Chapter 26
An Introduction to Physiological Player Metrics for Evaluating Games

Lennart E. Nacke

Take Away Points:

1. Provides a brief introduction to physiological game evaluation.
2. Discusses the benefits and limitations of physiological measures for game evaluation.

26.1 Introduction

Do you remember insult swordfighting in Monkey Island? The moment when you got off the elevator in the fourth mission of Call of Duty: Modern Warfare 2? Your romantic love affair with Leliana or Alistair in Dragon Age? Dancing as Madison for Paco in his nightclub in Heavy Rain? Climbing and fighting Cronos in God of War 3? Some of the most memorable moments from successful video games, have a strong emotional impact on us. It is only natural that game designers and user researchers are seeking methods to better understand the positive and negative emotions that we feel when we are playing games.

While game metrics provide excellent methods and techniques to infer behavior from the interaction of the player in the virtual game world, they cannot infer or *see* emotional signals of a player. Emotional signals are observable changes in the state of the human player, such as facial expressions, body posture, or physiological changes in the player's body. The human eye can observe facial expression, gestures or human sounds that could tell us how a player is feeling, but covert physiological changes are only revealed to us when using sensor equipment, such as

L.E. Nacke (✉)
HCI and Game Science Group, Faculty of Business and Information Technology,
University of Ontario Institute of Technology, 2000 Simcoe Street North,
Oshawa L1H 7K4, ON, Canada
e-mail: lennart.nacke@acm.org

M. Seif El-Nasr et al. (eds.), *Game Analytics: Maximizing the Value of Player Data*, 585
DOI 10.1007/978-1-4471-4769-5_26, © Springer-Verlag London 2013

Fig. 26.1 An overview of game user research methods grouped together by similarity in a quantitative or qualitative and objective or subjective focus based on Mandryk (2008)

electroencephalographs (EEG), electromyographs (EMG), or galvanic skin response (GSR) recording systems. These player-focused body-related responses or physiological metrics are at the heart of this chapter.

26.1.1 Limitations of this Chapter

This book chapter was written with two audience types in mind: user researchers and graduate students. My goal is to give a brief overview of the field of physiological emotion research in games (more interesting as pointers for graduate students) as well as some how-tos for physiological recording (probably more useful for game user researchers). Keep in mind that this chapter cannot cover everything that you need to know about the background of recording physiological signals (Stern et al. (2001) is a better resource for this purpose) or give you all the information you will need to run physiological tests as a games user researcher. It is by nature a primer, something that hopefully gets you interested enough in physiological game research to start asking the right questions and look for the latest results in this growing field.

Figure 26.1 gives you an idea of the methods available for games user research and helps you locate at which part of the spectrum game metrics and physiological measures are (two of the more quantitative approaches available to game evaluators).

All of the methods mentioned in the figure are possible options for evaluating the player experience in a game depending on where your focus lies. Game metrics together with self-reported data from questionnaires or interviews can provide an additional cross-reference to physiological measures, which are still largely lacking validations for their use in games. But before we get there, let us review the emotional and physiological fundamentals for this type of games user research.

26.2 What Are Emotions?

Rosalind Picard mentions two types of signals in her "Affective Computing" book that need to be differentiated for emotion-recording systems: (1) expressive signals directly originating from a person and (2) non-expressive signals from environment and context of a person (Picard 1997). A physiological recording, for example, will not necessarily be able to differentiate between these types of signals. This is a problem that is similar to "situational stereotypy" (Lacey 1959), which is the idea that physiological responses depend on the experimental context. So, from a psychological view, the context in which players experience their emotions is as important as the game-related cues that trigger their emotions (for a player experience model that incorporates the idea of context see Engl and Nacke 2013). Picard (1997) notes that our expectations will influence our emotional perception, meaning that our body responses are shaped by our mental ideas and vice versa. A player's own mood and emotions will influence their perceptions and cognitive processes (for an excellent review of how affective computing relates to psychological emotion literature, I recommend reading Calvo and D'Mello (2010)). At this point, the boundaries between user experience research and emotion research start to blur (Brave and Nass 2002). But let us keep the focus on emotions in the psychological sense for this introduction. So, where do we start, when we want to understand and distinguish emotions?

Emotion research is a huge field with journals such as *Emotion, Emotion Review, Cognition and Emotion*, or *IEEE Transactions on Affective Computing* at its heart, and the scope of this article forbids going into real depth here, but I want to give you some pointers about what the different views of emotions are. A general starting point for the interested emotion researcher are the following introductory articles: Kleinginna and Kleinginna (1981), Panksepp (2004), Bradley and Lang (2007), Russell (2003), Barrett (2006), Barrett et al. (2007), and Dalgleish et al. (2009). Of course, for those wanting to go in depth in this field, there are also several comprehensive handbooks available, on emotions (Lewis et al. 2010), on cognition and emotion (Dalgleish et al. 2000), on psychophysiological research (Cacioppo et al. 2007), on affective sciences (Davidson et al. 2003), on emotion and the affective sciences (Sander and Scherer 2009), and on emotion and mass media (Döveling et al. 2010), just to name a few. Finally, you should consult some comprehensive books on affective computing (Picard 1997; Scherer et al. 2010; Gokcay and Yildirim

2010; Pelachaud 2012) as well as affect and emotion (Panksepp 2004; Lane and Nadel 2002; Frijda 1986; Ekman and Davidson 1994) if you really want to get more familiar with the topic. By its nature, my overview is only brief and scratches the surface of more than a decade of emotion research.

There is no definitive taxonomy for emotions and there are many different ways of classifying emotions. One of the oldest theories of emotion is the James-Lange theory, which states that our emotion follows from experiencing physiological change first (James 1884; Lange 1912). According to this theory, when an outside event or object changes, it causes the physiological or visceral change, which then generates the emotional feeling. This theory has been challenged several times and continues to be criticized.

One of the first challengers was the Cannon-Bard theory, which offers an alternative sequence of emotion processing. After a feeling occurs, Cannon hypothesized that it triggers a behavior based on how the emotion is processed (Cannon 1927). The emotional perception influences the physiological reaction. How you think you feel will change your reaction to the feeling. This theory tries to account for a combination of high-level mental and low-level physiological responses when experiencing emotions.

Another emotional concept is the two-factor theory of emotions which is based on empirical observations (Schachter and Singer 1962). This theory considers mental processing to have a large influence on our individual interpretation of our body reactions to an event that caused them. According to Schachter, emotions stem from the interaction of two distinct factors: cognitive labeling and physiological arousal (Schachter 1964). Cognitive processes provide the framework in which individual feelings are processed and labeled, giving the state of physiological arousal positive or negative values according to the situation and past experiences. LeDoux (1998) provides an excellent overview of this in Chapter 3 of his book; specifically, he discusses some of these theories and the pathways of interpretation from the causing event to the resulting feeling.

For those more interested in modern theories of emotion that take into account that emotional processes can happen without the resulting emotional experience, I recommend Damasio (1994). It is also worth considering the multicomponent process emotion model (Scherer 1984), which constitutes that an emotion processing system consists of the five distinct subsystems: information processing, support, executive, action, and monitoring. This model is rooted in appraisal theory (Lazarus 1968), which denotes that emotional arousal from a stimulating event is ingrained in the meaning it has for the person perceiving it. Most studies regarding appraisal theory models of emotion have used verbal reports. This requires participants to engage in complex recall and imagination processes before they put their feelings into words.

To summarize, at the heart of most emotion theories are two basic concepts: (1) discrete emotional states and (2) dimensional (often biphasic) theories. Discrete emotional states date back to early ideas of the French philosopher René Descartes, who described basic emotions, such as joy, wonder, love, desire, hate and sadness. Later, Ekman (1972) would describe the appearance of the face for six distinct emotions: surprise, fear, anger, disgust, sadness and happiness. A list that he extended later (Ekman 1992a, b). Other notable lists of discrete emotional states where contributed by Izard (1972) and Plutchik (1980). Plutchik also described a structural model of emotions (Plutchik 1991) has eight prototypic dimensions in horizontal (maximum intensity emotions at the top: ecstasy, acceptance, amazement, terror, grief, loathing, vigilance, rage) and vertical levels (lower located emotions are closer together and less intense).

While discrete emotions have a place in psychophysiology, they are often seen as broader concepts of underlying factors of a more affective nature, such as stimulus[1] appraisal (Scherer 1984), tendencies for action (Frijda 1986), or emotional expression through facial muscles (Ekman 1972). In psychophysiological research, dimensional models of emotion are more commonly used to conceptualize emotional facets. The most common model is the two-dimensional valence-arousal circumplex model (Russell 1980). The main criticism with this model is that the dimensions are not completely bipolar (Larsen et al. 2001), so alternative models were suggested, such as the positive activation and negative activation structure, that account for approach and withdrawal behavior (Watson et al. 1999). In this vein, another theory of positivity (appetition) and negativity (aversion) is offered by Cacioppo et al. (1999). In contrast to these newer models, early discussions of emotions tended to be completely biphasic, distinguishing between good and bad, positive and negative, appetitive and aversive, or pleasant and unpleasant. Only recently are we beginning to understand the complexity of affective processes that are often a blend of positive and negative feelings.

26.2.1 How Game Metrics Relate to Psychophysiological Emotion Induction

Emotions in a psychophysiological context can be understood as connected physiological and psychological affective processes, which can be induced by perception, imagination, anticipation, or action triggers (see Fig. 26.2). Perceptual emotions can be triggered by sensory information, such visual, acoustic, tactile, olfactory, or gustatory signals (Bradley and Lang 2007).

