are masters at conveying intentionality and emotions through the physical movements of their characters. In The Illusion of Life, Thomas and Johnston (Thomas *et al*., 1981) argue that simple changes in bodily movement can convey radically different emotions to human observers.

Similarly, graphically represented autonomous agents could express a great deal of emotion and personality based on the way in which they move their bodies -- if it were not for the fact that these highly expressive bodily movements are surprisingly difficult to design and to

control in a way that is dynamically responsive to user activity. Expressive actions of synthetic characters in an interesting research problem for affective system design which has only begun to be explored.

Design problems

Crucial design problems that needs to be addressed in this area are the following:

- Agents need to display behaviour and affective expressions in such a manner that the user *understands* them. This means that they cannot always act in the most efficient rational way but instead they might have to act in ways that conveys to the user what is going on (Sengers 1998). This also holds for agents that work in multi-agent systems but where their results have to be communicated to a user in the end.

- Agents need to be timely. When an emotion is displayed to the user it has to come at the right point in time, and last for an appropriate length (Ruttkay et al. 2000). If an affective response from the user is the aim, then the interaction has to be carefully paced so that the user can follow it without being bored or puzzled.

- Agents sometimes need to have interesting personalities. Only then will their emotional behaviour be comprehensible and interesting to the user. Conveying the personality might be difficult if the interaction with the user is limited. This is where idle behaviour or interaction between several agents can come into play. When several agents interact, they can take the opportunity to show more of their personality traits.

- For some affective agent situations, it is necessary to create a narrative context (a situation, an interaction history) in order to understand the emotional behaviour (Isbister and Doyle, 2003).

- If the agent is used for a longer time span with a user, different personalities and attitudes might be needed in order to fit the needs of different users (Boyce 2000, Nass and Gong, 2000).

Evaluation of ECAs

When designing an Embodied Conversational Agent to act as interface with the user, one should answer a few questions. First of all, is the ECA the best interface? (Where 'best' needs to be defined in the context of the interface: the most convenient, the most natural, the easiest way to interact with the system, the most enjoyable, etc?) If the interface does require the use of ECA, how to design it? Should it be 2D, 3D? Should it be a realistic or a cartoon-like model? What should be its gender and its age, a young male or an older woman? Should it have a humanoid aspect or not? Should it be a puppy or a famous cartoon character? Should the agent look more like a French or like a Japanese? Once details on the geometry of the agent has been decided for a given application targeted toward given users' profiles, once has to wander not only how the agent should communicate information, but also how should the agent interact with the user(s)? One has to ponder if the agent would be extrovert or introvert? Would the agent be emphatic, or expressionless? How should the agent communicate, with which modality: with the face only, the gaze, gesture? Should the agent be sparing in its communication style and use one modality to communicate an information; or should it be very redundant and use as much modalities as possible when talking? Does redundancy affect the style of the agent, or does it also affect the interaction quality, and if so along which dimensions (learning, maintaining attention, attracting attention...)? Which role should play

the agent in the interaction? What should be its goals? Should it persuade the user, or simply provide information, or may be show empathy? From this long list of questions one can see the diversity of choices that need to be made in order to design an ECA. Evaluation of these choices and on their effect on the design of ECA is a necessity.

Evaluation should happen at different levels: the micro level in which one aspect of the agent is observed; the user level in which the reaction of the user is evaluated; the application level in which the agent within an application is studied.

Evaluation of agents and of interactive systems using agents are still relatively new. No benchmarks or standard evaluation methodologies exist, though several attempts to remedy such a lack have happened (Ruttkay et al., 2004; Christoph, 2004; Catrambone et al., 2004; Isbister and Doyle, 2004). In particular, Ruttkay et al. (2004) propose a common definition of terms that are largely used in the agent community, but with different meanings associated with them. They propose a taxonomy of the various dimensions upon which evaluation should be made (e g, agent's appearance, function, believability). On the other hand, Christoph (2004) proposes a methodology to evaluate user's interaction with the system. The author proposes ways to collect data and analyze them, to select the appropriate subjects of a study as well as to define the research strategy in order to set up properly the evaluation studies. She bases her guidelines on evaluation study from empirical research.

Evaluations at the micro-level of the design of an ECA involve the evaluation of aspects such as the muscular computational model of the face (muscle contraction, skin model, wrinkle), the lip model, and/or emotion model. Several muscular and skin models have been developed (Lee, 1995; Wu, 1995; Guenter, 1998; Pighin, 1998). Some are physically-based and propose an accurate model of the muscle and the skin; while others offer a rendering-based method which give realistic visual results. Researchers have also developed computation models for lip shape and co-articulation effects (Cosi, 2002; Bevacqua and Pelachaud, 2003; Cohen, 1993; Beskow, 1995; Reveret et al., 2000; LeGoff, 1997). Evaluation methodologies have been proposed by Massaro (1997). Computation models are evaluated at the syllable level, the word level, or even the sentence level (Massaro, 1997; Siciliano et al., 2003). Comparisons between models can be done using an objective evaluation: for example by comparing between real targets of lip parameters and predicted parameters of the computational model. On the other hand, intelligibility evaluation works at the perception level: by performing perceptual tests on subjects. Studies have shown (Beskow, Ph.D. thesis) that results using objective evaluation and intelligibility evaluation may not be correlated. Indeed models with hyper-articulation give better perceptual results even-though it shows much less natural movements.

Evaluation at the micro level looks also at behavior along one channel. For example, Krahmer and Swerts (2004) found that eyebrow movement on pitch accent does not play the same role for Dutch language as for Italian language. Gaze behavior has also been studied (Kraemer et al., 2004; Garau et al., 2003). This behavior has a great impact on the way the user perceives the agent. Buisine and Martin (2004) investigate if the strategies of using multimodalities in communicating have effect on the perception of ECAs as well as on the learning dimension. The authors designed a study in which three different looking agents use just one modality (say gaze OR gesture) to communicate an information or use redundant behaviors (say gaze AND gesture) to communicate. Moreover Cassell and Thorisson (1999) found that users prefer interacting with agents communicating multi-modally: users found the agent more natural if this one shows gaze, gesture and facial behaviors. The study reported below looks at how verbal and nonverbal behaviors may persuade users or not.

Other evaluation studies look more particularly at the user standpoint. Höök's research is a good example of such studies (Höök et al., 2000; Höök et al., 2003; Höök, 2004). Höök works at the user's perception level, investigating how agent design contributes to the users' experience of the interface. Using several applications, she examines how role, behavior, speech, and emotions of agents act on the users' emotions. She argues that developing an agent should be done in a loop of design and evaluation; that is to say, the design of an agent should be interwoven with evaluation studies looking at the effect the agent has on the user's experience. Nass and his colleagues have performed important studies on different aspects of agent's design: culture (Lee andNass, 1998), personality (Isbister and Nass), and appearance (Nass et al., 2004). They based all their research on psychological studies.

Given an application, the impact agents may have on the user has been reported (McBreen et al., 2001; Morton et al., 2004; Krenn et al., 2004; Darves and Oviatt, 2004). Studies have also looked at the effect an agent's clothing (McBreen et al., 2001), intonation (Darves and Oviatt, 2004), and gender (Buisine et al., 2004) may have on dimensions such as trust.

