Chapter 10
Evaluating User Experience Factors Using Experiments: Expressive Artificial Faces Embedded in Contexts

Michael Lankes, Regina Bernhaupt, and Manfred Tscheligi

Abstract There is an ongoing debate on what kind of factors contribute to the general positive user experience while playing a game. The following chapter introduces an experimental setting to measure user experience aroused by facial expression of embodied conversational agents (ECAs). The experimental setup enables to measure the implications of ECAs in three contextual settings called "still," "animated," and "interaction." Within the experiment, artificially generated facial expressions are combined with emotion-eliciting situations and are presented via different presentation platforms. Stimuli (facial expressions/emotion-eliciting situations) are assembled in either consonant (for example, facial expression: "joy," emotion-eliciting situation: "joy") or dissonant (for example, facial expression: "joy," emotion-eliciting situation: "anger") constellations. The contextual setting called "interaction" is derived from the video games domain, granting an interactive experience of a given emotional situation. The aim of the study is to establish a comparative experimental framework to analyze subjects' user experience on emotional stimuli in different context dimensions. This comparative experimental framework utilizes theoretical models of emotion theory along with approaches from human–computer interaction to close a gap in the intersection of affective computing and research on facial expressions. Results showed that the interaction situation is rated as providing a better user experience, independent of showing consonant or dissonant contextual descriptions. The "still" setting is given a higher user experience rating than the "animated" setting.

10.1 Introduction

Various methods and new methodological developments have been proposed to evaluate user experience in application domains ranging from user experience evaluation of mobile phones (Roto and Rautava 2008) to user experience for interactive

M. Lankes (✉)
Upper Austria University of Applied Sciences, Hagenberg, Austria
e-mail: michael.lankes@fh-hagenberg.at

R. Bernhaupt (ed.), *Evaluating User Experience in Games*, Human-Computer Interaction Series, DOI 10.1007/978-1-84882-963-3_10,

TV (Bernhaupt et al. 2008b) and several others (e.g. Law et al. 2007). Most of them did not take into consideration recent developments in the area of gaming, such as game play between thousands of players, multiplayer audio channels and the use of novel input devices to encourage physical activity (Bernhaupt et al. 2008a).

Digital games constitute a tremendously varied set of applications, with a wide range of associated player experiences, defying a one-size-fits-all approach to their conceptualization and measurement. One of the main challenges facing the gaming research community is a lack of a coherent and fine-grained set of methods and tools that enable the measurement of entertainment experiences in a sensitive, reliable, and valid manner. Taking a factor-structure approach to characterize user experiences, terms like fun, flow, and playability are most often used to explain user experience in game design. However, there is an open discussion to include other factors which might have relevance for games. Emotion is often cited as a key element of user experience (e.g., Hassenzahl and Tractinsky 2006).

On the other hand, the quality of the display of emotions portrayed by embodied conversational agents (ECAs) is perceived as a necessity to improve the user experience (Lee and Marsella 2006). The following chapter is looking in detail on the relation of user experience and emotions that are expressed by ECAs. It shall provide some insights concerning the relation between emotions displayed in a game (through the characters) in conjunction with emotion-eliciting situations in regard to the user experience. Results of this study are a first step in a series of experiments investigating the relationship between user experience and interaction with embodied conversational agents.

The chapter is organized as follows: Based on an overview on currently used models in emotion theory, the usage of emotional (factor) models in games is explained and how the modeling of nonplayer characters (NPCs) – a field of application of ECAs – is representing these emotional (factor) models. It shall be investigated how the display of emotions (still, animated, or interactive) in various contexts (either with a consonant or with a dissonant context description) is affecting the user experience (measured with a questionnaire). The section on the experimental study describes in detail three prestudies for selecting stimuli and material followed by the main study on the relation of displaying emotion in games via ECAs and perceived user experience. The conclusion shows how the findings can be used in game development in terms of designing positive user experiences.

10.2 Related Work

Incorporating emotional expressions for nonplayer characters in games is seen as an appropriate way to improve the gamers' experiences. The research area of emotions is a central topic in human–computer interaction and is approached from various perspectives. Subsequently, some of these perspectives, and their relation to current developments in user experience research, are presented. First, we look at how emotions can be integrated, seen either as part of the computing system or as part of the overall user experience. Second, we look on how emotions can be measured,

and third, we present some related work on how emotions are integrated into ECAs and how we measure the user experiencing in regard to the emotional expressions performed by ECAs in conjunction with emotion-eliciting situations.

10.2.1 General Description on Emotion

The implementation of emotional factors in systems received an increased interest by the human–computer interaction (HCI) community as researchers within this field aim to develop machines that are focused on human needs (Branco 2003). Emotions play a crucial role in our everyday life with computers (Crane et al. 2007) and have a significant impact on user experience as they influence actions, expectations, and future evaluations (Picard 1997). Technological advancements enable machines to perceive, interpret, express, and respond to emotional information. Traditionally, emotional factors were neglected as designers focused on usability aspects and developed systems with the aim to increase efficiency of required tasks (Picard et al. 2002). Although it might be argued that machines should be treated as mere tools that do not (or should not) require any emotions, results of Reeves and Nass (2003) showed that people tend to exhibit social and emotional behaviors toward machines. Picard et al. (2002) also note that interaction with machines is emotional even if the system was not designed to incorporate emotional aspects. Users should be enabled to utilize familiar communication mechanisms when interacting with computational systems. The human–machine interaction process should be designed to resemble human interpersonal interactions, in order to rely on skills obtained from human–human communication. Systems get easier to use if the interaction between human–machine is similar to human–human interaction (Bernhaupt et al. 2007a).