This distinction between emotional triggers is especially relevant when analyzing psychophysiological reactions together with game metrics. Only through the use of game logs that pinpoint exactly what game events were happening when, are we are able to contextualize physiological reactions of players (see Nacke et al. (2008) and Kivikangas et al. (2011b) for descriptions of a logging systems that used player interaction logs together with physiological responses). Finding out what cues in the game induce a physiological reaction can be done by logging game metrics and sending game events coded as voltage triggers directly to physiological hardware

[1] A stimulus in psychological research is something (could be an event or an object) that evokes a body or mind response.

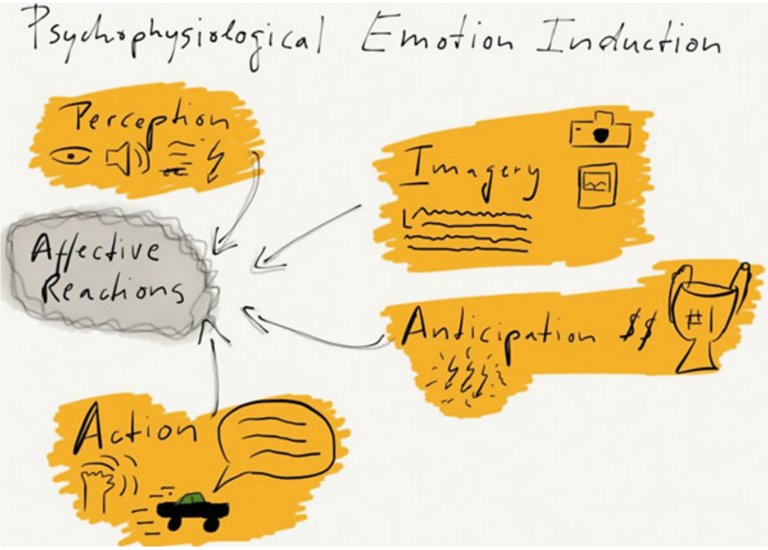

Fig. 26.2 Emotion inducers in the field of psychophysiology. Stimuli most often used in psychophysiological experiments come from these contexts

(Kivikangas et al. 2011b). Alternatively, one can triangulate player and event logs with physiological data as long as they contain a timestamp that is synchronized with the physiological timestamp (a procedure which can be difficult if several computers are involved; in this case networked time synchronization is suggested). Other approaches include a “manual” correlation of the physiological data with player events using video data. Here, several videos (usually of the player’s face, an in-game capture, an event log, and physiological graphs) are watched after a player session and events of interest are identified and scored manually in physiological data processing software. We will talk more about triangulation and data storage procedures for physiological data later in this manuscript. Before we talk about detailed physiological recording procedures, we need to understand the field of psychophysiology.

26.3 What Is Psychophysiology?

Psychophysiology is a research field where body signals, so called physiological responses, are measured to understand what mental processes are connected to those bodily responses (see Darrow 1964; Andreassi 2006; Hugdahl 1995; Cacioppo et al. 2007 for more definitions). I will refer to this as physiological metrics in this chapter. In this area of research, we are studying body signals to get an idea of what our mind was doing at that point. We are studying brain-behavior relationships that are guided by activity in the nervous systems (Hugdahl 1995). This makes psychophysiology a useful tool for evaluating excitement, emotion, or mental workload in games by conducting experiments. However, one has to keep in mind that valid experimentation requires much caution and preparation. In physiological experimentation, we have to balance ecological validity of our experimental environment with possible distractions that need to be controlled for.

Most of our body responses are spontaneous. This means they are difficult to fake, which makes physiological measures more objective than, for example, behavioral gameplay metrics, where a participant is able to fake doing an activity while cognitively engaging in another. One could say they allow the least biased assessment of how a player is reacting to gameplay actions compared to other game user research methods. They are also recorded continuously, meaning they do not interrupt a player's gameplay session. Physiological metrics are vast amounts of data, which become meaningful only when analyzed using the correct context and correct signal processing procedures (Mandryk 2008; Nacke 2009). For example, as a game designer we want to create meaningful decisions that involve some tradeoff of game resources (e.g., resource trades, weighing risk against reward, and choosing an appropriate action) (Brathwaite and Schreiber 2008). Here, emotional decisions can make playing games more fun. In these game decision situations, physiological metrics give you an objective way to assess a player's emotional response. We can get an idea about the emotional state of a player based on physiological metrics and this helps us inform game designers. In case of our example, we would know whether the designers have succeeded in causing an emotional response in the player with the decision options that they provided. However, again you have to keep in mind that this type of quantitative data has to be interpreted to make correct design suggestions, which leaves room for interpretation bias of the game user researcher.

Without a high level of experimental control, physiological data is volatile, variable, and difficult to interpret. For example, if a think-aloud protocol is applied when recording physiological metrics, a researcher risks influencing heart rate and respiration. When interpreting physiological metrics, it is important to understand the relationship between what happens in your brain (the psychological effect or mental process) and what your body tells us (the physiological variables, such as EEG, EMG, EDA). Cacioppo et al. (2007) note that there are five general relations between mental processes and body responses that we need to understand. The following relationships are distinguished:

- **The one-to-one relationship**. One mental process is directly associated with one body response and vice versa. This type of relationship would allow you to identify a mental process based on a body response and it is rarely possible.
- **The one-to-many relationship**. One mental process is associated with many body responses. Here, we cannot make draw a conclusion regarding mental processes.
- **The many-to-one relationship**. Many mental processes are associated with the same body response. While this scenario is worse than a one-to-one relation, it is the one most often used in physiological evaluation. It allows us to make assumptions of mental processes based on a body response.

- **The many-to-many relationship**. Many mental processes are associated with many body responses. Again, this type of relation does not allow for a conclusion of mental processes based on body responses.
- **The null relationship**. There is no relationship or association between mental processes and body responses.

The most common case in physiological evaluation is the many-to-one relationship, where one body response may be associated with many mental effects or processes. Therefore, we must keep in mind that a direct mapping of a discrete emotional state is not possible (and it is debatable whether discrete emotional states even exist, although we will not touch on this discussion) and body responses must be understood as elements of sets with fuzzy boundaries. When we measure body signals, we are measuring essentially the operation and activity of muscles, nerve cells, and glands.

26.4 Physiological Response Metrics of Players

To understand how physiological measures work on the human body, we need to take a quick neurobiological look at how our bodily reactions are organized. On a macro level, bodily operations are controlled by our nervous system, which is split into two parts, the central nervous system (CNS) and the peripheral nervous system (PNS). The CNS consists of big brain (cerebrum), little brain (cerebellum), and spinal cord. It manages all the information received from the whole body and coordinates body activity accordingly. The CNS is well protected by the skull and spine bones, which also makes it difficult to access outside of the body. The PNS includes all nerve cells outside of the CNS. You could say that its main job is to connect the CNS to the rest of our body. To use an example from musical theater, you can imagine the CNS having the same functions as the conductor in a concert, whereas the PNS would be the orchestra.

Since most of our physical sensations are transmitted through the PNS, we are able to measure its reactions on our skin. The skin is the place where most physiological sensors are applied. More on the micro level, the PNS is split into the somatic and the autonomic nervous system. It is enough to say here that the somatic nervous system regulates body activity that we have under conscious control such as deliberate movement directly through our muscles. The autonomic nervous system (ANS) is more exciting for physiological evaluation because it takes main care of our unconscious, visceral responses. These responses are hard to get with classic game user research methods and physiological metrics can really shine here. In the ANS, just like in a good game, we have two opposing players, the sympathetic nervous system and the parasympathetic nervous system. The former is our emergency response system that triggers fight or flight reactions while the latter manages our relaxation, resting, and digesting. It is important to keep those two players in mind, when we look at how we measure emotion with physiological sensors.

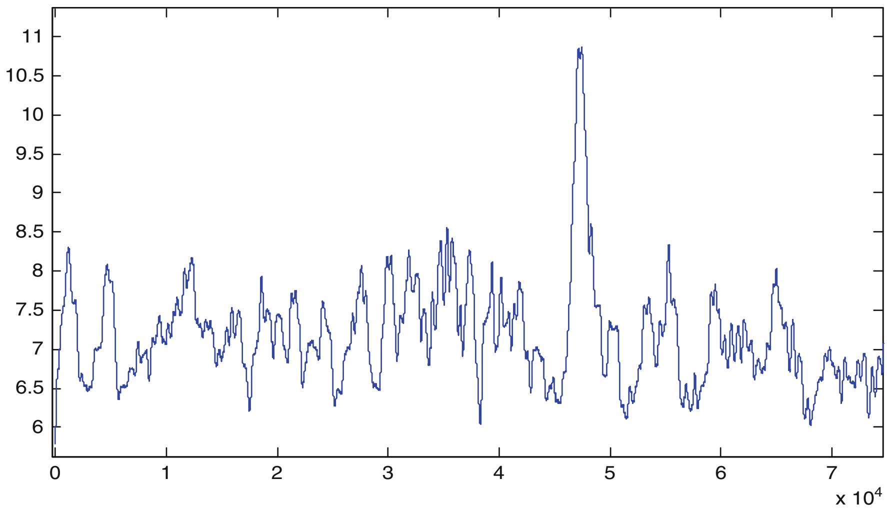

Fig. 26.3 Raw psychophysiological signal with a baseline offset (e.g., EMG)

The PNS is particularly useful for measuring stimulation, but not so much when it comes to measuring emotion itself. However, by detecting slightest muscular movements in the face with physiological sensors, we are able to assess emotion based on facial expression. For example, a frowning face would express negative emotion, whereas a smiling face would express positive emotion. Both reactions will shows as spikes in the data of the physiological sensor that is applied to the corresponding region of the face. As game user researchers, we are also interested in feelings and the experience players are having when interacting with a game. Therefore, we cannot solely rely on physiological metrics for player testing, but we have to accompany them with questionnaires or other contextual recording techniques (e.g., interviews, video observation, game metrics) to get a better idea of player experience. However, the basic tenet of physiological experimentation still holds true: we measure the physiological response (in addition to other subject responses) while manipulating a behavioural factor, often an element of gameplay. For an overview of recent game research studies that are investigating psychophysiology in games, see Kivikangas et al. (2011a).