General Affective System Research Summary

Some work has also been done to create systems that incorporate affective components without explicitly evoking anthropomorphic relationships with ECAs. Rosalind Picard and her research team at MIT's Media Lab have focused on both reading and generating affective cues in a wide range of projects (Picard, 1997). The Influencing Machine is an art-based exploration of how users will interpret and engage with a machine's internal affective model (Sengers et al, 2002, Höök et al., 2003).

Work has also been done to incorporate non-mouse-based interaction with systems, towards richer affective engagement. SenToy — a doll-controlled game (Höök et al., 2003, Paiva et al., 2003), and eMoto — a gesture-controlled device for colouring mobile phone messages with emotional cues (Fagerberg et al., 2003) are two examples of this approach. Cowie's work on analyzing affective cues from speech (Cowie and Cornelius, 2003) should help further efforts to allow use of one of the most natural affective ways of interacting with a system: human speech.

Non-character-based emotion synthesis and analysis research has also been carried out in the study of affect and music. Bresin and his colleagues (e.g. Bresin and Friberg, 2000) have isolated and parameterized dimensions of musical performance that have been incorporated into a system which can generate performances of the same piece of music with very different affective qualities.

Finally, work has been done to bring techniques from film and other narrative media to enhance affective engagement with interactive systems (Laaksolahti et al. 2003).

5. Review of key problems in the thematic area

Affective interfaces include those that express emotions, those that attempt to understand user emotions, and those that use affect as part of the system's "mind". User evaluations in this new field are still infrequent, and lacking an established theory providing similarly well defined targets to assess as in traditional human computer interaction. At the same time, most researchers have been focussing on the natural phenomenon of affect and its expression, communication, perception, and hypothesised processing, without being overly concerned with whether qualities such as believability or naturalness of synthetic characters do in fact contribute to the overall success of a system. Neglecting the relationship between natural source of inspiration and actual engineered systems, designers and researchers are all too easily misled into not questioning whether inter-human communication really forms the best model for interaction between human and machine. Results from natural language interfaces, adaptive interfaces, or intelligent user interfaces indicate that there exist principles and idiosyncrasies that are specific to the design of human machine interaction (Dahlbäck et al., 1993, Höök 2000). Among what little results exist today, there already are a number of findings underlining the importance of empirical verification and questioning of intuitions and working hypotheses of the appropriateness of the deployment of affective functionalities for specific purposes, e.g. (Issroff and del Soldato, 1996).

On the other hand, the well-established goals of usability research, including effectiveness, efficiency, or safety either are probably not adequate for scenarios centred on subjective user experiences of being satisfied, rewarded, pleased or motivated; at least there is no clear understanding of what should be measured in which way so as to assess the standard usability qualities with respect to these new targets. Many other difficulties ensue from the social and biological grounding of the phenomenon to the investigated. Affective interfaces address directly inherently adaptive faculties of humans; the methods for user studies therefore have to take into account that not only test users, but also "emotionally enabled" artefacts will alter their behaviour within and across evaluation runs. As also mentioned in many other parts of this document, in different contexts, the manifestations of greater or actual interest are not those of extreme states, but rather the subtle emotions; the challenges posed by these difficult to detect and highly susceptible targets are complemented by those of rarely occurring emotional states and of those that take significant time to develop. The latter are only two examples making it rather evident `that` usability assessments concerning affective interfaces hardly can be sped up at will or carried out in a conclusive fashion within a limited timeframe using standard approaches. To mention just one last example, the simple circumstance that users are *always* emotional immediately points to the difficulties in clearly identifying affect that is due to and related to the application at hand, excluding reliably exogenous causes; conversely, the scope of usability assessments is difficult to contain when consequences outside the intended reach of the application are hard to rule out or prevent.

The situation is further complicated by the fact that studies of affective interfaces are being carried out against a background of theory and technology that is still embryonic. Appropriate utilisation of reported usability results therefore needs to be carefully considered among the alternatives of performing tuning or adaptation of the currently employed system components, altering the architectural design introducing different components that replace or complement the current ones, or initiating a new research thread defined by the intersection of open issues and indications from the usability study.

Taken jointly, these three views on usability of emotion-related systems document vividly how usability research forms a particularly well suited hub that not only invites but

downrightly requires collaboration and exchange of information across all of the multifarious research community. The circumstance that every discipline can both contribute as well as gain from the interactions is an important facilitating aspect.

The aim of emotional systems is to get users affectively touched by the interaction with the system and even to maintain this relationship or illusion over a longer interaction session. This is, obviously, a hard and very difficult goal to achieve. We know for sure that movies, novels, television shows, arts, music, are indeed able to get people affectively involved. But we want to make end-users affectively touched by interacting with systems that model emotions, reason using emotions and express emotions.

There has been quite some research in Europe as well as in the states on how to recognise users' emotional states through singular, one-off, readings of biosensor data, facial expressions, body posture, interaction with devices, such as mouse or keypad, or props, such as plush toys, but repeatedly there seems to be the same conclusion: while this may recognise some basic emotions (fear, stress, and arousal), they fail to get the whole picture and oftentimes contradictory results between users' self-reports of what they think and feel and their physical expressions arise (e.g. Höök et al., 1999). They also fail to understand any more complex and interesting emotional states that users might be in – such as shame, guilt, positive arousal, or flow.

Affective interaction as an interaction problem

In WP9 our aim is instead to see usability of emotional systems as an interaction problem that concerns the whole interaction cycle and where emotions arise from an active act of interpretation and participation from the end-user side. We can compare this to, for example, how an empathic successful teacher works in a classroom (Cooper et al., 2000). An affective teacher will not only recognise the affective state of the pupil, but also rather try to influence the pupil's state through communicating the state that is most desirable in order to solve a problem. An emphatic teacher will illustrate that a math problem is difficult and requires lots of attention through a strong physical facial expression showing of effort. The pupil imitates this expression – imitation is strong, innate, ability – which in turn induces the right kind of starting point for working hard. The imitation of the facial expression together with the cognitive understanding that the problem is hard achieves the right state for working hard. This is an *interaction* between the pupil and the teacher. It is not an attempt at recognising the pupils state solely. They interact and the goal is to achieve the desired state – to learn math and solve the problem at hand – not to recognise the pupil's initial state.

The other half of the problem of the affective loop lies in producing the right output from the system to maintain end-users involvement, and to do this over a whole interaction. The proposal pf WP9 is to tackle this part through the user-centred approach to design. The field of affective computing has mainly been driven by AI researchers with little or no knowledge of how to work with end-user applications. The field of HCI on the other hand, have addressed this problem for a couple of decades. Many important lessons can be learnt from their work. One such lesson is that human-computer interaction is not, and cannot be, modelled from human-human interaction. A computer system is a designed artefact – not a "natural" thing. While the field of HCI certainly recognises that there are design considerations that should be built from knowledge of human abilities and limitations (see e.g. Norman, 1990), they also recognise that computers are part of human culture, and thus subject to change. Over and over, artefacts are designed that users then take into use in ways that are quite different from what the designer expected (Suchman, 1987, 1997). A design process that fails to involve end-users in the design loop, will fail to recognise the particular quirks and problems of how to design these artefacts.

A main part of WP9 will aim to take the lessons learnt and methods and theories of HCI into account in the design of emotional systems. Through involving users in the design process at several stages of the development of a showcase, the aim is to find ways to ensure the development of emotional systems that users find usable and able to create experiences that are meaningful and engaging to them.