When dealing with the various objectives within HCI research in the field of emotion, we can choose from a tremendous amount of research approaches. Mahlke (2005) provides a taxonomy dividing emotion in HCI into affective computing and emotional design. The concept of affective computing postulates to develop systems that are able to perceive the emotional state of the user, interpret the affective state, adapt to the user's state, and generate an expressed emotion (Minge 2005). Emotional design claims that emotion is considered as an important factor of the user's experience with interactive systems and it is aimed to incorporate emotional aspects in the interactive system design process (Norman 2002). From the perspective of User Experience (UX) research, emotions are investigated to understand their role as antecedent, as a consequence and a mediator of technology use (Hassenzahl and Tractinsky 2006). Researchers in the field of user experience evaluation thus try to concentrate on integrating emotional processes of the user experience into the evaluation procedure of the interactive systems.

Our experiment addresses the factor emotion concerning user experience by raising the questions how emotional stimuli (facial expressions by ECAs and emotion-eliciting situations) in interactive system affect the (more general) user experience?

Concerning the factor emotion, there are two major research foci: the assessment of emotional dispositions aroused by games and the incorporation of emotion into the game world (see the following two sections).

10.2.2 Games and User Experience

Several tools are available to investigate the factor emotion: the Self Assessment Mannequin (SAM) (Fischer et al. 2002), Emoticons (Desmet and Hekkert 2002), or Affective Grid (Russell and Fernandez-Dols 1997) are just a few examples. Until now, no commonly accepted method for measuring emotions is available. Ravaja and colleagues (2006) presented 37 subjects different types of computer games (Tetris, James Bond, Nightfire, and others). To measure emotional response patterns they employed categorical (fear, joy, etc.) and dimensional measurement methods (arousal and valence dimensions). They conclude that different types of games elicit different types of emotional dispositions. Furthermore, the researchers believe that developers will increase the commercial success of a game by incorporating emotional aspects while testing different computer game concepts. Pleasant emotional episodes during game play are deemed to be an indicator to provide positive (and desirable) user experience (Ermi and Mäyrä 2005).

To understand the overall user experience, we decided to focus on a general perception instead of only looking at the elicited emotion. User experience in games is evaluated using a large variety of approaches ranging from questionnaires to physical measurements (Mandryk et al. 2006). As we wanted to have a simple and flexible to use measurement, we decided to measure user experience with the AttrakDiff questionnaire (www.attrakdiff.de) that has been used in various studies to investigate pragmatic and hedonic quality of users interacting with a system.

The AttrakDiff questionnaire was developed to measure implications of attractiveness of a product. Users indicate their impression of a given product by bipolar terms that reflect four dimensions. The first dimensions, the pragmatic quality (PQ), describes traditional usability aspects, while the dimension Hedonic Quality-Stimulation (HQ-S) refers to the need of people for further development concerning themselves. By supporting this aspect, products can offer new insights and interesting experiences. Hedonic Quality-Identification (HQ-I) allows to measure the amount of identification a user has toward a product. Pragmatic and hedonic dimensions are independent from each other and share a balanced impact on the overall judgment. The two aspects contribute equally to the overall judgment of the situation/product and is referred to as hedonic quality (HQ). Attractiveness (ATT) resembles an overall judgment based on the perceived quality.

10.2.3 Embodied Conversational Agents

According to Bartneck (2000), computer games were one of the first applications that incorporated interactive virtual characters. One main driving force in the games industry is innovation in computer technology, which enables the development of

more visually elaborated game entities (here: characters). As video game systems have become more powerful from a technical point of perspective, the gaming community has demanded games that push the technical capabilities of the platforms (Pruett 2008). At the beginning of video game, history game elements were displayed as very abstract and simple forms, while nowadays players are confronted with rather highly realistic virtual actors inhabiting complex virtual worlds. A lot of effort is put in the creation of NPCs by game companies. NPCs can include capabilities of verbal and nonverbal communication and may aid the player in a gaming situation. Players may encounter NPCs as enemies that try to interfere to reach game goals, or as characters that serve them as tutors or supporters. Isbister introduces in her book "Better Game Characters by Design" (Isbister 2006) a classification of NPCs based on their social roles within the game.

NPCs can be seen as a field of application regarding ECAs. In general terms, an embodied agent can be understood as a specific type of agent whose behaviors are executed by some type of perceivable digital representation (Bailenson 2008). Lieberman (1997) describes agents, in contrast to traditional interfaces, as any program that serves as an assistant or helper to aid users during the interaction process. Bates (1994) adds emotional aspects when defining embodied agents. Nonverbal signals form an essential part in the communication process, which incorporate the portraying of emotional dispositions via facial expressions, gestures, voice, etc. With the implementation of emotional aspects, agents are more attractive to users because they communicate in ways we are used to (Elliott and Brzezinski 1998). Agents containing knowledge about the conversational process and capabilities to perceive and express emotional signals can be summarized under the term ECAs. They are characters that visually incorporate, or embody, knowledge about the conversational process (Prendinger and Ishizuka 2004). ECAs are virtual humans able to perform conversations with humans by both understanding and producing speech and nonverbal signals (Cassell 2008). They form a type of multimodal interface where the modalities are the natural communication channels of human conversation. The visual representation of ECAs of interacting is intrinsic to its function, meaning that visual information (for example, display of facial expressions) is crucial in the process (Bickmore and Cassell 2001). Nonverbal channels are necessary for both conveying information and regulating the communication process (Bickmore and Cassell 2001). They can be utilized to provide social cues as attentiveness, positive affect, and attraction. For investigating the affect of displayed emotions on the users (players) experience, the definition by Mancini et al. (2004, p. 1) shall serve as the basis: "ECAs are virtual embodied representations of humans that communicate multimodal with the user (or other agents) through voice, facial expression, gaze, gesture, and body movement."