26.4.1 Physiological Signal Processing Primer

Raw physiological signals, such as EEG and EMG (EDA is a bit different as we will note later), represent an assembly of positive and negative (i.e., an oscillating) electrical voltage. Important traits of a physiological signal are the frequency (number of oscillations) and their amplitude (maximum positive or negative voltage). An example of a raw psychophysiological signal can be seen in Fig. 26.3.

In a psychophysiological recording graph, we can see two dimensions. The abscissa shows the recording time (often ms) and the ordinate displays the amplitude (often in μV). Raw EMG signals, especially when the sensors are applied to larger muscle sites are characterised by activity bursts (when the muscle contracts) and baselines during the muscle resting periods. For facial EMG, the signal is not quite as distinguishable in experimental situations, since an alert participant is rarely completely relaxed and there is always some activity visible in the raw EMG (depending on the resolution and filters of the recording hardware). Recording baseline noise before exposure to an experimental stimulus and subtracting this from the signal is a common method used for removing unwanted noise from the recordings before starting to filter the signal. When first looking at physiological signals, do not be surprised by seeing many negative numbers in your first recording (negative values are plotted upwards as a neurophysiological convention). Signal filtering from the raw signal toward a normalized signal usually follows the following steps (Tassinary et al. 2000):

1. Rectify the raw signal using a technique called RMS or root mean square, which equals the quadratic mean of a number. Applying an RMS transformation to the raw signal folds makes it easier to view and understand.
2. Sometimes the physiological recording hardware (the so-called amplifier) already has a bandpass filter (10–500 Hz) built in. If not, then at least for EMG signals a lowpass filter of 500 Hz should be applied to remove noise from the amplifier hardware. The decision of whether to use a 10, 20 or 30 Hz high pass filter depends on how much the researcher wants to attenuate weak signals (30 Hz gets rid of cross-talk and other noise such AC power). The 10–500 Hz bandpass filter is sufficient for most EMG applications and can easily be implemented in MATLAB using a digital third order Butterworth bandstop filter. For EEG data, using a smaller range between 1 and 40 Hz would be advisable depending on whether gamma frequencies (30–50 Hz) are used in the analysis. Since EEG analysis works on lower frequencies, a notch filter can be applied as well to remove 60 Hz noise.[2]
3. Finally the signal is often smoothed using a moving window technique, where based on a time window defined by the researcher data is averaged within the moving window. This often called moving average or average rectified value.
4. For EEG, a next step would be to calculate average power estimates with a Fast Fourier Transformation (FFT). Since most EEG analysis is more complicated and warrants a chapter of its own, we will not go into depth here.

Depending on what kind of statistical analysis is later done with the physiological signal, it can be logarithmically (log or ln) transformed to eliminate skew in the data distribution. The rest of the analysis is done depending on the experimental setup and using statistical methods.

[2] 50/60 Hz is the electrical energy frequency that can come from lights, power supplies and other devices in your experiment environment.

When designing a games user research study that could involve physiological sensors, you have to pick wisely which sensors to use and whether to use sensors at all. Some game user research questions can be answered with other methods, depending on what you want to know about the user experience. Skin conductance level is correlated with psychological arousal (Prokasy and Raskin 1973), but so are cardiovascular measures, such as blood volume pulse, higher heart rate, and respiration. If you wanted to look at mental effort and task load, you could resort to a subjective measure like the task load index (Hart and Staveland 1988), look at multivariate EEG measures (Smith et al. 2001), decreased heart rate variability, or more dilated pupils (using eye tracking technology), brow or jaw muscle activity (Waterink and van Boxtel 1994). We will also later in this chapter talk about using facial EMG to asses positive or negative valence of emotions to indicate whether a game action is perceived as positive or negative. With all these measures available and each of them using different signal processing, the task of choosing the right one might seem daunting at first. For people getting started with physiological measures, I recommend sticking to the basic skin conductance and EMG measures presented in this chapter.

For some games sensors might be a better fit than others. In our experience, action games that produce a visceral experience are generally a good fit for physiological measures (Nacke 2009). It remains to be shown whether this is a useful tool for casual games as well, since recent reports have yet to make a strong argument for the method in this context (Gualeni et al. 2012). If you decide that physiological sensors are your method of choice for your games user research question, you should choose a sensor that will not alter the player experience through its application, but one that is sensible to the effects that you want to measure.

26.4.2 Electroencephalography (EEG)

There are many myths surrounding EEG as a measure of brainwave activity of the human body. Participants unfamiliar with this technique may assume that you are able to find out exactly what they are thinking or even get graphic representations of their thoughts. While recent research in the latter (i.e., reconstructing visuals from brain activity using magnetic resonance imaging [MRI]) has been impressive (Nishimoto et al. 2011), the reality of EEG measures is a little bit more abstract than one might think. Compared to other techniques of analysing the CNS response, such as functional MRI or positron emission tomography (PET) scans, EEG, can be considered less invasive and easier to apply. The advantage of EEG for brain activity measurement over these other techniques is its millisecond resolution, which allows studying physiological responses in real time. A slight disadvantage of EEG to the other approaches is its spatial resolution, which is constrained, for example, by a low signal-to-noise ratio and limited spatial sampling.

Example Experimental Protocol for Attaching EEG Electrodes

1. When selecting or inviting participants for physiological studies that use surface electrodes such as EEG or EMG, it is a good idea to **screen participants for hair growth**. EEG usually has to be applied directly on the scalp (head skin surface) of the participant, the more hair a participant has, the more electrode gel you are likely to use. When using male participants for facial EMG, a beard might also be problematic when trying to apply the electrode directly to the skin. Adhesion is reduced when much or thick hair is present, especially when recording under humid conditions or with skin types prone for sweating. Also, no chewing gum for participants.
2. Most dry electrodes do not require extensive **cleaning of the skin**, although it is recommended for hygienic reasons to clean the skin before and after attaching electrodes, but not with soap. A soft cleaning with an alcohol pad or conductive cleaning paste is usually sufficient for removing dead skin cells.
3. If not using pre-gelled electrodes, the **electrode needs to be gelled** (or wetted for some toy EEG devices) for optimal skin contact. EEG electrodes are often snapped into a cap that is worn (and needs to be correctly placed) on a participant's head to ensure correct alignment of all electrodes.
4. Depending on thickness of the electrode cables, having **surgical tape** on site is invaluable for making sure that the **cables** are closely **attached** to the participant and do no move around during recording.
5. After the recording or experiment is done, the **electrodes need to be removed** from the participant as soon as possible to minimize discomfort.
6. All **equipment** that was in contact with the participant needs to be **washed** (sensor cap and straps) or thrown away (disposable electrodes).

In EEG, electrodes are placed on a participant's head. Their location and alignment is standardized in the 10–20 system (Jasper 1958) (or the 10–10 EEG sensor placement system (Chatrian et al. 1988)). Often EEG systems ship with caps that take care of this alignment by having electrode inlets sewn into headgear that looks like a swimming cap. EEG measures slight electrical activity, such as the signals generated by neural activity in the brain. There is a wide range of different measurement devices available for this type of physiological measure, ranging from a more sophisticated medical grade headcap setup with large density electrode arrays (from 32 to 256 electrodes) and simpler devices that have less electrodes and therefore less spatial accuracy but similar time accuracy. Some really cheap EEG devices sell for lower than $1,000 (e.g., Neurosky, Emotiv). Most of these devices compute affective and cognitive states such as attention, engagement, boredom, meditation, frustration, or long- and short-term excitement. Be aware that these computations are not openly available and they are mostly a black box for researchers.

EEG lets us record electrical activity on the head that relates to brain activity. We usually distinguish brain activity by using the amplitude and frequency of the signal in comparison to a reference location. Amplitude describes the size of the signal, while frequency refers to the speed of signal cycles. EEG devices compute brain waves in different frequency bands, such as alpha (e.g. 8–13 Hz), beta (e.g. 13–30 Hz), theta (e.g. 4–8 Hz), delta (1–4 Hz), and sometimes gamma (30–50 Hz).[3] Alpha activity is associated with relaxation and lack of active cognitive processes; it has also been tied to information and visual processing. Beta activity is related to alertness, attention, vigilance, and excitatory problem solving activities. Theta activity has been related to decreased alertness and lower information processing, however, frontal midline theta activity in the anterior cingulate cortex scalp area is linked to mental effort, attention, and stimulus processing. Delta is most prominent during sleep, relaxation or fatigue. Gamma activity is still largely unexplored. While these associations come from research in medicine and psychology, they make it easier to evaluate a game based on the EEG activity. For example, if you notice increased beta activity during gaming, it could be linked to player attention and increased arousal during a focused gaming task.

A major disadvantage of early EEG methods was the placement of the electrodes with gel. Many budget-type EEG devices got rid of the gel and have dry electrodes. This minimizes discomfort by providing a comfortable fit on the head.

EEG is difficult or impossible to measure when there is movement involved. The electrodes might move on the head while the player is moving. This leads to artifacts in the EEG data. Therefore, some games are not very suited for EEG evaluation (e.g., Guitar Hero, Kinect, or Wii games). Movement artifacts are a problem of all physiological measures, but are especially problematic with EEG as we are interpreting very low electromagnetic activity. It is important to apply proper filters to your EEG data, so that no interferences are recorded in the EEG signal (e.g., often a 50/60 Hz notch filter is used to exclude interference signals).