6. Assessment of the key development goals in the thematic area

The first goal of WP9 is to explore appropriate ways for understanding user experiences – *Ends*. To address this area, we plan to explore development of evaluation instruments that supplement the current state of the art—questionnaires, video recording, and bio-sensors—with nonverbal, subjective measures of affective success. Along with this work will be included a review of current user evaluation strategies. An end goal for this work is the creation of a new evaluation tool that can be used by creators of affective systems.

The second goal of WP9 is to evolve the philosophy and craft needed to design for affective experiences--*Means*. Initially, we shall bring in existing methods for user-centred design from the Human-Computer interaction field. As these are applied in the exemplars (see below for exemplar ideas) experiences will be gathered and the methods will be revised.

A user-centred design framework for affective interaction will be developed. It will consist of methods for:

- how to describe and understand characteristics of the end-user group (e.g. Cooper, 1999),
- brainstorming, such as "Random Words" (http://www.randomwordgenerator.com/index.html),
- early idea evaluation, such as "Six Thinking Hats" (deBono, 1985),
- user-centred design, such as "Contextual Design" (Beyer and Holzblatt, 1999) providing real-life (light-weight ethnography) input to the specific scenarios or into specific settings, such as the home (Gaver and Dunne, 1999),
- early (drama and paper-based) development of ideas for user-testing, such as "Prototyping with Tiny Fingers" (Rettig, 1994), drama (Iacucci et al., 2002), staging happenings (Iaccuci and Isomursu, 2004),
- design approaches, such as making use of ambiguity for open interpretation of affective expressions (Gaver et al., 2003),
- fake system testing for end-user interaction, such as the Wizard-of-Oz method (Dahlbäck et al, 1993, Andersson et al., 2002),

The aim is to change and refine the methods throughout the lifetime of the WP9 in order to arrive at a set of efficient and viable methods for design that help forward the craft and knowledge of how to design emotional systems. Each exemplar will provide feedback on how well the methods worked for creating affective loops.

7. Relation to other workpackages

WP6 – Two shared exemplars (Gossiping Face Project, and Facial Parameters Project. See our exemplars for general descriptions of the projects). By working with Work Package 6 in the development of these systems, we hope to test and refine our evaluation techniques and instruments, which will serve as our final exemplars.

WP5 – Development of database and ways of annotating material that can serve as input to the design process, using materials gathered in proposed Exemplar 2.

WP3 – Absorb and build upon current affective theory to create appropriate evaluation instruments.

8. Preliminary ideas about possible exemplars

Note about exemplars: Work Package 9 seeks to innovate means for evaluation of affective systems. These will be our Exemplars, and may be used in the development and evaluation of one ore more systems developed by our group or by other Work Packages.

There has also been some discussion among a subset of Work Package 9 about developing assessment instruments and comparisons for examining the effectiveness of particular components or variables within affective systems, see exemplar suggestion F below.

A. First Possible Exemplar: Affective Evaluation Instrument

Description of problem:

Neurophysiologists and psychologists have in recent years proposed that our brains, rather than operating in a wholly logical, conscious verbal manner, actually process information and make decisions using various layers working in parallel, complementary ways. They have demonstrated, for example, that we can learn something new and 'intuitively' put it into action before we are able to consciously verbalize it (Myers, 2002).

Affective processing, in particular, has been shown to occur at levels other than the cognitive/word-oriented level of the brain (e.g. the primal nature of fear: http://www.nimh.nih.gov/events/ledoux.htm).

Yet we rely heavily on *verbal* reports from users for evaluation of the affective effects of systems. Likert scale items (e.g. How enjoyable was this interaction? Circle the number from 1-7 that best fits.) are still the most common subjective quantitative measures used.

Of course it is possible to supplement verbal reports with some form of observation of affective reaction—videotaping or physiological tracking. However, we wonder if it is possible to develop a form of subjective report mechanism that 'end-runs' the verbal part of a person. The exploration of this terrain is the goal of our project.

Such an instrument could provide several key benefits:

- Preserves benefits of subjective measures (easy to use, portable, empowering to user).
- Transcends language and cultural barriers.
- Provides results more in line with current research about affective processing—potentially less distortion through the 'lens' of the verbal mind.
- More fun for the user!

Description of idea:

We would like to explore the possibilities for subjective, self-report of affective state in *nonverbal* ways. We are especially interested in developing an instrument that could be calibrated and re-used to measure affective reaction to systems and interactions, across cultures—an **affective instrument**.

Design criteria:

- Intuitive to use
- Nonverbal (makes use of visual/tactile/auditory and other nonverbal cues instead)
- Provides quantitative results
- Easy to calibrate for different user groups
- Portable, durable

Initial design suggestions:

Our first brainstorming session led us to the following suggestion:

We would create (in partnership with product design experts) a kit of non-representational objects that make use of affective cues (e.g. colour, texture, shape).

Before using this kit to evaluate an interaction, we would have users calibrate the objects by arraying them along the valence and arousal scale of affect (see Russell for a discussion of these dimensions and how they correspond to emotional states). The positioning of the objects would be tracked and each object would then receive a value along the scales for use in subsequent interface evaluations. From the eMoto-project (Fagerberg et al, 2003, 2004), we have picked up a version of Russell's circumplex model of affect including colour, shapes and animations that illustrate different parts of the circular model and the dimensions of arousal and valence, see figure 1.

An alternate model worth exploring could be the crafting of multiple representational objects for the basic emotion states—assertion/anger, aversion/fear, satisfaction/happiness, and disappointment/sadness (see Turner), which could be arrayed along a magnitude scale for subsequent use in evaluation.

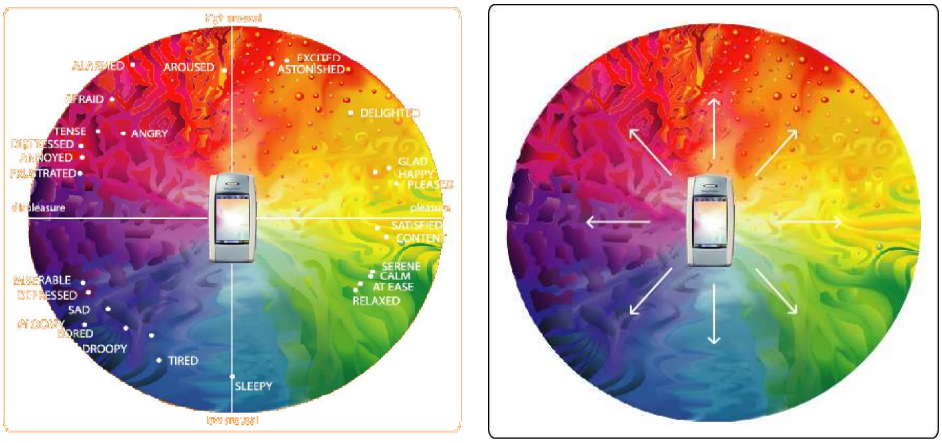

Figure 1: Russell's circumplex model of affect with the dimensions arousal and valence. Illustration with colour, shape and animations from the eMoto-project (Fagerberg et al. 2003), with emotion labels to the left, without, as used in eMoto, to the right.

Of course, these are only initial proposals, which we plan to test and evolve through iterative prototyping working with users as well as product designers, using our design criteria as signposts.

Context/background:

There is some history of the standardization and use of nonverbal scales in psychology (e.g. PONS (the Profile of Nonverbal Sensitivity), from which we can draw lessons.