10.2.4 Facial Expressions Performed by Embodied Conversational Agents

Emotion theory offers a variety of approaches including perspectives of social constructivism, cognition, or theories based on the work of William James or Charles

Darwin (Cornelius 1996). The "Darwinian approach" focuses on facial expressions and propagates a limited number of basic, fundamental, or discrete emotions that are directly linked to the motivational system (Scherer et al. 2004). Followers of this tradition assume that specific eliciting conditions would automatically trigger a pattern of reactions such as peripheral physiological responses. It is postulated that mechanisms of emotion mixing or blending occur, which lead to a great variability of facial expressions. Russell and Fernandez-Dols (1997) summarized the discrete emotion approach to outline basic assumptions. First, there is a small set of basic (or fundamental) emotions that are genetically determined and discrete. Each of these emotional states is composed of behavioral patterns like the portrayal of specific facial expressions. The encoding and decoding of emotional signals developed based on adoption processes. States that are not linked to facial signals are not considered as basic emotions. Evidence is present for the basic emotions happiness, surprise, anger, contempt (some uncertainty), disgust, sadness, and fear. These emotions are recognized by all humans (innate) independent from their cultural background. Emotions that share nonfundamental states are considered to be blends (mixtures) of basic emotions. Cultural restrictions may inhibit or mask certain behavioral patterns called display rules.

Based on these assumptions, Ekman and Friesen (1972) developed Facial Action Coding System (FACS). It serves as a high-level description of motions by feature points (Jaimes and Sebe 2007). Each facial muscle is assigned a numeric value that is modified when muscles move. Thus, facial expressions could be synthesized by relating to FACS codes. FACS allows measuring facial expressions objectively, which enables the synthesis of specific expressions by applying the required FACS codes. Movement of individual facial muscles sections lead to observable alternations within the overall appearance of the face.

Fernandez-Dols and Carroll (1997) emphasized the importance of context, claiming that the perception of emotional signals is significantly influenced by situational factors and vice versa. Wallbott (1990) also supports this position by noting that subjects were confronted with isolated stimuli to indicate their perception of the presented emotion. He propagated the explicit incorporation of context-related information in the investigation of facial expression. Without context, subjects are forced to simulate (or construct) the missing information, which inevitably will lead to invalid research results. According to Wallbott (1990), three factors are relevant when judging the emotional quality: the stimulus (for example, photos showing facial expressions), the background (or context), and the emotional disposition of the observer. Contextual aspects are not only embedded in emotions, but also the cause for emotional dispositions. Context in facial expression can be subdivided into a situation-related context (modification of the current emotion), a comparative context (the relation of one nonverbal communication channel between others), the static context (captured via photos), and the dynamic context (involved channels in a given time frame).

To summarize the findings of Wallbott (1990), person-related aspects (here: facial expressions) have more influence on emotion judgments than situational components. However, the analyzed data revealed that (although visual stimuli are

dominating the perception) descriptions of emotion-eliciting situations will gain importance if the constellation of stimuli is dissonant. Furthermore, subjects employ different strategies when being confronted with different types of stimuli constellations. Person-related aspects do not completely dominate situational factors, as they are always integrated into the judgment of emotional stimuli. An important factor that determines the importance of components is the type of presentation medium. The increase of visualization in regard to situational aspects leads to a shift of dominance. The more visualized a situation is presented, the more it will influence the judgment on emotion. Wallbott (1990) assumes that dynamic stimuli material (descriptions of emotion-eliciting situations) in "still" settings (presentation of facial expressions via film clips) grants more clear information on a given situation than static presentations (presentation of facial expressions via photos).

The following proposal for evaluating UX in the context of ECAs will build on these findings as it tries to extend the framework by employing a new (interactive) presentation medium. The relative importance of information channels shall not be addressed, but a novel experimental setting to investigate the perceived user experience in regard to facial expressions and their relation to contextual aspects shall be presented. The introduced theoretical considerations on emotion should serve as a foundation, as well as to provide some insights into this multidimensional research topic. It should have been pointed out that situational aspects have to be considered when investigating the perception of facial expressions as they determine the quality of the interpretation process.

10.3 Evaluation

The goal of this experiment was to understand the influence of emotional facial expressions of ECAs and descriptions of emotion-eliciting situations in three interaction conditions (still, animated, or interactive) on the user experience. We see user experience as a concept that is best described as a property of the human interacting with the game. The overall user experience during game play is consisting of some key components. Emotions are the most prominent component together with immersion, playability, or flow. As we only wanted to understand how changes in the emotional expression of an ECA and emotion-eliciting situations might influence the general perception, we decided to focus on a general measurement of user experience, based on the AttrakDiff questionnaire.

10.3.1 Methodological Considerations

To investigate user experience in games, a set of methods has been developed. Following traditional HCI approaches of classifying evaluation, methods can be grouped in expert- and user-oriented evaluations (Dix et al. 2004), other classifications are based on development cycles or more social science-oriented approaches (Bernhaupt et al. 2007b). How emotional expressions of ECAs and

emotion-eliciting situations affect the user experience in a game can be evaluated using several of these methods. But what the relationship between these methodologies is stays rather unclear.

To understand the influence of emotional expressions of ECAs and the influence of emotion-eliciting situations on UX, a more rigid approach is necessary. An experiment (including three prestudies) was set up to investigate the relationship between emotional expressions of ECAs and emotion-eliciting situations in interactive settings compared to still and animated settings.