In addition, as with all physiological measures, EEG measures should be recorded with a baseline. For example, you could record this at the start of your session and let the player do nothing but stare at a cross on a grey background. This allows getting rid of the noise in your EEG signal. A final problem with EEG as a method is the difficult interpretation of the data. For example, when delta activity is increased in a playing session, do we argue that the game is relaxing or that it is boring and fatigue-inducing? It is quite important to keep one's game design goals in mind when doing this type of evaluation. Relating this data with other measures is paramount for a solid interpretation. Table 26.1 shows the pros and cons of EEG.

26.4.3 Electromyography (EMG)

An EMG measures whether our muscles are active or not. Therefore an EMG electrode attached to the surface above a muscle is able to sense even the slightest activation of this muscle (Bradley et al. 2001; Lang 1995). Whenever we flex a

[3] Another way of analysing EEG is through Event-Related Potentials or Mu Rhythm, which I do not cover in this chapter.

Table 26.1 Pros and cons of measuring EEG

PRO	CON
Great time resolution	Low space resolution
Deep cognitive insights	Gel-based caps and conductivity
Quantitative data	Movement artifacts
Small system setup	Data needs proper filtering
Different analyses possible with the same data set	Difficult to interpret
	Expensive

muscle on our body, this produces a difference in electrical activity or isometric tension which is measurable by EMG. While EEG measures activation in the CNS, EMG is all about measuring PNS activation. Since most muscles can be directly controlled, EMG is a measure of high interest for interacting with computers in a more natural way (Nacke et al. 2011).

However, the most common use for evaluating games is through facial EMG (Fridlund and Cacioppo 1986), which measures the activation of specific facial muscles responsible for displaying our positive or negative reactions to an emotional moment in a game (Hazlett 2006). In particular, physiological game research has focused on using brow (corrugator supercilii) to indicate negative emotion and cheek muscle (zygomaticus major) to indicate positive emotion (Mandryk et al. 2006) or even fun and flow in a game (Nacke and Lindley 2008). For longer term evaluation (say over a few minutes of gameplay), the eye muscle (orbicularis oculi) has also proven helpful in registering high arousal pleasant emotions (Ravaja et al. 2008).

In game user research using facial EMG to assess emotions, we have recently found the threshold of the total signal average with added standard deviation (Hazlett 2008) helpful to identify significant positive and negative gameplay moments (see Fig. 26.4 for a visualization). Hazlett used this to calculate an EMG ratio for a game with the total time spend in brow muscle activation (negative) or cheek muscle activation (positive). In my research group,[4] we have used this positive and negative gameplay time measure together with video observation and biometric storyboards (Mirza-Babaei et al. 2012) to identify key positive and negative moments during gameplay. Combined with gameplay logs, we can correlate this negative and positive gameplay response time with behavioral events, such as button presses, navigational interactions and gameplay actions. We are working on automating this scoring process and are working on validating this procedure and making the tools available for the game industry. While this will not allow the fine grained level of details that a mixed methods gameplay video analysis will provide, a gameplay metrics based scoring system that takes into account physiological responses will definitely be useful for the game industry.

Similar to EEG, EMG uses silver-silver chloride electrodes (see Fig. 26.5) because they have only a small measurement error, little drift potential, and minimal polarization. EMG electrodes are applied to the surface of the skin and will also

[4] http://hcigames.businessandit.uoit.ca

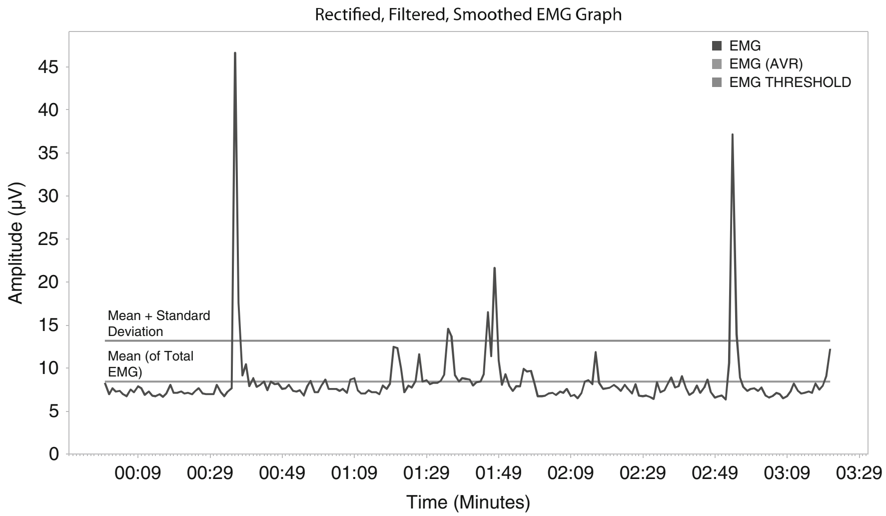

Fig. 26.4 Example of a filtered, smoothed and rectified EMG graph, showing thresholds for EMG analysis. Hazlett (2008) suggested using a threshold of average (i.e., Mean) EMG Amplitude plus Standard Deviation for finding positive measures

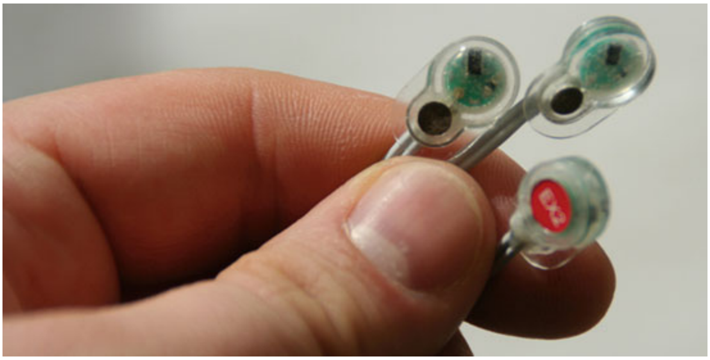

Fig. 26.5 An example of EMG electrodes (silver-silver chloride)

need a reference (if part of a larger system this reference can be on the head or close to the EMG electrodes). In measuring facial EMG one risk is to pick up muscle activity that is not related to the muscles that you would like to measure, such as cheek muscle for positive emotions or brow muscle for negative emotions.[5] In clinical settings, EMG electrodes might be placed under the skin surface to eliminate muscle interference, but these are not appropriate in a game user research setting. However, screening for facial hair is recommended, since body hair can cause interference

[5] For example, participants in an experiment cannot chew gum, laugh, or talk during facial EMG, because this will introduce large artifacts in your EMG data.

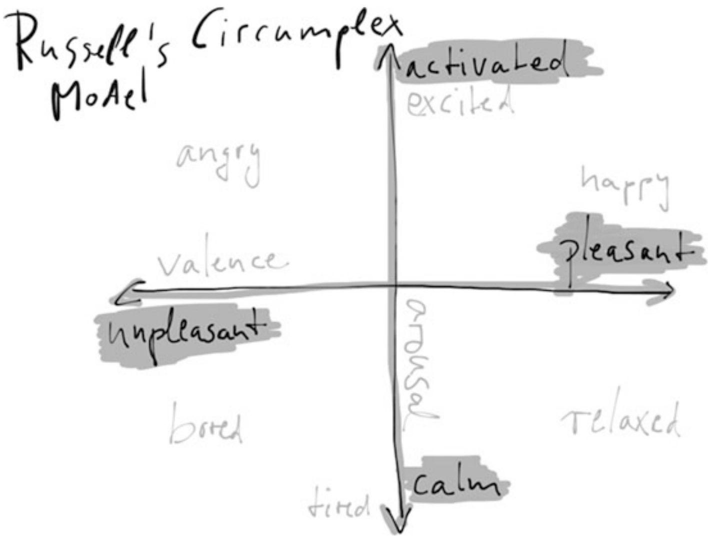

Fig. 26.6 Two emotion dimensions (valence and arousal) in the circumplex model from Russell (1980)

Table 26.2 Pros and cons of measuring EMG

PRO	CON
Great time resolution	Muscle and movement interference
Best way to measure emotion	Data needs proper filtering
Quantitative data	Electrode placement in the face
Easy signal analysis	Difficult to get a natural measurement
More precision than face cameras	Expensive

with EMG signals. Since muscular signals are amplified from microvolts, careful signal processing has to be done on EMG data before it is interpreted.[6]

Figure 26.6 shows how emotions are usually interpreted in psychophysiology on a two dimensional model (Russell 1980). We find that by measuring these face muscles we are able to get an idea of pleasant or unpleasant emotions along one axis of this model. This is called emotional valence assessment as we are able to show whether an emotion was evaluated by a player as pleasant or unpleasant.

While facial recognition software or direct observation would also allow the analysis of facial expressions and therefore the mapping on emotions, the software or the observer often miss less salient expressions, which are picked up by physiological measures. See Table 26.2 for pros and cons of EMG.

[6] The usual processing procedure is signal smoothing (often at half of the recording frequency, for example 0.5 s at 2 kHz recordings), baseline subtraction, and sometimes a logarithmic normalization. Depending on the system, an additional bandpass filter (high: 10Hz, low: 400Hz) or a Butterworth lowpass filter of 500Hz are necessary.