There are popular uses of nonverbal affective scales (thumbs-up and down movie ratings, the 'little man' on the San Francisco Chronicle movie review page (http://www.sfgate.com/eguide/movies/reviews/) indicating that calibration and use of nonverbal scales is possible and appealing in everyday contexts.

There has been some work in the product design community on mapping product qualities to affective reactions, which we hope to draw upon (see Green and Jordan, (2002) for some recent work in this area).

Next steps:

- Pursue relevant literature for background.
- Conduct preliminary tests of our calibration and evaluation-during-interaction strategies.
- Partner with design schools/researchers toward further brainstorming and prototyping of the kit of objects.
- Work toward larger-scale, cross-cultural testing of the Affective Instrument, and production of several of the final kit, for ongoing use in the field.

B. Second Possible Exemplar: Using Casual, Contextual Video Recording for Affective System Development. System Context: Gossiping Face Project (Work Package 6)

Description of problem:

Talking faces are one output of affective interface research. As of yet, there are not compelling applications for these faces, that truly take advantage of their expressive power and have demonstrably favourable user response.

We believe this may be in part due to a focus on tactical versus social conversation applications for such faces. Emotional expressions are layered onto essentially functional dialog (e.g. about health care or real estate), without regard for whether the expressions truly provide additional value to users.

We also believe that talking faces may gain in appeal and believability to the extent that we ground their development in situated observation of facial expression. That is to say, if we want to design a social application for the face, we should look at how people really use their faces in such social contexts, and build the expression set and model from this sort of contextualized data. This is in opposition to working from generalized models of emotional expression and then layering them onto a given social situation.

Description of idea:

We would work in partnership with WP6 to develop a talking face application that is primarily social in nature, building from observation of real human faces in an appropriate context.

The target communication context is engaging in what is known as 'shit talk' about others—gossiping about celebrities or other people, and either praising or criticizing them.

Our plan is to videotape young women engaging in this kind of talk, in several countries (currently, Sweden, the U.S., Japan, and possibly also France and Ireland). We plan to analyze the results of this taping with the assistance of WP3, to discern patterns of emotional expression.

We would work with WP6 to use our results to evolve an engaging application with a talking face provided by one of WP6's participants, Cantoche.

WP9 would provide the evaluative expertise to the team as the design moved all the way to trials with real users.

Next steps:

- The collaboration partners (WP9, and 6) will meet to brainstorm based on the results of the initial videotaping, in early June.
- Cantoche, a partner in WP6, has agreed to provide a face and tools we can use for developing the application, to be available in September.
- WP3 and 9 need to meet to talk about analysis techniques and tools.

C. Third Possible Exemplar: WoZ Puppet. System Context: Greta Face System Iteration (Work Package 6)

Description of problem:

Catherine Pelachaud and her research team (in WP6) plan to add dynamic qualities to the greta synthetic face that they have created, in order to provide a wider range of expressive ability. They would like to use user-centred design methods and evaluation to improve the process and outcome of this addition.

They would also like to explore potential application areas for the face that take full advantage of its expressive capabilities.

Description of idea:

We suggest that the Greta team could benefit from using some innovative brainstorming and testing strategies, including:

- Early brainstorming with potential users of desirable dynamics for the face (to help define the feature set)

- The creation of a physical puppeting interface for early versions of the dynamic controls, which could then be used in WoZ (Wizard of Oz) studies to determine which controls are used under what circumstances (to help give direction to how the dynamics should affect the face, before computational modelling even begins)

- Brainstorms with users, using the puppeting tool as a prop and aid, about potential application areas for the new version of the face (to come up with a strong application for the face that takes advantage of its new capabilities).

Context/background:

Work Package 9 team has an interest in physical brainstorming tactics, and believes providing potential users with a kinaesthetic method for playing with the face will help them understand and evolve its capabilities, saving the WP6 team initial modelling work and allowing for more targeted and user-centric iteration and system development.

Next steps:

Further meetings to refine the research direction and to discuss initial brainstorming plans.

D. Fourth Possible Exemplar: Dramatic Gaming

Description of problem:

Action games have established themselves as the dominating genre in contemporary computer gaming. Games striving to provide other kinds of experiences, notably emotional and social experiences, are still rare. The success of games such as The SIMS however, indicates that there is a demand for other types of games.

One promising genre in this respect is interactive drama that focuses on providing a dramatic experience rather than an action packed one. Instead of monster bashing players can engage in navigating the socio-emotional web linking characters together. This is something that requires players to be both sensitive and attentive to social and emotional cues displayed by characters. In an interactive drama the motivating factor for playing a game switches from "winning" to exploring the social and emotional relations between characters.

How to design for interactive drama is an open question. So far research has focused mostly on how to create believable characters (discussed elsewhere in this deliverable) and to a lesser extent story sequencing. Undoubtedly believable characters are an important part of interactive drama as described here. Creating characters with the necessary attributes to become interesting counterparts in emotional and social interactions is challenging. However, there are also other aspects that can influence a players experience and function as emotional boosters e.g. how characters are presented (cinematography), sound effects and music that have received less attention.

Description of idea:

We propose to design and implement an interactive drama system partly based on existing technology developed at SICS (Swedish Institute of Computer Science). By using existing technology an early prototype of the system can quickly be developed.

The basic idea is that the player takes on the role of one character, and interacts with other – computer controlled – characters (NPC's) in the scenario. The game will be realized as a simulated dialogue between the player and the NPCs in the game. NPCs will be equipped with emotional models developed within WP3 but also social models that partly determines their behaviour in the world. Depending on how the player acts or reacts emotions and relations will change and influence the unfolding of the drama.

The 'story' of the game is intended to be about changing relationships between the characters. As the game progresses we want the system to push events in a direction that creates a dramatically interesting story. This can for instance mean favouring actions that sow discord between the characters.

An important step on the way will be to refine the scenario for the game. The idea is to use methods developed within the WP (e.g. the affective instrument) to incrementally improve on the system as well as the scenario. Results from the evaluations will feed back into the iterative design/evaluation/redesign cycle.

As suggested earlier in this deliverable designing and controlling highly expressive bodily movements of characters is a difficult task. Hence we wish to explore cinematography as another, possibly less complicated, method for expressing emotions. Cinematography refers to *how* something is filmed – in contrast to *what* is being filmed – and typically involves three factors:

- *Photographic aspects of a shot,* e.g. how a shot is illuminated. For instance a shot can be very dark and gloomy, light from the sides casting sinister shadows, or it can be bright and happy.
- *Framing of a shot,* i.e. what is included in the camera rectangle and its location within the rectangle. For instance, a shot can be centred on a person talking to someone outside the frame, or both persons can be visible on opposite sides of the frame.

- *The duration of a shot*. A shot can be very long, e.g. showing a person giving a speech, or short, showing the person giving the speech and then rapidly cutting to the audience's reactions to the speech.

By altering these factors different dramatic and emotional effects can be constructed although the content remains the same. Inspired by comic artists, colour theory and industrial design we also want to explore how colour and shape of shot frames can *boos* the emotional content of the drama.

The purpose of the game is to provide a pleasurable experience to the player. To determine whether the game has succeeded we need to understand what it means for something to be fun or pleasurable and how to capture it. Ideally we would like to induce a sense of flow in the player. It is a concept often used in games research to describe a very desirable state:

> "It is the state in which people are so intensely involved in an activity that nothing else seems to matter; the experience itself is so enjoyable that people will do it at great cost, for the sheer sake of doing it" (Csíjszentmihályi, 1990).