The main study investigates how a given emotional facial expression and an emotion-eliciting situation combined in either a consonant (for example, facial expression: "joy," emotion-eliciting situation: "joy") or a dissonant (for example, facial expression: "joy," emotion-eliciting situation: "anger") constellations presented in either a still, animated or an interactive format is influencing the overall user experience. The goal was to investigate the influences on the overall user experience to understand how the design of ECAs influences the game play.

10.3.2 Prestudy 1: Evaluation of Emotion-Eliciting Situations

The goal of prestudy 1 was to identify and validate emotion-eliciting situations. The purpose was to identify emotion-eliciting situations with "pure" emotions (weak or no presence of other emotions) and high intensity that will be utilized in the main study. As emotion descriptions set up by the researchers influence heavily the outcome (Wallbott 1990, p. 37), a categorized and standardized emotion-eliciting situation experienced in real life was used. Projects (for example, Summerfield and Green 1986, Scherer et al. 2004) were carried out for years in different cultures to identify emotion-eliciting situations that are culturally independent from their meaning. The "International Survey on Emotion Antecedents And Reactions" (ISEAR) database (ISEAR 2008), which was made freely available for researchers interested in this field, contains data files and explanations for a major cross-culturally comparative study on the cognitive antecedents of emotion (based on appraisal notions) and the reaction patterns reported for seven basic emotions (joy, fear, anger, sadness, disgust, shame, and guilt) by close to 3000 respondents in 37 countries.

We used 200 randomly chosen database entries as a basis and then removed descriptions that did not refer to the emotion categories of Summerfield and Green (1986). Within prestudy 1, the applicable descriptions were filtered using three criteria. Criterion 1 identify the dominating emotions by analyzing the intensity of all six basic emotions. The second criterion should reveal the presence of "pure" (one emotion present) and "blended" (mixture of emotions) emotions, as the questionnaire allowed multiple choice answers. Only pure emotions are considered applicable to the main experiment. The purpose of criterion 3 is to filter out pure emotions that have a fairly low intensity. Descriptions that are employed in the main experiment have to fulfill all three criteria.

Table 10.1 Selection of stimulus material based on prestudy 1, criterion 1. The mean values are shown for each emotion type

Stimulus	Joy	Fear	Anger	Sadness	Disgust	Surprise	Dominance
Aa1	5.76	0.76	0.00	0.10	0.06	1.26	Joy
Aa4	7.10	0.00	0.00	0.00	0.00	0.17	Joy
Ab1	6.16	0.00	0.00	0.00	0.00	0.93	Joy
Ab3	7.20	0.00	0.00	0.00	0.00	0.26	Joy
Bb1	0.00	0.86	0.26	4.76	0.06	1.73	Sadness
Bb2	0.00	0.13	1.13	7.33	0.00	0.43	Sadness
Bb3	0.00	0.33	0.63	7.50	0.00	0.80	Sadness
Bb4	0.00	0.03	0.66	7.26	0.00	0.33	Sadness
Ca4	0.06	3.40	0.00	0.80	0.23	1.86	Fear
Cb1	0.30	4.90	0.26	0.00	0.00	0.10	Fear
Cb3	0.00	6.66	0.86	0.13	0.13	1.33	Fear
Da1	0.00	0.03	5.46	0.63	0.00	1.16	Anger
Da2	0.20	0.00	3.83	0.93	0.96	1.00	Anger
Da4	0.00	0.33	6.40	0.76	0.06	120	Anger
Db1	0.00	0.03	6.20	0.93	0.00	1.13	Anger
Db4	0.00	0.00	6.73	0.80	1.66	0.70	Anger

Thirty participants (15 male, 15 female) aged 22 to 61 took part in the study. To validate the experimental descriptions based on the three criteria, a simple questionnaire was used. Participants rated the evoked emotion (joy, fear, anger, sadness, disgust, and surprise) and the dominance for each description (scale from 0 to 8 [emotion not present at all, emotion intensively present]).

Based on the ratings of the participants, 11 descriptions are applicable for the main experiment. For criteria 1, the dominating emotion was analyzed (see Table 10.1), followed by criteria 3 selecting only emotion-eliciting situations that were rated on average higher than 4.75 (on the scale from 0 to 8) and finally excluding blended emotions (ratings of two emotions that were higher than 2 on average). Based on these criteria, finally 11 situations were showing pure emotions. For the following prestudies, we used two situations for each of the four emotions from the category of Summerfield and Green (1986): sadness, joy, anger, and fear. Table 10.2 shows these eight situations that were used in the following steps of the experiment.

10.3.3 Prestudy 2: Evaluation of Artificial Facial Expressions

It proved to be a difficult undertaking to find appropriate stimuli material mainly due to license or quality issues of available virtual actors. Since no appropriate stimuli material was at hand, it was decided to create the actors and facial expressions. Constructing six ECAs performing four basic emotions along with a neutral one leads to a total number of 30 stimuli images.

Table 10.2 Descriptions selected based on prestudy 1

Stimuli	Description
Aa4	My girlfriend was arriving back from overseas and I picked her up from the airport. She finally appeared from customs and we came into contact again
Ab3	I went back home after a long trip and met beloved people and close friends
Bb3	A close friend was involved in an accident and passed away instantly. He had gone to buy a new car and had asked me to wait at his house so that I could see his new car
Bb4	I hear about the death of somebody I liked very much and I was not present either to see that person or to try to share my emotions with other friends
Cb1	At about midnight I had to go by bike through the city alone. On the whole it was a distance of several kilometers. A car followed me through the streets. Only when I went into a one-way street the car disappeared
Cb3	I was living with my brother and one day he went away on business. I was left alone to look after the house and the property. At night, thieves came and wanted to break into the house
Da1	I had arranged with a friend to go with him to the city by car. We had arranged a place where to meet. I was a bit late and my friend had left already. I had no money to go by train. It was very important for me to go to the city
Db1	The headmaster of the job appointment committee in charge explained me that teacher (of the opposite gender) was more suitable for a particular post. I had more years of service than the male/female

After selecting the emotion-eliciting situations in prestudy 1, it was necessary to construct and investigate facial expressions performed by the constructed ECAs. The stimuli should be presented to subjects without any additional information and should communicate one of the four chosen basic emotions (joy, fear, anger, and sadness). The constructed expressions should convey pure emotions with a rather high intensity. As in the previous study, participants of prestudy 2 rated stimuli by answering via a multiple-choice questionnaire containing the six basic emotions and were rating the emotional intensity on a scale from 0 to 8 (emotion not present at all, emotion intensively present).