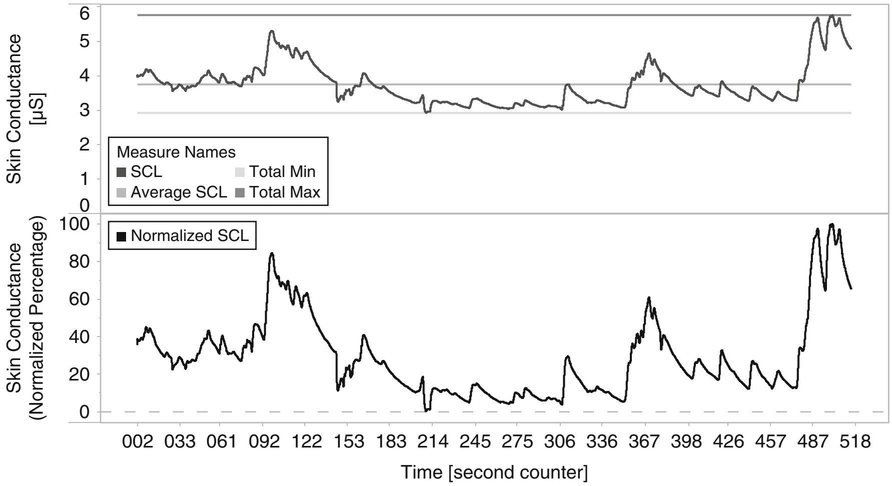

Fig. 26.7 A skin conductance level (SCL) graph for an example EDA recording. The *upper graph* shows the raw skin conductance level measured in µS sampled at 32 Hz over a total time of ~513 s (16,415 samples) together with the total average SCL as well as total minimum and total maximum. The *second graph* shows a normalized version of the skin conductance level based on the equation described below

The analysis of EMG signals is straightforward, as usually after application of some filters, we are already able to compare the signals. More activity on the cheek muscle relates to positive emotion, more activity on the brow muscle relates to negative emotion. However, EMG measures in the face of a player mean that there are electrodes attached to the player's face while playing, which make this measure intrusive although often the electrodes and cables can be easily taped to the player's head to remove discomfort and reduce movement artifacts. One thing to keep in mind is that just by feeling the electrode on their face, players might be feeling the need to elicit more pronounced muscle movements when playing. This might lead to unnatural signals, which could make data interpretation more difficult (if no video recording is available to check for this problem).

26.4.4 Electrodermal Activity (EDA)

EDA relates to how excited we are when we are exposed to a stimulus, such as playing a game. When measuring the skin conductance level (SCL) over time, we refer to this as measuring the EDA of the skin (see Fig. 26.7), but when measuring the direct response to an event, we call this galvanic skin response (Boucsein 1992). In any case, EDA measures changes in the passive electrical conductivity of the skin relating to increases or decreases in sweat gland activity. These fluctuations are caused by a person getting aroused by something that they see or do.

Most of us have seen EDA measures in movies branded as lie detector tests. EDA measures are attached to the fingers, palms or toes because the sweat glands in those body areas are more likely to react to changes in the PNS (sympathetic vs. parasympathetic activity). Since we are measuring the differences in conductivity, we only need two electrodes, which make EDA a very easy physiological measure to prepare and apply. EDA electrodes are prone to movement artifacts just like other psychophysiological measures, but because of their location (hand or feet) special care has to be taken in the preparation of steering controls of the game. If a regular game controller is used and the EDA sensor is applied on the palm of the hand, movement artifacts are likely to occur, so using the feet or fingers of the non-dominant hand would be a better location (or the side of the palm of the hand). EDA is also very easy to interpret, since it almost has a one-to-one relationship with physical arousal. However, individuals are different in their SCL, so a comparison between people is only possible with normalized data. SCL can be normalized by calculating each sample as a percentage of the entire span of EDA, using the min and max values over all samples for one participant (Mandryk 2008; Lykken and Venables 1971). The equation below shows how to normalize your SCL data at any point in time (SCL_{now}) as a percentage, given that you know the maximum (SCL_{max}) and minimum (SCL_{min}) value of your EDA data.

$$SCL_{normalized} = \frac{SCL_{now} - SCL_{min}}{SCL_{max} - SCL_{min}} \times 100$$

An equation of normalizing skin conductance level (SCL) based on Mandryk (2008).

Another benefit of this measure is the inexpensive hardware that usually comes at a fraction of the cost of a research-grade EEG setup. Many modern EDA systems use dry electrodes and some EDA setups for example allow quickly attaching the electrodes to the little and ring finger with a Velcro strap. This makes EDA are very handy measure for game user research.

Analyzing SCL can be done in a macro (EDA over larger chunks of playtime) or micro fashion (GSR related to events). When analyzing the response to a direct event, one needs to take into account that EDA is a relatively noisy signal that also has some delay in response to a stimulus (often around 5 s). After a galvanic skin response is registered, there is also a decay or recovery time during which no further event responses will be registered (or the responses are registered together). In addition to this, EDA tends to drift over time, possibly a result of the hands or feet getting sweatier. A good way to make sure this drift does not affect the EDA data too much, is to make sure to have resting periods between different gameplay sessions. While it is pretty clear that EDA indicates physical arousal, there is still some interpretation effort required as to what this stimulation comes from. Is it really from a game stimulus or are environmental factors contributing to the response? This is why planning and controlling physiological experiments is very important. Possibly confounding factors, such as high physical activity, loud noise, caffeinated substances, bright light, and things moving in the background should be avoided at all costs when running a physiological study. Table 26.3 shows the pros and cons of the EDA physiological metric.

Table 26.3 Pros and cons of measuring EDA

PRO	CON
Cheap hardware	Noisy signal
Easy to measure	Large individual variation
Easy to interpret	Baseline and response fluctuations
Less intrusive than other biosensors	Slow decay over time

Table 26.4 Pros and cons of cardiovascular measures

PRO	CON
Heart rate is easy to measure	Intrusive sensor
Heart rate hardware is cheap	Affected by many different things
Cardiovascular measures are established and prominent	Heart rate variability has a complex analysis procedure

26.4.5 Cardiovascular Measures

There are many cardiovascular measures available for physiological evaluation and all of them relate to the heart rhythm, its changes, and how this influences the physiological state of a human. The most common measures are electrocardiography (ECG), heart rate (HR), interbeat interval (IBI), heart rate variability (HRV), blood volume pulse (BVP), and blood pressure (BP). While physiological electrodes are necessary for all measures, blood pressure is not a real-time measure and also usually used in a medical context and has not been shown to be of relevance to game user research.

ECG measures the electrical activity caused by the heart pumping blood and is measured with three electrodes or leads, which are positive, negative and neutral and are usually applied to the upper body area. This can be considered a somewhat intrusive area for sensor placement depending on a participant's comfort level.

Heart rate is understood as the number of heart beats per time unit (usually measured in beats per minute). The amount of heart beats during a time unit is an interesting metric as is the time between the beats, the IBI. If IBI decreases, HR increases and this has been tied to increased information processing and emotional arousal. So, IBI and HR are two related measures. However, HR variability is a more complicated measure with a complex analysis procedure. In HRV, we are looking at differences in the IBI over time and analyze frequency changes. In general, we need to keep in mind that cardiovascular measures are intrusive to measure accurately and they are affected by many things, such as physical activity. Table 26.4 shows the pros and cons of physiological cardiovascular measures.

26.4.6 Other Physiological Measures

There are a number of other physiological measures not covered in this introductory chapter, such as respiratory sensors, eye trackers, temperature sensors, and brain imaging techniques. Another chapter in this book deals with eye tracking techniques in depth. And there is good other introductory literature available for this field (Duchowski 2007) as well. In addition, for more details on respiratory sensors and cardiovascular measures, other sources are available (Mandryk 2008).

26.5 Case Study: Physiological Measures of Sonic Gameplay Experience

This case study explains an experiment I conducted together with colleagues during my Ph.D. studies and a part of which was published in the journal *Interacting with Computers* (Nacke et al. 2010). Our initial research question behind this study was whether we can investigate the effects of sound and music in games on physiological measures and subjective measures of player experience. The research was conducted using a modified version of the first person shooter *Half-Life 2* (Valve Corporation, Bellevue, WA, USA). The modified level was designed for a playing time of 10 min. The game mod was played four times in different sound and music conditions. A first-person shooter is an excellent environment to conduct this type of research, since gameplay is highly arousing and visceral, which we hoped to be likely to yield physiological responses.

26.5.1 Metrics Used in This Study

Facial electromyography (EMG) was used to record the activity from left orbicularis oculi (eye), corrugator supercilii (brow), and zygomaticus major (cheek) muscle regions using BioSemi flat-type active electrodes with sintered Ag-AgCl (silver/silver chloride) electrode pellets having a contact area 4 mm in diameter. The electrodes were filled with low impedance highly conductive Signa electrode gel (Parker Laboratories, Inc., Hellendoorn, The Netherlands). The raw EMG signal was recorded with an ActiveTwo AD-box at a sample rate of 2 kHz and using ActiView acquisition software, and afterwards filtered in BESA (MEGIS GmbH, München, Germany) using a low cutoff filter (30 Hz; Type: forward, Slope: 6 dB/oct) and a high cutoff filter (400 Hz; Type: zero phase, Slope: 48 dB/oct). Electrodermal activity (EDA) was measured using two passive Ag-AgCl (silver/silver chloride) Nihon Kohden electrodes (1 mA, 512 Hz). The electrode pellets were filled with TD-246 skin conductance electrode paste (Med. Assoc. Inc., St. Albans, VT, USA) and attached to the thenar and hypothenar eminences of a participant's left hand. EMG data were rectified and exported together with EDA data at a sampling interval of 0.49 ms to *SPSS* (SPSS Inc., Chicago, IL, USA) for further analysis. Data were considered to be invalid when no signal was recorded for long periods (e.g., electrode fell off or equipment error). These data were excluded from further analysis: this was the case for seven participants.