Like believability, flow in a game seems to describe a situation in which a subject has been so engaged that they 'disregard' the physical circumstances of the playing situation or the technical features of the computer game, and instead 'enter' the story world, focusing on the events and become cognitively, emotionally and morally engaged in the lives of the characters.

Context/background:

There has been some work in automated cinematography. However, most work has so far focused on techniques for describing and solving the geometric constraints of a scene, such as where to place the camera given two characters talking to each other so that none of the characters are occluded (Hornung et. al, 2003). In contrast very little work seems to be available on how cinematography can actually be used to make a scene more expressive, understandable or believable. Our proposed work belongs to this second category in that less will be spent on the geometric constraints than on exploring the expressiveness afforded by cinematographic techniques.

Regarding the definition of what is pleasurable or not, Wiberg (2003) provides an overview of attempts to operationalise this difficult concept. According to Tiger (2002) and Jordan (2001) pleasure can be divided into four different dimensions:

- Physio-pleasures
- Socio-pleasures
- Psycho-pleasures
- Ideo-pleasures

Physio-pleasures relate to pleasures derived from sensory organs, such as smell or touch. Socio-pleasures are derived from relationships between people or relationships to society in large. Psycho-pleasures are derived from individual activities that often involve using skills to perform something that is emotionally rewarding, e.g. painting. Finally ideo-pleasures are derived from consumption of – and reflection upon – values and ideas from various sources such as books, films or newspapers. The importance of these various dimensions for evaluating fun in games will likely vary depending on the design of the game, the values conveyed by the game, which modalities are used to interact with the game (text, speech, mouse, keyboard, gestures) and so forth.

Next steps:

- Create a scenario for the drama. As a starting point we will use an existing scenario which focuses on three girls arranging a party for their friends
- Evaluate the drama using the "affective instrument". This will provide feedback on the drama as well as the instrument.
- Work together with WP3 to define the necessary models?
- Work together with WP6 in developing believable characters?

E. Fifth possible exemplar idea: WoZ environment for multimodal emotional interaction

Affective factors may influence the design of human-machine interaction in several ways. Typically, natural language messages generated so as to include consideration of the emotional situation of either (or both) the Speaker and the Hearer. This situation may affect the content, the order of presentation of information and the 'surface' formulation of one-shot messages (deRGrasso, 2000). It may affect, as well, the way dialogs proceed (who takes the initiative, which goals and plans prevail in every phase of the dialog and so on (Caval, 2004).

More recently, Embodied Conversational Agents (ECAs) extended with 'emotion expression modes' the range of features that were considered in natural language interactions.

Evaluation of affective interfaces, and ECAs in particular, were focused, so far, mainly on one-shot messages and on short-length interactions. The 'affective usability' measures applied were of two sorts:

- on one side, 'subjective evaluations' of aspects like 'believability', 'likeability' etc;
- on the other side, measures of 'user performance' (level of memory in information systems, level of persuasion in advice-giving systems and similar).

ECAs are a prototypical case for systems generating multimodal affective behaviour. A crucial criterion for ECA design is whether the multimodal affective behaviour simulated by the ECA is adequately perceived by the human spectator/listener. In other words, the system designers need to know according to which criteria they should model and assess the ECA behaviour so that a certain experience is created by/for the user.

A main problem, however, is that existing emotion theories mainly account for single modalities, especially the face (e.g. Ekman, 1993) and the voice (for an overview of theories on the latter see (Schröder, to appear)), while there is a discrepancy between emotion modelling and testing the adequacy of emotion display. Human emotion display as well as perception of emotion is inherently multimodal, and situationally and socio-culturally mediated. See for instance (Scherer, Ceschi 2000) for a respective study on emotion perception. Thus we, first of all, have to identify those modality-specific model parameters that are relevant in multimodal emotion display. Here a close connection to WP3 exists. Second, we need to define how we can appropriately ask for / assess the emotion perceived by the human through multimodal display via the ECA, and how to map between (single modality) model parameters and items to be assessed in human perception of multimodal affect display. Ideally, we also want to control for personality, situational and socio-cultural effects in the animated scenes as well as in the spectator's (= subject's) individual background. To which extent this can be achieved within HUMAINE is not yet fully clear.

Within HUMAINE, we wish to make a step further in the kind of evaluation studies which have been performed so far, by considering the following aspects:

- uni- and multimodal dialogs in medium-term interactions;
- variation in emotion expression in information content, order, linguistic and nonverbal forms,
- evaluations by single users and groups of users;
- natural and immediate forms of (individual and group) subjective evaluation (different from questionnaires).

As our interest is focused on *methods* rather than on evaluation of specific products, we wish to build a 'testbed' for evaluation of affective dialogs in the form of a tool for Wizard of Oz studies, and also use an already existing platform for simulating multi-modal interaction.

a. The **Wizard of Oz platform** will enable us to evaluate affective dialogs in natural language or with embodied agents: it will be domain and character-independent.
System moves will be represented with an appropriate mark-up language (APML: (DeCar03)) and will be rendered, by means of a wrapper, by one of the available Haptek characters with different ages, cultures, etc (HaptekWS). Simulated dialogs will be user driven vs. character-driven, empathic vs. more 'cold'; etc, to test the effect of various variables. This platform will be employed to compare *evaluation methods*: individual Likert-scale style evaluations vs. individual evaluations with other forms (colours, shapes, etc as suggested above) vs. group evaluations, with methods to be discussed with some psychologists. Individual evaluations will be collected in the interface, while group evaluations will be collected by either video recording or with some form of 'cooperative interaction tool'. These tools enable users grouped around a touch-screen type of table to interact with a virtual scene that is projected on the table. They may point to items in the scene (individual ECAs or their parts) and may express their viewpoint on them.

b. An evaluation version of the **NECA platform** will enable us to experiment with different channels of communication and their integration. This platform is a confederation of system components for generating and presenting multimodal dialogues that have been developed in the NECA project, see http://www.oefai.at/NECA, (Krenn et al. 02). Some of the current HUMAINE members were also involved in NECA, and therefore we have the ability to

directly manipulate the system code if required. In particular, animated conversations can be manipulated at the levels of (i) scene generation and dialogue modelling, especially the lexical realization of the dialogues; (ii) speech synthesis, including the manipulation of parameters relevant for the manifestation of personality traits (i.e., extrovert versus introvert speech), and the manipulation of parameters relevant for the vocal expression of emotional state; (iii) facial expression and gesture.

As in the WoZ system, we have the possibility for completely scripting animated scenes using the NECA RRL which is an XML-compliant multimodal markup language, cf. (Piwek et al. 02). A first step here will be to define a mapping between APML employed in the WoZ platform and the NECA RRL This will also feed in WP6, and fosters the further unification of multimodal representation languages.

Combining the WoZ and the fully automatic system in our evaluation studies opens up a broad range of possibilities to explore the ground of multimodal HCI, with and without human intervention, in the multimodal presentation of affective behaviour. As different parameters may be manipulated in the two tools, we will have the opportunity of testing the influence of a broad range of factors on the affective involvement of users in dialogs. Finally, we will be able to reflect on advantages and disadvantages of individual and group interactions with the technology, and to test the interest of employing cooperative design and evaluation tools in affective HCI.