In contrast to prestudy 1, the questionnaire was not printed out on paper, but was shown via an LCD display to match the presentation as close as possible in regard to the main experiment. Therefore, a tablet PC was utilized in order to resemble study 1 as close as possible and to grant a certain amount of mobility. With the aid of the lime survey or application (Lime 2008), an online questionnaire tool, the questionnaire was set up and images were implemented along with emotion type and intensity scales. The application also enabled the scrambling of picture order. The evaluation of facial expressions is carried out by applying the three criteria from prestudy 2. This step was necessary to verify if constructed facial expressions are perceived as intended.

Thirty participants took part in prestudy 2 (16 male, 14 female), age ranging from 20 to 63. The filling out of the questionnaire via the tablet PC took about 15–20 minutes. Material was selected based on the same criteria as in prestudy 1. All stimuli were perceived as intended (criterion 1: most intensive emotion), the rates for perceiving blended emotions followed the reported recognition rates of emotions

Fig. 10.1 Female and male ECAs showing the emotions joy, sadness, anger, and a neutral face

in faces (Ekman and Friesen 1972). For criterion 3, the intensity of the emotion was over 4.75 for all presented stimuli. The constructed material thus fulfilled the intended purpose. Figure 10.1 shows examples of female and male ECAs showing different emotions.

10.3.4 Prestudy 3: Evaluation of Settings and Text Fragments

Prestudy 3 deals with the assessment of virtual settings mainly utilized in interactive condition of the main experiment. The virtual settings should indicate the physical context in which emotion-eliciting situations are embedded. The step of creating virtual settings is necessary as conversations in real life take place in physical contexts.

The introduction of a physical context layer may cause unwanted artifacts. The interpretation of facial expressions and emotion-eliciting situations may be influenced by the color of settings. The work of Suk (2006), who investigated emotional responses to color to analyze the relationship between color attribute and emotional dimensions (dimensional approach), helped to overcome this issue. He found out that emotional responses to color vary more strongly with regard to tone than to hue

categories. The overall color of settings and its lightness are slightly toned down to avoid affective-related influence. The blue prints for the construction of virtual settings are derived from the situations of study 1. Most descriptions of prestudy 1 indicate a physical context by containing words such as “airport,” “home,” or “town.” We thus constructed eight contextual settings, one airport setting, two town settings, and five home settings for the eight emotion-eliciting situations. Eight settings were available as text (for the still images), and eight virtual settings were additionally constructed for the interactive setting of the main experiment.

Similar to study 1, the fragments of emotion-eliciting situations should be judged by presenting the six basic emotions, and intensity was rated on a scale from 1 to 8. Thirty participants (14 male, 16 female) aged 22 to 57 took part in the evaluation. The evaluation of text fragments and virtual settings led to a total number of 16 stimuli (eight images of virtual settings plus eight split up emotion-eliciting situations). As in prestudy 2, the questionnaire presents stimuli via an LCD display to match the presentation as close as possible in regard to the main experiment.

Table 10.3 presents the emotion ratings for the eight text fragments describing the context of the eight emotion-eliciting situations (see again Table 10.2) and the rating of the emotional judgment of the eight virtual scenarios (which should not influence the experiment, thus ratings should be below 2.00 on average).

10.3.5 Experiment: Facial Expression and User Experience

The goal of the main experiment was to evaluate the impact emotion-eliciting situations and facial expressions on the overall user experience. We manipulated two conditions: condition one was the influence of a consonant/dissonant contextual descriptions. Condition two was the influence of either a “still” situation (frame), the “animated” (animation clips) situation, or the “interactive” situation (game environment).

The following hypotheses are related to the general research question on how emotional facial expressions of ECAs and emotion-eliciting situations influence the user experience:

- (H1) The overall user experience in the interactive situation is rated higher than the user experience in animation scenario and the still scenario.
- (H2) The animation scenario in terms of user experience is rated higher than the still scenario.
- (H3) The perceived user experience will be the higher for consonant settings than for settings with dissonant stimuli.

The experiment is based on 576 possible stimuli. Four consonant scenarios presenting the four selected basic emotions (joy, fear, anger, and sadness) and 12 dissonant settings, each performed by one out of six possible virtual actors, embedded in eight different emotion-eliciting situations. Since the experimental setup consists

Table 10.3 Ratings of the stimuli for the emotion-eliciting situations and ratings for the virtual settings showing that the design does not influence the emotion elicitation