Different components of game experience were measured using the gameplay experience questionnaire GEQ. It combines several game-related subjective measures (with a total of 36 questions): immersion, tension, competence, flow, negative affect, positive affect and challenge. Each dimension has five items (except immersion which has six items). Each item consists of a statement on a five-point scale ranging from 0 (not agreeing with the statement) to 4 (completely agreeing with the statement). Example statements are "I forgot everything around me" (Flow), "I was good at it" (Competence), "I felt that I could explore things" (Immersion), "I felt frustrated" (Tension), "I had to put a lot of effort into it" (Challenge), "I enjoyed it" (Positive Affect), and "I was distracted" (Negative Affect). The questionnaire was developed based on focus group research and subsequent survey studies (Cronbach's alpha values ranged from .71 to .89 in the original study).

26.5.2 Experimental Design

We employed a 2×2 repeated-measures factorial design using sound (on and off) and music (on and off) as independent variables, using a counter-balanced order of sound and music game-level stimuli. Thus the conditions were: (1) Sound on, Music off, (2) Sound off, Music off, (3) Sound on, Music on, (4) Sound off, Music on. EMG and EDA responses were measured together with questionnaire items indicating the overall game experience for the different playing conditions. Questionnaire item order was randomized for each participant.

Data were recorded from 36 undergraduate students (66.7 %) and University employees. Their age ranged between 18 and 41 (M=24, SD=4.9). Gender was not evenly distributed, since only 19.4 % of all participants were female. All participants played digital games regularly, and 94.4 % reported they play games at least once a week. 94.4 % believed they had full hearing capacity. 41.7 % saw themselves as hobby musicians, while only 33.3 % played an instrument, which can be explained by people working with sound recording and programming but not playing an instrument. All participants considered sound at least "somewhat important" in games.

Although Half-Life 2 allows the control of game audio features internally, sound and music were controlled externally for this experiment. For example, a music track was triggered externally, which was audible during playing and a software trigger controlled whether the game engine would play game sound or not.

26.5.3 Data Processing

We approached data processing in two different ways. First, we were investigating tonic or cumulative responses of the cumulative time period when playing the game in each of the condition, so after filtering and rectifying the signal, we compared the average values of each individual condition using inferential statistics. Second, we opted for an event-based analysis approach, where data was clustered 7 s around player death events. We averaged seven 1-s means, 1 s before (baseline; Second 1) and 6 s after the event (the death of the player; Seconds 2–7). To normalize the distributions of physiological data a natural logarithm was taken from EDA and EMG signals. All data were analyzed in SPSS (SPSS Inc., Chicago, IL, USA) by the linear mixed model procedure with restricted maximum likelihood estimation and a first-order autoregressive covariance structure for the residuals. Participant ID was specified as the subject variable, while the game audio conditions (sound on/music off; sound off/music off; sound on/music on; sound off/music on), the sequence number of the event, and second (1–7) were specified as the repeated variables. When examining the main effects of game events, the condition, sequence number of an event, and second were selected as factors, and a fixed-effects model that included the main effects of these variables was specified. When examining the interaction effects of condition and game events on physiological activity, the condition, sequence number of an event, and second were selected as factors, and a fixed-effects model that included the main effects of these variables and the condition × second interaction was specified.

Main effects of event-related changes in physiological activity were tested using the following contrasts:

- *Contrast 1*: baseline (Second 1) vs. response (Seconds 2–7).
- *Contrast 2*: linear trend across Seconds 1–7.
- *Contrast 3*: quadratic trend across Seconds 1–7.

Interactions were tested for both quadratic and linear trends. However, since the interaction contrasts with quadratic trends yielded no significant associations, only those using linear trends are reported as follows:

- *Interaction Contrast 1a*: sound vs. no sound × linear trend across Seconds 1–7. Interaction Contrast 1b: sound vs. no sound × change from baseline (Second 1 vs. Seconds 2–7).
- *Interaction Contrast 2a*: music vs. no music × linear trend across Seconds 1–7. Interaction Contrast 2b: sound vs. no sound × change from baseline (Second 1 vs. Seconds 2–7).
- *Interaction Contrast 3a*: both music and sound vs. neither × linear trend across Seconds 1–7. Interaction Contrast 3b: sound vs. no sound × change from baseline (Second 1 vs. Seconds 2–7).
- *Interaction Contrast 4a*: only music vs. only sound × linear trend across Seconds 1–7. Interaction Contrast 4b: sound vs. no sound × change from baseline (Second 1 vs. Seconds 2–7).

26.5.4 Some Findings from the Tonic Analysis

We used a two-way repeated-measures factorial analysis of variance (ANOVA) using sound and music as independent variables with two levels (on = audible or off = inaudible) and facial EMG (brow, eye, cheek), EDA, and GEQ dimensions as dependent variables. Before the analysis, average values of psychophysiological measures were normalized using logarithmic transformation. Using the tonic data (in this context, we understand tonic as measured over a period of time albeit responding to an experimental condition) we tested significant differences between the factors sound and music. We ran 2×2 ANOVAs for each EMG measure. The ANOVAs showed no statistical differences for EMG measurements. We have to assume that neither sound nor music, nor the interaction of sound and music, had a significant accumulative effect on EMG measurement. No threshold values existed to classify EDA results as either activation or deactivation in the arousal dimension. Higher arousal values could indicate a more exciting experience in any of the conditions. Using a 2×2 ANOVA, we tested the effects of the independent variables, as for the results of EMG, but no significant cumulative effects were found. Thus, we assumed that neither sound nor music, nor the interaction of sound and music, had a significant effect on EDA measurement.

For the questionnaire data, we tested the effects of sound and music with a 2×2 ANOVA and found a main effect of sound on all seven dimensions of the GEQ and an interaction effect of sound and music on tension and flow. Absence or presence of sound influences all subjective GEQ dimensions, but we could further determine this effect by an interaction of tension and flow. The more positive dimensions of the GEQ (Flow, Positive Affect, Competence, Immersion, Challenge) were rated higher when the sound of the game was playing, while the more negative dimensions (Negative Affect, Tension) were rated lower. When sound was turned off, we could see the opposite effect. Based on these subjective results, game sound is crucial for a subjectively positive gameplay experience.

We also found an interaction of sound $\times$ music on the GEQ dimensions tension and flow. Flow was rated highest when sound was on and music was off, but received the lowest scores when everything was turned off. When sound and music were on, the flow rating was a slightly lower than in the sound on/music off condition. The experience ratings were even lower when sound was turned off and music remained on. Turning on non-diegetic music (that is music that does not directly relate to gameplay) seemed to dampen the flow experience (which was more polarized in positive and negative dimensions) when the differences of sound were taken into account. Regarding tension ratings, when music was on and sound was on, tension was experienced lowest while when music was on and sound was off, tension was rated highest. There was not much difference when music was off.

26.5.5 Some Findings from the Event-Based Analysis

For the event-based analysis, we used a linear mixed model analysis procedure. We found that regardless of the condition, EMG activity for all investigated muscle areas (brow, eye, and cheek) presented a statistically significant quadratic increase. In Condition 1 (sound on, music off), Contrast 1 (first second vs. seconds 2–7) revealed that the response to a death event was a significant increase in EMG activity

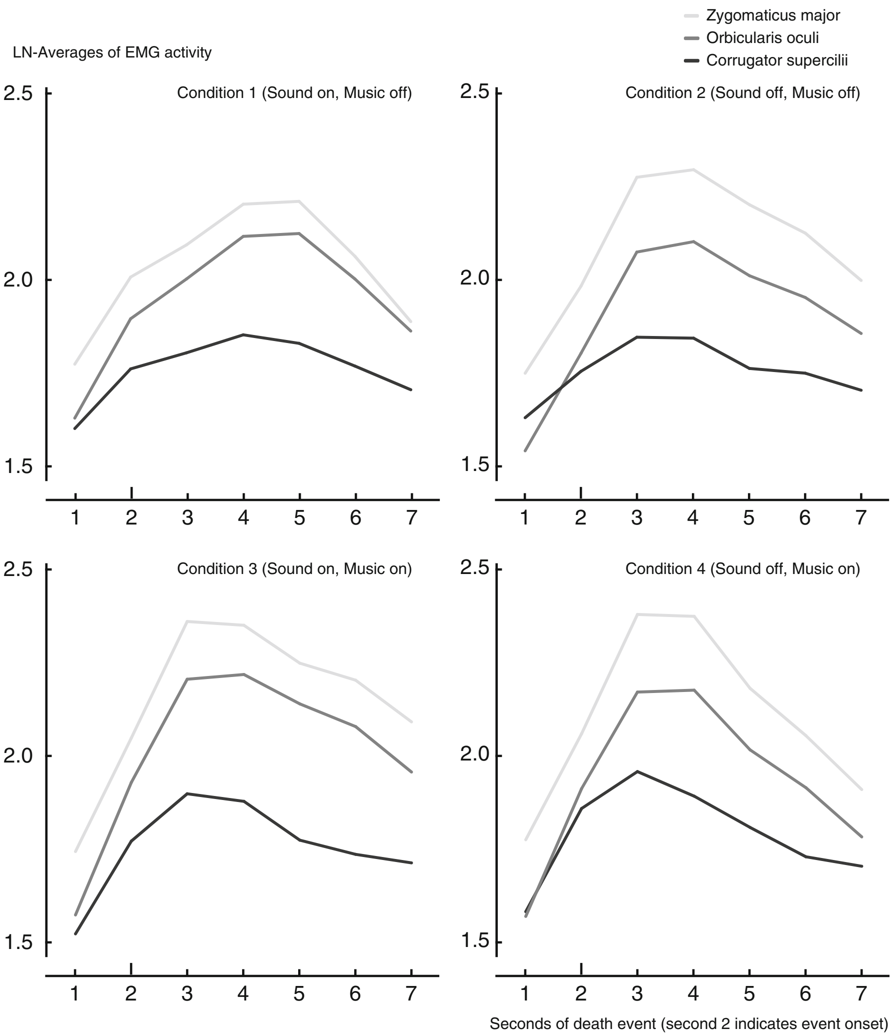

Fig. 26.8 Averages of EMG activity during each of the four conditions for cheek (ZM), eye (OO), and brow (CS) muscles

for CS, OO, and ZM muscle areas ($p<.001$). Contrast 2, testing linear trend, was significant for OO EMG activity ($p=.003$), but not for others.