9. Conclusions and Way Forward

At present, Work Package 9 has proposed several initial exemplar methodologies, and has already begun work with other Work Packages toward incorporating these methodologies into the design of affective systems that will be part of those WPs.

The primary research emphasis (innovating means for successful affective system design, borrowing from participatory design and other HCI tactics) has been set, and the exemplar's proposed reflect this approach. There may also be some sub-set of the Work Package group that decides to additionally pursue more component-based comparison of evaluation methodologies.

Work Package 9 has plans to carry each of the exemplar proposals forward to the next steps to help elucidate which will be the most promising to focus our efforts upon. This includes meeting with the relevant WP leaders and team members, as well as early data-gathering and prototyping.

10. References

A. Cited references

Andersson, G., Höök, K., Mourão, D., Paiva, A., Costa, M., (2002). Using a Wizard of Oz study to inform the design of SenToy", Designing Interactive Systems (DIS), ACM Press.

Beskow, J. (1995). Rule-based visual speech synthesis. In ESCA - EUROSPEECH '95. 4th European Conference on Speech Communication and Technology), Madrid, September 1995.

Beskow, J. (2003). Talking Heads: Models and Applications for Multimodal Speech Synthesis. Ph.D. thesis, Centre for Speech Technology, KTH, Stockholm, Sweden, 2003.

Bevacqua, E. and Pelachaud, C. (2003). Modelling an Italian talking head. In Auditory-Visual Speech Processing AVSP'03, Annecy, France, 2003.

Beyer, H, and Holtzblatt, K, (1999) Contextual Design, interactions, vol 6, iss 1 Jan/Feb, ACM 1999.

DeBono, E. (1985) Six Thinking Hats. Little, Brown & Company, Boston, MA, 1985

Buchenau, Marion, and Fulton Suri, Jane (2000), Experience prototyping, proceedings of DIS'02, Designing Interactive Systems, ACM Press, New York, pp 424-433, 2000.

Bødker, Susanne, Pelle Ehn, J. Kammersgaard, Morten Kyng and Yngve Sundblad (1987). A UTOPIAN Ex-perience: On Design of Powerful Computer-Based Tools for Skilled Graphical Workers. Computers and Democracy - A Scandinavian Challenge. Gro Bjerknes, Pelle Ehn and Morten Kyng. Aldershot, England: Avebury: 251-278.

Bresin, R. and Friberg, A. (2000). Emotional coloring of computer-controlled music performances, Computer Music Journal 24(4), 44-63.

Bretan, I. (1995) Natural Language in Model World Interfaces. Licentiate thesis, Stockholm University, Report no 95-017.

Buisine, S., Abrilian, S., and Martin, J-C. (2004). Evaluation of multimodal behaviour of embodied agents. In Z. Ruttkay and C. Pelachaud, editors, From Brows till Trust: Evaluating Embodied Conversational Agents). Kluwer.

De Carolis, B., M Bilvi, C Pelachaud (2003). APML, a mark-up language for believable behavior generation. In H Prendinger and M Ishizuka (Eds): *Life-like Characters. Tools, Affective Functions and Applications*. Springer, 2003

Cassell, J., and Thorisson, K. (1999). The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. Applied Artificial Intelligence, 13(3).

Catrambone, R., Stasko, J. and Xiao, J. (2004). ECA as user interface paradigm. In Z. Ruttkay and C. Pelachaud, editors, From Brows till Trust: Evaluating Embodied Conversational Agents. Kluwer.

Cavalluzzi, A. V Carofiglio and F de Rosis (2004). Affective Advice-Giving Dialogs. ISCA Tutorial and Research Workshop on "Affective Dialogue Systems" (ADS04). Kloster Irsee, Germany, June 2004

Christoph, N. (2004). Empirical evaluation methodology for embodied conversational agents. In Z. Ruttkay and C. Pelachaud, editors, From Brows till Trust: Evaluating Embodied Conversational Agents. Kluwer,

Cohen, M.M., and Massaro, D.W. (1993). Modeling coarticulation in synthetic visual speech. In M. Magnenat-Thalmann and D. Thalmann, editors, Models and Techniques in Computer Animation), pages 139--156, Tokyo, Springer Verlag.

Cooper, A. (1999). The Inmates Are Running the Asylum: Why High-Tech Products Drive Us Crazy and How to Re-store the Sanity, SAMS, 1999.

Cooper, B. Brna, P. and Martins, A. (2000). Effective Affective in intelligent systems - building on evidence of empathy in teaching and learning, In Paiva, A. and Martinho, C., (eds.), Affective Interactions: Towards a New Generation of Computer Interfaces, pages 21-30. Proceedings of Workshop at i3 Annual Conference in Sienna, Italy, 2000.

Cosi, P., Magno Caldognetto, E., Perin, G., and Zmarich, C. (2002). Labial coarticulation modeling for realistic facial animation. In Proceedings of ICMI 2002, Pittsburgh, PA, USA, October 14-16, 2002.

Cowie R. and Cornelius R. (2003) Describing the emotional states that are expressed in speech. Speech Communication 40, 5—32, 2003.

Csíjszentmihályi, M., (1990). Flow: The Psychology of the optimal experience. Harper and Row, NY.

Dahlbäck, N., Jönsson, A. and Ahrenberg, L. (1993). Wizard of oz studies-why and how. In Proc. ACM International Workshop on Intelligent User Interfaces, ACM Press.

Darves, C., and Oviatt, S. (2004). Talking to digital fish. In Z. Ruttkay and C. Pelachaud, editors, From Brows till Trust: Evaluating Embodied Conversational Agents. Kluwer.

Ebling, M.R. and John, B.E. (2000). On the contributions of different empirical data in usability testing. Pp. 289-296 in Proceedings of the conference on Designing Interactive Systems: Processes, practices, methods and techniques.

Ekman, Paul (1993). Facial expression of emotion. In American Psychologist, 48:384—392.

Elliott, Clark. (1992) The Affective Reasoner: A process model of emotions in a multi-agent system. Institute for the Learning Sciences Tech. Report #32.

Elliott, Clark (1997) Hunting for the Holy Grail with "emotionally intelligent" virtual actors, ACM Intelligence Magazine.

Elliot Clark and Brzeinski, Jacek (1998) Autonomous Agents as Synthetic Characters, Communications of the ACM.

Fagerberg, P., Ståhl, A. and Höök, K. (2003) Designing gestures for affective input: an analysis of shape, effort and valence. In Proceedings of the conference on Mobile Ubiquitous Multimedia (MUM'03), Norrköping, Sweden, ACM Press.

Fagerberg, P., Ståhl, A. and Höök, K. (forthcoming) Affective Messages – Emotionally Engaging Interaction. To appear in Journal of Personal and Ubiquitous Computing, Special Issue on Tangible Interfaces in Perspective, Springer (to appear).

Garau, M., Slater, M., Vinayagamoorhty, V., Brogni, A., Steed, A., and Sasse, M.A. (2003). The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In Proceedings of the SIG-CHI conference on Human factors in computing

Gaver, W, Beaver, J and Benford, S. (2003). Ambiguity as a Resource for Design, in proceeding of CHI 2003, Com-puter Human Interaction, Fort Lauderdale, April, 2003.

Green, W.S. and Jordan, P.M. (2002). Pleasure with Products: Beyond Usability. London: Taylor and Francis.