Stimulus	Joy	Fear	Anger	Sadness	Disgust	Surprise	Stimulus	Joy	Fear	Anger	Sadness	Disgust	Surprise
T_Aa4	6.87	0.40	0.17	0.03	0.70	0.00	V_Aa4	0.53	0.80	0.17	0.63	0.43	0.13
T_Ab3	6.73	0.17	0.20	0.40	1.33	0.20	V_Ab3	1.30	0.23	0.23	0.23	0.63	0.37
T_Bb3	0.80	0.90	0.63	6.57	1.37	0.30	V_Bb3	1.03	0.27	0.07	0.10	0.53	0.10
T_Bb4	0.17	1.03	1.37	6.93	1.37	0.33	V_Bb4	0.37	1.17	0.20	0.33	0.27	0.17
T_Cb1	0.13	5.97	1.23	0.23	0.87	0.27	V_Cb1	1.30	0.23	0.10	0.17	0.77	0.13
T_Cb3	0.10	0.47	5.80	1.27	1.63	0.00	V_Cb3	1.67	0.20	0.17	0.07	0.80	0.10
T_Da1	0.10	0.47	5.80	1.27	1.63	0.00	V_Da1	1.10	0.03	0.23	0.03	0.37	0.03
T_Db1	0.03	0.40	6.83	1.50	1.53	0.83	V_Db1	1.57	0.47	0.23	0.10	1.23	0.27

of three different scenarios, 192 constellations have to be multiplied by 3, resulting in 576 stimuli. Out of these 576, six scenarios are randomly chosen for each subject. Subjects have to indicate the perceived user experience via the AttrakDiff (Hassenzahl et al. 2003) questionnaire.

In order to avoid material effects caused by repeated measurements, items of the AttrakDiff were randomized result in six different versions of the AttrakDiff questionnaire. Furthermore, each participant received the six generated Attrakdiff questionnaires in a randomized order when providing information on their impression concerning user experience factors.

For the experiment, each participant was presented three scenarios (still/animated/interactive) two times (one consonant/one discrepant). Each ECA appeared only one time for each participant. Repetition of consonant and dissonant emotion constellations was avoided and none of the emotion-eliciting situations was repeated for a participant.

Concerning scenario 1 (still), a picture was shown containing a facial expression performed by a virtual actor (evaluated in prestudy 2), along with descriptions of the emotion-eliciting situations (evaluated in prestudy 1). An image of a virtual setting, referenced to the emotion-eliciting situation, is also displayed in the background. After subjects indicated that they were finished with their observations, the experimenter handed out the AttrakDiff questionnaires.

Scenario 2 was structured similar to scenario 1, as facial expressions are presented along with emotion-eliciting situations. In contrast to scenario 1, the faces of virtual actors were animated by performing eye-blinking animations, slight head rotations, and minor changes of emotion intensity to grant a vivid impression. The emotion-eliciting situations were presented in three information chunks (see study 3) along with an animated background showing one of the eight virtual settings. The animation of background contained short clips with camera tilts. Each sequence played in looped cycles and lasted 20 seconds. The animation is stopped when participants finished their observation task during the experiment.

Scenario 3 involved the assessment of user experience in an interactive setting. At the beginning of the experiment, subjects had the possibility to get used to the input controls (Wii-controller) by carrying out a tutorial. The tutorial contained an example scene made up of one ECA and three hotspots conveying dummy information. By performing the basic controls with the Wii-controller, the experimenter showed interaction possibilities. Afterward, participants were asked to maneuver within the scenery by themselves. They were told of the structure of interactive scenes and the purpose of information hotspots. Next, the stimulus was initialized by loading the required scenery. The scenery contained one ECA showing one out of four basic emotions, and three information hotspots incorporating descriptions of an emotion-eliciting situation divided into three information chunks (study 3). Participants observed the facial expressions and entered hotspot areas to read the situation-related information. As in scenarios 1 and 2, the experimenter handed out the AttrakDiff to subjects after the observation task.

Twenty participants took part in the study, 9 female and 11 male aged between 21 and 54 years (M = 31.8).

10.4 Results

We present the findings according to the three hypotheses:

(H1) The overall user experience in the interactive situation is higher than the user experience in animation scenario and the still scenario.

The evaluation of how users experience the various emotional stimuli in the varying conditions showed that the overall user experience in the interactive situation is rated at a higher level by users than in the animation scenario and still scenario (see Fig. 10.2). Comparing the overall judgment on the hedonic quality (HQ) of the users with a factor analysis shows that users rate the interactive scenarios presenting emotions as higher (M = 0.81, SD = 0.95) than for the still (M = 0.24, SD = 0.78) and the animated scenarios (M = 0.29, SD = 0.89). The post hoc analysis shows that the HQ is significantly different for the interactive scenario compared to the still scenario (LSD = 0.51, p = 0.01) and the animated scenario (LSD = 0.57, p = 0.04). We can conclude that emotions that are presented in an interactive setting lead to a higher user experience.

(H2) The animation scenario in terms of user experience is rated higher than the still scenario.

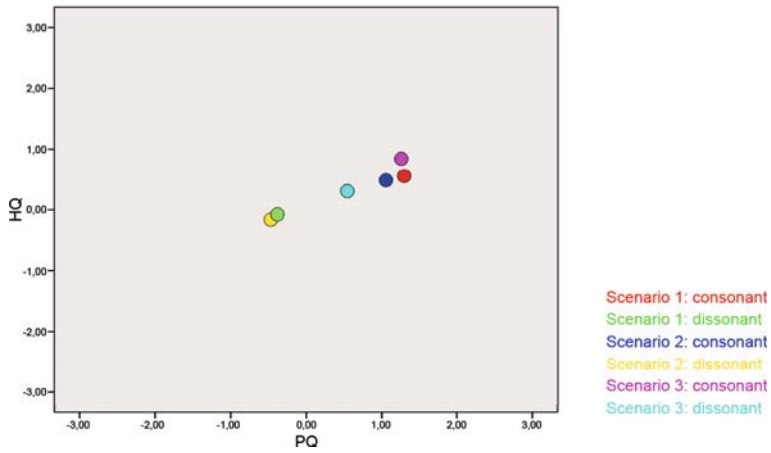

Fig. 10.2 Results of the AttrakDiff: overview of values concerning the hedonic and pragmatic quality of the three employed scenarios (consonant/dissonant)

Comparing still and animated scenario, no significant difference could be found. The animated scenario is thus not contributing to a positive user experience in the setting. The employment of animation in stimuli did not have the anticipated positive effect on UX.