Results of Contrast 3 showed that the trend was quadratic (first rising and then declining) for all EMG activity measures ($p<.001$, see Fig. 26.8). This tendency was repeated in Condition 2 (both sound and music off), Condition 3 (both sound and music on), and Condition 4 (sound off, music on): Contrast 1 showed that the response was increasing in relation to the baseline second for EMG activity over all muscle areas (all $p<.001$), and Contrast 3 that the response was quadratic, that is, first increasing but decreasing within 7 s (all $p<.001$). Contrast 2 yielded significant associations in Condition 2, where sound and music were both turned off, for eye EMG ($p < .001$) and cheek EMG activity ($p = .010$), like it did in Condition 3, where both sound and music were turned on, ($p < .001$ and $p = .007$, respectively). In Condition 4 (sound off, music on) none of the Contrast 2 tests showed significant associations. However, in all the cases where Contrast 2 was significant suggesting linear trend, the t-value for contrast 3 was higher, revealing a trend that was predominantly quadratic. Thus, the results of Contrasts 2 and 3 together indicate that in the most cases, the response to the death event is not an increase in long-term EMG activity level, but rather a transitory peak in EMG activity. The response peaked around Second 3 or 4 in all conditions except Condition 1, where the peak occurred approximately 1 s later. In summary, no condition had an effect on the EMG responses elicited by the death event, since the response was significant but similar in all conditions.

The main effects of death events on electrodermal activity showed less uniform responses. In Conditions 1 (sound on, music off) and 3 (sound on, music on), none of the contrasts revealed significant trends; that is, there was no change from the baseline (first second), no linear, and no quadratic trend in response to the event (see Fig. 26.9).

The trend for EDA in Condition 3 appears linear in visual inspection, but because the amount of death events in this condition was lower than in others and it did not quite reach statistical significance ($p = .064$).

In Condition 2 (sound off, music off), both positive linear and negative quadratic trends ($p = .023$ and .013 for contrasts 2 and 3, respectively) were found, latter being stronger. Only Contrast 2 showed a significant trend in Condition 4 (sound off, music on) ($p = .031$), revealing a linear increase in EDA as a response to the event.

We tested the interaction between the condition and linear trend of EDA over 7 s using contrast analyses. None of the interactions between condition and quadratic trends showed significant associations. Whereas all interactions using linear and quadratic trend EMG activity were non-significant, Interaction Contrasts 2b and 3b testing change from baseline (Second 1) to response (Seconds 2–7) showed that the brow EMG activity level rose in response to the death event more when music was on than when it was off ($p = .018$), and more when both music and sound were on vs. when they were both off ($p = .017$).

For EDA, Interaction Contrast 1a, testing the interaction of linear trend and the effect of sound vs. no sound, showed that the event prompted a greater linear increase when the sound was off compared to condition where the sound was on (regardless of music). That is, the participants responded with greater arousal to the death event when there were no sounds. Interaction Contrast 2a, testing the interaction of linear trend and the effect of music vs. no music, similarly showed that the event prompted a greater linear increase (greater arousal) when the music was on, as compared to when music was off (regardless of sound). Interaction Contrast 3a was not significant, suggesting that there was no difference in the EDA response whether both music and sound were on or off. Interaction Contrast 4a, testing the interaction of linear trend and the effect of only music vs. only sound, revealed that the event elicited a greater linear increase in EDA when music was on and sound

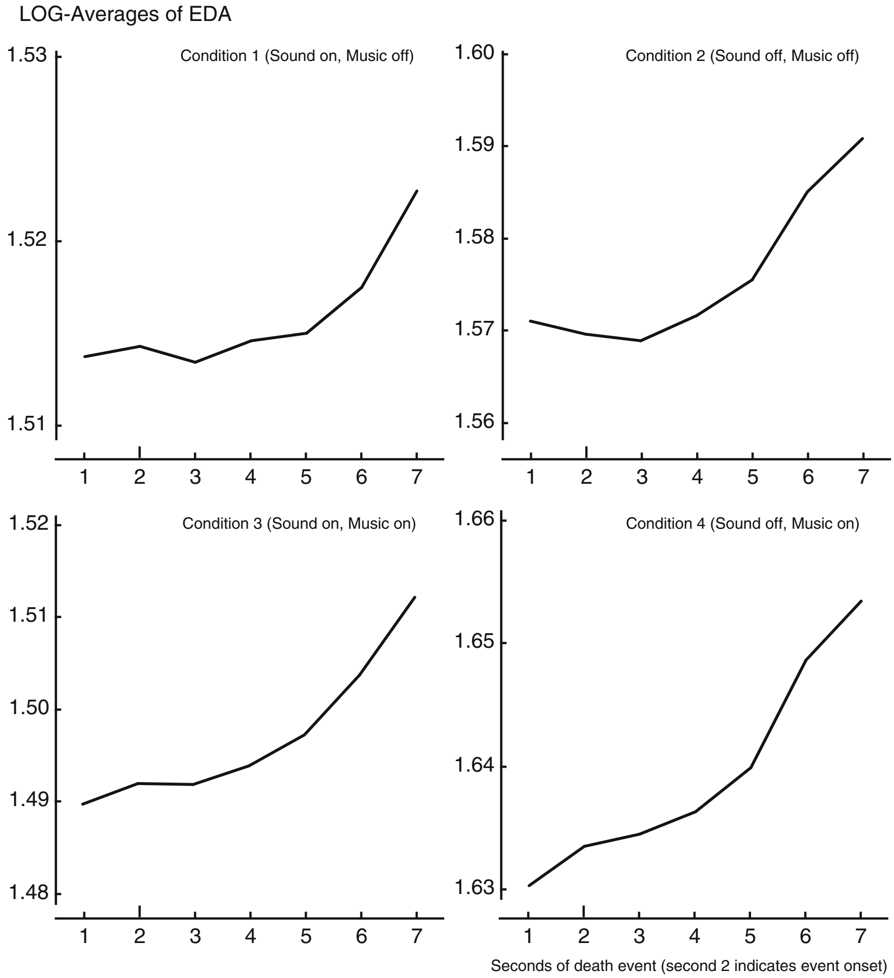

Fig. 26.9 EDA averages during each of the four conditions for cheek (ZM), eye (OO), and brow (CS) muscles. Note the different scales for each condition in the figure

was off, compared to the opposite condition. Interaction Contrast 4b showed that this linear increase was also a significant increase as compared to the baseline (Second 1; $p=.021$). In conclusion, the EDA response to the death event increased when the sound was off or music was on.

26.5.6 Takeaway from the Case Study

This case study demonstrates two different approaches to analyzing physiological data and the different the results that these types of analyses yield. For this particular study design regarding sound and music in games, we could take away that tonic effects (meaning effects that are measured throughout a game session) can be better gauged using questionnaires (which again emphasizes the importance of using mixed methods in Games User Research). For gameplay related events, in this case death events, physiological measures seem to provide some insights into short physiological experience peaks that seem to occur 3–4s after an event onset. This is particularly interesting since we are currently moving toward mixed-method automatic marker-driven analysis of physiological metrics and gameplay metrics. While the conditions did not yield a significant response, the death event was meaningful to the gameplay in terms of its physiological response. The interpretation of psychophysiological measures and data depends on the context and research paradigm that is being followed. We should evaluate other statistical tests and a non-linear analysis of the relationship between physiology and player experience in the future. For example, applying artificial neural networks to the analysis process might yield some more interesting relationships than the standard inferential tests used here. The interpretation of these measures in a gaming context requires more validation before we can ensure consistent results and guidelines useful to the industry. One approach in this direction has recently been reported in a study by Mirza-Babaei et al. (2013), which addresses the usefulness of biometric measures in a games user research scenario and provides user testing guidelines for games user researchers interested in physiological measures.

26.6 How Can Physiological Metrics Be of Value to the Game Industry?

Given that this is only a brief introduction of psychophysiological measures for the game industry, we also have to discuss in which scenarios physiological game research is useful. Fairclough (2011a, b) suggested a thought experiment, where he outlines ten suggestions for improving the use physiological metrics in game user research. Many of these suggestions should be implemented by game user researchers for physiological metrics to work most effectively in an industry setting.

- **Physiological metrics can be recorded continuously during a game user research session without interrupting play.** This makes these methods superior to subjective measures that either break the experience (by interrupting and prompting with questions) or introduce memory bias (by asking questions about the game in retrospect). The only downside of physiological metrics is that the player has to wear sensors and that some might find this intrusive (although based on personal experience, many players forget that they wear sensors a few minutes into the game).
- **A game user researcher interested in physiological assessment of players needs to be well-informed about what each sensor type measures.** Company executives and the marketing department need to understand this is no emotion quantifier or thought printer. Sensors measure electrical activity that comes from

motor, skin, or brain activity and depending on the area of application allows some conclusion of the activity of the body area being measured. This also means that we need an experience vocabulary working from a high level psychological concept (engagement) toward the low-level body response (sympathetic activation → higher heart rate). Inferences made from low-level body responses to high level concepts in comparison are always difficult to withstand closer scientific inspection.