Guenter, B., Grimm, C., Malvar, H., and Wood, D. (1998). Making faces. In Computer Graphics Proceedings, Annual Conference Series. ACM SIGGRAPH, 1998.

Haptek website: http://www.haptek.com

Hazlett, R. (2003). Measurement of user frustration: A biologic approach. Pp. 734-735 in CHI 03 Proceedings.

He, L, M. F. Cohen, D. H. Salesin. (1996) *The Virtual Cinematographer: a paradigm for automatic real-time camera control and directing* in Proceedings of SIGGRAPH 96, Computer Graphics Proceedings, Annual Conference Series, 217-224.

Helander, M. (eds) (1988). Handbook of HCI. North-Holland: Amsterdam.

Hornung, A., Lakemeyer, G., Trogemann, G., (2003) *An Autonomous Real-Time Camera Agent for Interactive Narratives and Games*, Lecture Notes in Artificial Intelligence, vol. 2792, pp. 236 – 243.

Höök, K. (2004). User-centred design and evaluation of affective interfaces. In Z. Ruttkay and C. Pelachaud, editors, From Brows till Trust: Evaluating Embodied Conversational Agents. Kluwer,

Höök, K., Bullock, A., Paiva, A., Vala, M., Chaves, R., and Prada, R. (2003). FantasyA and Sen Toy. In Proceedings of the conference on Human factors in computing systems, pages 804—805, Ft. Lauderdale, Florida, USA, 2003. ACM Press.

Höök, K., Persson, P., and Sjölinder, M. (2000). Evaluating users' experience of a character-enhanced information space. Journal of AI Communications, 13(3):195—212.

Höök, K., Sengers, P., and Andersson, G. (2003). Designing design: Sense and sensibility: evaluation and interactive art. Pp. 241-248 in CHI 2003 Proceedings.

Iacucci, G, Iacucci, C, and Kuutti, K. (2002) Imagining and experiencing in design, the role of performances, Proc. Of the second Nordic Conference on Human-Computer Interaction (NordiCHI), Aarhus, 2002.

Iacucci, G. and Isomursu, M (2004). Facilitated and Performed "Happenings" as Resources in Experience Design of Ubiquitous Computing, Presented at a workshop held in conjunction with Computer Human Interaction (CHI) conference in Vienna, April, 2004.

Isbister, K. and Doyle, P. (2003). Web guide agents: Narrative context with character. Pp. 229-243 in (M Mateas and P. Sengers, Eds.) Narrative Intelligence, Amsterdam: John Benjamins Publishing Company.

Isbister, K., and Doyle, P. 2004. Evaluating interdisciplinary ECA research. In Z. Ruttkay and C. Pelachaud, editors, From Brows till Trust: Evaluating Embodied Conversational Agents. Kluwer,

Isbister, K. and Nass, C. 2000. Consistency of personality in interactive characters: Verbal cues, non-verbal cues, and user characteristics. International journal of human-computer studies 53(2), 251-267.

Krenn B., Grice M., Piwek P., Schröder M., Klesen M., Baumann S., Pirker H., van Deemter K., Gstrein E. (2002). Generation of Multi-modal Dialogue for Net Environment, In Proceedings of KONVENS-02, 30 September - 2 October 2002, pp.91-98, Saarbrücken, Germany.

Jordan, P.W. (2001). Pleasure with Products: Human Factors for Body, Mind, and Soul. Pp. 206-217 in (W.S. Green and P.W. Jordan, Eds.) Human Factors in Product Design: Current Practice and Future Trends, London: Taylor and Francis.

E. Krahmer and M. Swerts (forthcoming) More about brows. In Z. Ruttkay and C. Pelachaud, editors, From Brows till Trust: Evaluating Embodied Conversational Agents. Kluwer, To appear.

B. Krenn, B. Neumayr, C. Schmotzer, and M. Grice (forthcoming). What we can learn from avatar-driven Internet communities. In Z. Ruttkay and C. Pelachaud, editors, From Brows till Trust: Evaluating Embodied Conversational Agents. Kluwer, To appear.

Laaksolahti, J., Bergmark, N., Hedlund, E., (2003). Enhancing Believability Using Affective Cinematography. LNAI 2792, pp. 264 – 268.

E.-J. Lee and C. Nass.(1998) Does the ethnicity of a computer agent matter? An experimental comparison of human-computer interaction and computer-mediated communication. In WECC'98, The First Workshop on Embodied Conversational Characters, October 1998.

Y. Lee, D. Terzopoulos, and K. Waters (1995). Realistic modeling for facial animation. In Computer Graphics Proceedings, Annual Conference Series), pages 55--62. ACM SIGGRAPH, 1995.

B. LeGoff (1997). Synthese a partir du texte de visage 3D parlant francais. PhD thesis, Institut National Polytechnique, Grenoble, France, 1997.

McBreen, H., Morton, H., and Jack, M. (2004). Experimental evaluation of the use of ECAs in e Commerce applications. In Z. Ruttkay and C. Pelachaud, editors, From Brows till Trust: Evaluating Embodied Conversational Agents. Kluwer.

H. McBreen, J. Anderson, and M. Jack (2001). Evaluating 3 D embodied conversational agents in contrasting VRML retail applications. In Multimodal Communication and Context in Embodied Agents, Proceedings of the AA'01 workshop, Montreal, Canada, May 2001.

D. Massaro (1997). Perceiving Talking Faces: From Speech Perception to a Behavioral Principle. Bradford Books Series in Cognitive Psychology. MIT Press, 1997.

Mentis, H. and Gay, G. (2003). User recalled occurrences of usability errors: Implications on the User Experience. Pp. 736-737, CHI 03 Proceedings.

Morkes, J., Kernal, H.K., and C. Nass (1998). Humor in task-oriented computer mediated communication and human computer interaction. Pp. 215-216, CHI 98 Proceedings.

Muller, M. and Kuhn, S. (1993). Participatory Design. In Communications of the ACM 36(6), 24-28.

Myers, David G. (2002). Intuition: Its Powers and Perils. Yale University Press.

Nass C., Li Gong (2000). Speech interfaces from an evolutionary perspective. Communications of the ACM 43(9) (2000) 36—43

C. Nass, E. Robles, and Q. Wang (forthcoming). Apparent attention. In Z. Ruttkay and C. Pelachaud, editors, From Brows till Trust: Evaluating Embodied Conversational Agents. Kluwer, To appear.

Paiva, A., Costa, M., Chaves, R., Piedade, M., Mourão, D., Sobral, D., Höök, K., Andersson, G., and Bullock, A. (2003). SenToy: an Affective Sympathetic Interface, International Journal of Human Computer Studies, Volume 59, Issues 1-2, July 2003, Pages 227-235, Elsevier.

Picard, R. (1997). Affective Computing. Cambridge, MA: MIT Press.

F. Pighin, J. Hecker, D.H. Salesin, D. Lischinski, and R. Szeliski (1998). Modeling realistic facial expressions from photographs. In Computer Graphics Proceedings, Annual Conference Series). ACM SIGGRAPH, 1998.

Piwek P., Krenn B., Schröder M., Grice M., Baumann S., Pirker H. (2002). RRL: A Rich Representation Language for the Description of Agent Behaviour in NECA, In Proceedings of the Workshop "Embodied conversational agents - let's specify and evaluate them!", held in conjunction with AAMAS-02, July 16 2002, Bologna, Italy.