(H3) The perceived user experience will be higher for consonant settings than for settings with dissonant stimuli.

Overall, the perceived user experience is rated different for consonant and dissonant settings. The ANOVA performed showed significant differences for consonant and dissonant settings ($F = 13.6$, $p < 0.000$).

Looking into the differences in more detail, a second ANOVA showed that all scenarios were rated significantly different in terms of hedonic quality (HQ) ($F = 5.25$, $p < 0.000$), as well as for attractiveness ($F = 9.16$, $p < 0.000$) and pragmatic quality ($F = 14.23$, $p < 0.000$). Figure 10.2 shows these results in more detail: showing higher ratings for consonant scenarios (upper right) and lower ratings for dissonant scenarios (lower left). The interactive scenario is rated best in both conditions.

The type of the presentation medium is heavily influencing the overall user experience in which an emotional stimulus is shown (in this case, a facial expressions of an ECA). For research on emotional aspects in psychology, it can be concluded that the context description can have a significant influence on how an emotion is perceived, as well as the scenario the stimuli are presented in.

For the games industry, the user experience in terms of facial expression of ECAs can be enhanced by providing consonant stimuli (consonant facial expression and description) and allowing direct interaction with the ECA (not only still or animated sequences).

10.5 Conclusions and Future Work

Investigating user experience and possible influences on user experience is a difficult task. We were interested in the relationship between emotions and perception of emotions and how the user experience for such a setting would be. To investigate this aspect, we (had to) carefully construct(ed) the stimuli material (emotion descriptions, ECAs showing these emotions) and to control any unwanted influences balanced various influencing factors (female/male ECA, etc.). The experiment showed that people do perceive ECAs that display emotion differently depending on the context description (emotion-eliciting situation) given. If the presented material is consonant, the overall user experience is higher, if the presented material is dissonant, the user experience is lower (even negative). User experience can thus be influenced (and can be designed positively in a game) giving congruent information in form of (written) scenario and presented emotions of the ECA. Second, user experience is not higher for semi-animated facial expressions. In general, user

experience is higher for interactive settings. The implications of this first results show that investment in small animations (for improving still images presenting emotions) is not improving the user experience, except the user is allowed to interact with the ECA. Creators in the games domain are interested in establishing an entertaining and intense gaming experience for their audience. A lot of production time is spent on creating realistic characters in detailed environments. Most situational aspects are defined without any theoretical foundation by focusing on artistic procedures. The work should show the necessity to relate to character-based communication channels (such as facial expression) with the current event the player confronts.

From the methodological perspective on how to evaluate user experience in games, we can summarize that experiments are one way to better understand the more general aspects of UX in games. As user experience is consisting of a wide variety of factors, it is difficult to find an experimental setup limiting the possible influencing experimental components. A careful experimental setup (including many prestudies) is thus a long process, and results for influencing factors on user experience are not immediately available. In general, experiments are a necessary means to understand the scientific basics of user experience, for an industrial context this kind of methodology might not be applicable. On a long-term basis, we see this kind of experiments as a necessary means to lay the foundations for understanding user experience.

References

Bailenson JN (2008) Avatars. http://www.stanford.edu/~bailenso/papers/avatars.pdf. Accessed 22 May 2008.

Bartneck C (2000) Affective expressions of machines. Master's Thesis, Stan Ackerman Institute – III, Eindhoven, Netherlands.

Bates J (1994) The role of emotion in believable agents. Communications of the ACM 37(7): 122–125.

Bernhaupt R, Boldt A, Mirlacher T et al. (2007) Using emotion in games: Emotional flowers. In: Proceedings of ACE 2007, ACM, New York, pp. 41–48.

Bernhaupt R, Ijsselsteijn W, Mueller F, Tscheligi M, Wixon D (2008) Evaluating user experiences in games. In: Proceedings of CHI 2008, ACM, New York, pp.3905–3908.

Bernhaupt R, Palanque P, Winkler M, Navarre D (2007) Model-based evaluation: A new way to support usability evaluation of multimodal interactive applications. In: Law E et al. (eds) Maturing Usability: Quality in Software, Interaction and Value. Springer, London, pp. 95–127.

Bernhaupt R, Sloo D, Migos C, Darnell M (2008) Towards new forms of iTV user experience. Workshop During EuroiTV 2008, 2nd July 2008, Adjunct Proceedings of EuroiTV 2008.

Bickmore T, Cassell J (2001) Relational agents: A model and implementation of building user trust. In: Proceedings of CHI 2001, ACM Press, New York.

Branco P (2003) Emotional interaction. In: Proceedings of CHI 2003, ACM, New York, pp. 676–677.

Cassell J (2008) Justin cassell: Research. http://www.soc.northwestern.edu/justine/jc_research.htm. Accessed 8 May 2008.

Cornelius RR (1996) The Science of Emotion: Research and Tradition in the Psychology of Emotions. Prentice Hall, Upper Saddle River, NJ.

Crane EA, Shami NS, Peter C (2007) Let's get emotional: Emotion research in human computer interaction. In: Proceedings of CHI 2007, ACM, New York, pp. 2101–2104.

Desmet P, Hekkert P (2002) Pleasure with Products, Beyond Usability, Chapter: The Basis of Product Emotions. Taylor Francis, London, pp. 60–68.

Dix A, Finlay J, Abowd G, Beale R (2004) Human–Computer Interaction. Prentice Hall, Essex, England.

Ekman P, Friesen W (1972) Emotion in the Human Face: Guidelines for Research and an Integration of Findings. Pergamon Press, New York.

Elliott C, Brzezinski J (1998) Autonomous agents as synthetic characters. AI Magazine 19(2): 13–30.

Ermi L, Mäyrä F (2005) Challenges for pervasive mobile game design: Examining players' emotional responses. In: Proceedings of ACE 2005, ACM, New York, pp. 371–372.

Fernandez-Dols JM, Carroll JM (1997) Is the meaning perceived in facial expression independent from its context. In: Russell JA, Fernandez-Dols JM (eds) The Psychology of Facial Expression (Studies in Emotion and Social Interaction). Cambridge University Press, New York, pp. 275–295.

Fischer L, Brauns D, Belschak F (2002) Zur Messung von Emotionen in der angewandten Forschung: Analysen mit den SAMs: Self-Assessment-Manikin. Pabst Science Publishers, Göttingen.

Hassenzahl M, Burmester M, Koller F (2003) AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: Ziegler J, Szwillus G (eds) Mensch & Computer 2003. Interaktion in Bewegung. BG Teubner, Stuttgart, pp. 187–196.

Hassenzahl M, Tractinsky N (2006) User experience – A research agenda. Behaviour & Information Technology 25(2): 91–97.

ISEAR (2008) http://www.unige.ch/fapse/emotion/databanks/isear.html . Accessed 4 April 2009.

Isbister K (2006) Better Game Characters by Design: A Psychological Approach (The Morgan Kaufmann Series in Interactive 3D Technology). Morgan Kaufmann Publisher, San Francisco, CA.

Jaimes A, Sebe N (2007) Multimodal human-computer interaction: A survey. Computer Vision and Image Understanding 108(1–2): 116–134.

Law EL, Vermeeren AP, Hassenzahl M, Blythe M (2007) Towards a UX manifesto. In *Proceedings of the 21st British HCI Group Annual Conference on HCI 2008: People and Computers Xxi: Hci. But Not As We Know It - Volume 2* (University of Lancaster, United Kingdom, September 03–07, 2007). British Computer Society Conference on Human-Computer Interaction. British Computer Society, Swinton, UK, pp. 205–206.

Lee J, Marsella S (2006) Nonverbal behavior generator for embodied conversational agents. In: Lecture Notes in Computer Science: Intelligent Virtual Agents. Springer, Berlin/Heidelberg, pp. 243–255.

Lieberman H (1997) Autonomous interface agents. In: Proceedings of CHI 1997 ACM, New York, pp. 67–74.

Lime (2008) Lime surveyor. http:// www.limesurvey.org . Accessed 27 September 2008.

Mahlke S (2005) Studying affect and emotions as important parts of the user experience. http://www.emotion-inhci.net/workshopHCI2005/Mahlke_StudyingAffectAndEmotionsAs ImportantPartsOfTheUserExperience.pdf. Accessed 28 February 2008.

Mancini M, Hartmann B, Pelachaud C (2004) Non-verbal behaviors expressivity and their representation. http://pfstar.itc.it/public/doc/deliverables/pelachaud_tech_rep3.pdf . Accessed 10 February 2007.

Mandryk RL, Atkins MS, Inkpen KM (2006) A continuous and objective evaluation of emotional experience with interactive play environments. In: Proceedings of CHI 2006. ACM, New York, pp. 1027–1036.

Minge M (2005) Methoden zur Erhebung emotionaler Aspekte bei der Interaktion mit technischen Systemen. Master's Thesis, FREIE UNIVERSITÄT BERLIN Fachbereich Erziehungswissenschaften und Psychologie.

Norman D (2002) Emotion and design: Attractive things work better. Interactions 9(4): 36–42.
Picard RW (1997) Affective Computing. MIT Press, Cambridge, MA.
Picard RW, Wexelblatt A, Nass CI (2002) Future interfaces: Social and emotional. In: Proceedings of CHI 2002 ACM, New York, pp. 698–699.
Prendinger H, Ishizuka M (2004) Life-Like Characters: Tools, Affective Functions, and Applications. Springer, Heidelberg.
Pruett C (2008) The evolution of videogames. http://bcis.pacificu.edu/journal / 2003/07/pruett.php Accessed 5 April 2008.
Ravaja N, Saari T, Turpeinen M, Laarni J, Salminen M, Kivikangas M (2006) Spatial presence and emotions during video game playing: Does it matter with whom you play? Presence: Teleoperators and Virtual Environments 15(4): 381–392.
Reeves B, Nass C (2003) The Media Equation. The University of Chicago Press, Chicago, IL.
Roto V, Rautava M (2008) User experience elements and brand promise. http://research.nokia.com/files/UXelements-v2.pdf. Accessed 5 April 2009.
Russell JA, Fernandez-Dols JM (1997) What does facial expressions mean. In: JA Russell and JM Fernandez-Dols (eds) Psychology of Facial Expression (Studies in Emotion and Social Interaction), Chapter Introduction. Cambridge University Press, New York, NY, pp. 3–31.
Scherer KR, Wranik T, Sangsue J et al. (2004) Emotions in everyday life: Probability of occurrence, risk factors, appraisal and reaction patterns. Social Science Information 43(4): 499–570.
Suk HJ (2006) Color and emotion: A study on the affective judgment of color across media and in relation to visual stimuli. PhD Thesis, Sozialwissenschaften der Universität, Mannheim.
Summerfield A, Green EJ (1986) Experiencing Emotion: A Cross-Cultural Study. Cambridge University Press, Cambridge, MA.
Wallbott HG (1990) Mimik im Kontext. Verlag für Psychologie, Dr. C. J. Hogrefe, Göttingen.