Preece, J., Rogers, Y. Sharp, H., Benyon, D., Holland, S. and Carey, T. (1994) Human-Computer Interaction, Addison Wesley Part VI.

Reeves, B. and Nass, C. (1996). The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. Stanford, CA: CSLI Publications.

L. Reveret, G. Bailly, and P. Badin (2000). MOTHER: A new generation of talking heads providing a flexible articulatory control for video-realistic speech animation. In X. Tang B. Yuan, T. Huang, editor, Proceedings of ICSLP'00: International Conference on Spoken Language Processing), volume II, pages 755—758, Beijing, China, 1996.

F de Rosis and F Grasso, (2000). Affective Natural Language generation. In A Paiva (Ed): *Affective Interactions*. Springer Verlag.

Russell, J.A. (1980). A Circumplex Model of Affect, Journal of Personality and Social Psychology 39(6), pp. 1161—1178, American Psychological Association.

Ruttkay Z.M. et al. (2000) A facial repertoire for avatars, in Proceedings of the "Learning to Behave: Interacting Agents" workshop, CELE-Twente Workshops on Natural Language Technology, Twente University, October 2000.

Z. Ruttkay, C. Dormann, and H. Noot (forthcoming). ECAs on a common ground. In Z. Ruttkay and C. Pelachaud, editors, From Brows till Trust: Evaluating Embodied Conversational Agents. Kluwer, To appear.

Schank, Roger (1991). Where's the AI?, AI Magazine, 12(4), pp. 38—48.

Scherer, K. R. and Ceschi, G. (2000). Criteria for emotion recognition from verbal and nonverbal expression: Studying baggage loss in the airport. Personality and Social Psychology Bulletin, 26(3), 327-339.

M. Schröder (to appear). Speech and Emotion Research: An overview of research frameworks and a dimensional approach to emotional speech synthesis. PhD thesis, Institute of Phonetics, Saarland University. To be published.

Sengers P. (1998). Anti-Boxology: Agent Design in Cultural Context. Carnegie Mellon University Department of Computer Science, PhD Thesis (1998)

C. Siciliano, G. Williams, J. Beskow, and A. Faulkner, (2003). Evaluation of a multilingual synthetic talking face as a communication aid for the hearing impaired. In Proc. Intl. Cong. of Phonetic Sciences (ICPhS'03)), 2003.

Suchman, L. (1987). *Plans and Situated Actions: The Problem of Human-Machine Communication*, New York: Cambridge University Press.

Suchman L. A. (1997) From Interactions to Integrations. Howard S., Hammond J., Lindegaard G. (eds.): *Proceedings of Human-Computer Interaction INTERACT'97*. Chapman & Hall.

Suchman, L. (2002). Replicants and Irreductions: Affective encounters at the interface, published by the Centre for Science Studies, Lancaster University, Lancaster LA1 4YN, UK, at http://www.comp.lancs.ac.uk/sociology/papers/Suchman-Replicants-and-Irreductions.pdf

Thomas, F. and Johnston, O. (1981). *Disney Animation: The Illusion of Life*, Abbeville Press.

Tiger, L. (2002). *The Pursuit of Pleasure,* Transaction Publishers. New Brunswick (USA) and London (UK)

Wiberg, C., (2002). A Measure of fun: Extending the scope of web usability. Unpublished PhD Thesis, Umeå University

Y. Wu, N. Magnenat-Thalmann, and D. Thalmann, (1995). A dynamic wrinkle model in facial animation and skin ageing. The journal of visualization and computer animation), 6:195--205, 1995.

Zhou, X., and Conati, C. (2003). Inferring user goals from personality and behavior in a causal model of user affect. Pp. 211—218, Proceedings of the 8th International Conference on Intelligent User Interfaces.

B. General bibliography

Survey/Questionnaire Instruments (standard and custom measures)

Campbell, D.T. and Stanley, J.C. 1963. Experimental and Quasi-Experimental Designs for Research. Boston: Houghton Mifflin Company.

Carmines, E.G. and Zeller, R.A. 1979. Reliability and Validity Assessment. Newbury Park: Sage Publications.

McIver, J.P. and Carmines, E.G. 1981. Unidimensional Scaling. Newbury Park: Sage Publications.

Singer, E. and Presser, S. 1989. Survey Research Methods: A Reader. Chicago: The University of Chicago Press.

Stempel, G.H. and Westley, B.H. 1989. Research Methods in Mass Communication. New Jersey: Prentice Hall.

Sudman, S. and Bradburn, N.M. 1982. Asking Questions: A Practical Guide to Questionnaire Design. San Francisco: Jossey-Bass Publishers.

Weiss, A.J. and Leets, L.L. 1998. Introduction to Statistics for the Social Sciences. New York: The McGraw-Hill Companies, Inc.

Participatory Design Tactics

Beyer, H. and Holtzblatt, K. 1995. Apprenticing with the Customer: A Collaborative Approach to Requirements Definition. In Communications of the ACM, May.

Binder, T. and Buur, J. 1999. Video as Design Material in Collaborative Design. School of Art and Communication, Malmo.

Suri, J.F. 2002. Designing Experience: Whether to Measure Pleasure or Just Tune In? Pp. 161-174 in Pleasure with Products: Beyond Usability (W.S. Green and P.W. Jordan, Eds.). London: Taylor and Francis.

Ethnography

Fetterman, D. 1998. Ethnography Step-by-step (2nd Edition). Sage Press.

Henley, P. 1998. Film-making and Ethnographic Research. In Image-Based Research: A Sourcebook for Qualitative Researchers. Prosser, J. (Ed.). London: Falmer Press.

Taylor, K. 2002. Using Video Ethnography to Inform and Inspire User-Centered Design. Pp. 175-187 in Pleasure with Products: Beyond Usability (W.S. Green and P.W. Jordan, Eds.). London: Taylor and Francis.

Van Maanen, J. 1988. Tales of the Field: On Writing Ethnography. Chicago: The University of Chicago Press.

Social Psychological and Communication Theory

Bandura, A. 1997. Self-Efficacy: The Exercise of Control. New York: W.H. Freeman and Company.

Bandura, A. 1977. Social Learning Theory. Englewood Cliffs, NJ: Prentice Hall.

Bryant, J. and Zillman, D. 1991. Responding to the Screen: Reception and Reaction Processes. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Cialdini, R.B. 1993. Influence: The Psychology of Persuasion. New York: Quill, William Morrow.

Csikszentmihalyi, M. 1990. Flow: The Psychology of Optimal Experience. New York: Harper Perennial.

Dourish, P. (2001) *Where the action is. The Foundations of embodied Interaction*, MIT Press, Cambridge, MA, USA.

Fiske, S.T. and Taylor, S.E. 1991. Social Cognition. New York: McGraw-Hill, Inc.

Greene, J.O. and Burleson, B.R. 2003. Handbook of Communication and Social Interaction Skills. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Knapp, M.L. and Hall, J.A. 2002. Nonverbal Communication in Human Interaction. Australia: Wadsworth Thomson Learning.

Petty, R.E. and Cacioppo, J.T. 1981. Attitudes and Persuasion: Classic and Contemporary Approaches. Dubuque, Iowa: Wm C Brown Company Publishers

Reeves, B. and Nass, C. 1996. The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. Stanford, CA: CSLI Publications.

Zimbardo, P.G. and Leippe, M.R. 1991. The Psychology of Attitude Change and Social Influence. New York: McGraw-Hill, Inc.