- **For capturing player experience, a hypothesis-driven approach is suggested**, where only one particular aspect of experience is under investigation. Ideally this aspect is well-defined in related literature so that, for example, we only investigate positive and negative emotional responses to a certain game event or game area or that we investigate cognitive workload during a game tutorial.
- **To establish a link between ideal player experience and the corresponding physiological responses, we should investigate responses to key aspects when naïve participants play the most successful games of the industry** (in terms of financial and critical success). If we could find out what physiological responses relate to the player experiences that drive the success of these games, we could work towards establishing a physiological success metric. This would be truly valuable for the game industry.
- **In every aspect of physiological experimentation we need to be aware that the human body is still present in the real world while playing a video game.** Our nervous system therefore responds to real-world stimulation coming from our environment as well as to cognitive stimulation. These contextual influences (that may be overlooked during the screening of a participant) can result in changes in emotion or motivation during the experiment. Influences such as room temperature, movement, drugs, chemicals, noise, and many more can also introduce contextual bias into our interpretation of physiological activity. In the end, it is important to keep in mind how sensitive our nervous system really is when interpreting physiological metrics.
- **Physiological metrics do not distinguish between physical activity and psychological events.** Three components are involved in recording physiological metrics: external physical activity, internal emotional activity, and internal cognitive activity (Stemmler et al. 2007).
- **Given what we now know about physiological responses, we will always have a certain signal-to-noise ratio in our physiological metrics.** We can counteract the amount of noise by enforcing a strict experimental protocol in a very controlled environment or by recording all possible confounds with additional sensors (e.g., temperature, noise, light) to remove their influence during analysis.
- **Before testing players, it is important to carefully record their demographic background**, including their skill level and past game preferences and experiences. Novelty and habituation can impact physiological responses considerably.

- **It is important to create the different experimental conditions carefully** within a systematically manipulated environment (e.g., a game engine). Ideally, we only change one variable at a time.
- **Metrics should be tracked together.** Other gameplay tracking metrics can be considered overt behavior markers in the game world as they are visible instantly whereas physiological metrics are covert measures that are not always visible directly. Both metrics should be tracked together and a possible relationship between them should be explored using statistical analyses. Subjective responses are best recorded after physiological measurement.

We can conclude that psychophysiological measures in games should not be used alone, but always in conjunction with other measures to establish relationships between player experience and physiological responses. Much work remains to be done in this area before it becomes part of the everyday testing of game user researchers. However, given recent advances by sensor manufacturers, this technology will eventually be more common in game user research. When we start using it to improve our games, we will always need to remember its sensitivity and the possible contextual influences, so that all interpretations should be understood with a grain of salt.

26.7 Next Steps in Using Physiological Metrics

Imagine you would like to use these measures in your company, a good way to get started with most of them would be to consider that if you are measuring people, you always have to account for individual differences and you have to be very clear on your research goal. Can you obtain a result for your research question in an easier and more cost-effective way? If you answer is yes, then physiological measures might not be for you at this point. I would suggest the easiest way to get started with psychophysiological measures would be to obtain or build a system to measure electrodermal response, not only because it has a direct relationship to arousal or excitement, but also because this type of data can be easily processed and analyzed. We are currently working on methods to make the interpretation of electrodermal activity within a gaming context easier to interpret and more accessible for the games industry (e.g., using biometric storyboards) (Mirza-Babaei 2011; Mirza-Babaei et al. 2012).

Another option that we have worked with in the past and that we will continue working on in the future is building physiological measurement systems that track physiological metrics and player events together (Nacke et al. 2008; Kivikangas et al. 2011b). For these systems, a few key issues have to be considered:

- Synchronized time data. Often the timestamp is used as the identification key to merge different log data. If you are logging physiological data on a different system than your other metrics, make sure you timestamp is synchronized across the network.

- Data storage. Physiological log files become large over time (similar to video recording files). It is a good idea to have a backup or extended storage solution in mind when starting to seriously collect physiological metrics data.
- Choose the right gameplay hooks. Depending on your analysis, you will need to make a decision what gameplay hooks are relevant in your log data to be related to physiological measures.
- Set up a demo version of your game where players can focus on one action at a time. This will make your psychophysiological analysis much easier since you are concentrating on one key variable.

If you plan to use more elaborate physiological metrics, such as EEG, an expert opinion is often helpful to get started. EEG is hard to interpret and there is still not enough evidence to show its usefulness for researching interesting game situations. In the end all the research needs to tie the ideas back to improving game design. EEG has a lot of potential to investigate meaningful decisions at certain key points in your game and once we have better signal processing and automated analysis techniques, it will likely see a larger adoption in the game industry.

This chapter might have shown to you that physiological evaluation is a field that requires some consideration and is not as simple just sticking a sensor on persons and getting precise emotional readouts of their activities. Therefore, I can only recommend considering the goal of your game user research study before you decide on whether or not you can use sensors to solve your problem. Physiological sensors provide a great addition to other quantitative game metrics, because of their potential for adding emotional meaning to your data. If you want a complete and robust picture of the user experience when playing games, you should consider adding physiological measures to your quantitative toolbox since they will provide rich evaluation possibilities over time.

Acknowledgements Thanks to Pejman Mirza-Babaei for his initial feedback and Anders Drachen for his encouragement and feedback on this chapter. This research was supported by the Network of Centres of Excellence (NCE), Graphics, Animation and New Media (GRAND), NSERC, and SSHRC IMMERSe (895-2011-1014).

Additional Reading for Psychophysiology

1. Andreassi, J. L. (2000). Human behavior and physiological response. *Psychophysiology 4*, 1–12.
2. Hazlett, R. L., & Benedek, J. (2007). Measuring emotional valence to understand the user's experience of software. *International Journal of Human-Computer Studies 65(4)*, 306–314. doi:10.1016/j.ijhcs.2006.11.005
3. Baumgartner, T., Valko, L., Esslen, M., & Jäncke, L. (2006). Neural correlate of spatial presence in an arousing and noninteractive virtual reality: An EEG and psychophysiology study. *CyberPsychology & Behavior 9(1)*, 30–45. doi:10.1089/cpb.2006.9.30
4. Hudlicka, E. (2003). To feel or not to feel: The role of affect in human–computer interaction. *International Journal of Human-Computer Studies 59(1–2)*, 1–32. doi:10.1016/s1071-5819(03)00047-8

5. Hoffman, J. E. (1998). Visual attention and eye movements. In H.E. Pashler (Ed.) *Attention*. New York: Psychology Press (Taylor & Francis).
6. Fried, R. (1982). On-line analysis of the GSR. *Integrative Physiological and Behavioral Science 17(2)*, 89–94. doi:10.1007/bf03002004

About the Author

Dr. Lennart E. Nacke is an Assistant Professor for Human-Computer Interaction and Game Science in the Faculty of Business and Information Technology at the University of Ontario Institute of Technology (UOIT) in Canada and currently also a visiting lecturer at the Department of Informatics at the University of Sussex. He is working on projects that deal with the cognitive and emotional side of playing games. His goal is to describe, analyse and have computers and machines react to the thoughts and feelings you have when you are fully engaged in playing a video game. With this research we can start understanding the effects of video games (positive and negative). Using body sensor technology, for example, sensors for brainwaves, heart rate, or muscle tension, Dr. Nacke taps into some of video gaming's most motivating features to improve our physical fitness and mental wellbeing. His research is funded nationally and internationally and his publications have won best paper honourable mention (awarded to the top 5%) and best paper awards (awarded to the top 1%) at the premier human-computer interaction conferences CHI 2011 and CSCW 2012. He also writes a blog at www.acagamic.com and you can find out more about his research group at http://hcigames.businessandit.uoit.ca

References

Andreassi, J. L. (2006). *Psychophysiology: Human behavior and physiological response*. London: Routledge.

Barrett, L. F. (2006). Are emotions natural kinds? *Perspectives on Psychological Science, 1*(1), 28–58. doi:10.1111/j.1745-6916.2006.00003.x.

Barrett, L. F., Mesquita, B., Ochsner, K. N., & Gross, J. J. (2007). The experience of emotion. *Annual Review of Psychology, 58*(1), 373–403. doi:10.1146/annurev.psych.58.110405.085709.

Boucsein, W. (1992). *Electrodermal activity*. New York: Plenum Press.

Bradley, M. M., & Lang, P. J. (2007). Emotion and motivation. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (3rd ed., pp. 581–607). New York: Cambridge University Press.

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion, 1*(3), 276–298.

Brathwaite, B., & Schreiber, I. (2008). *Challenges for game designers*. Boston: Charles River Media.

Brave, S., & Nass, C. (2002). Emotion in human-computer interaction. In J. A. Jacko & A. Sears (Eds.), *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications* (pp. 81–96). Mahwah: Lawrence Erlbaum Associates.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1999). The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology, 76*(5), 839–855. doi:10.1037/0022-3514.76.5.839.

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. G. (Eds.). (2007). *Handbook of Psychophysiology*. Cambridge, UK: Cambridge University Press.

Calvo, R. A., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing, 1*(1), 18–37. doi:10.1109/t-affc.2010.1.

Cannon, W. B. (1927). The James-Lange theory of emotions: A critical examination and an alternative theory. *The American Journal of Psychology, 39*(1/4), 106–124. doi:10.2307/1415404.

Chatrian, G. E., Lettich, E., & Nelson, P. L. (1988). Modified nomenclature for the "10%" electrode system. *Journal of Clinical Neurophysiology: Official Publication of the American Electroencephalographic Society, 5*(2), 183–186.

Dalgleish, T., Power, M. J., & Power, M. (2000). *Handbook of cognition and emotion*. Chichester/New York: Wiley.

Dalgleish, T., Dunn, B. D., & Mobbs, D. (2009). Affective neuroscience: Past, present, and future. *Emotion Review, 1*(4), 355–368. doi:10.1177/1754073909338307.

Damasio, A. R. (1994). *Descartes' error*. New York: G.P. Putnam.

Darrow, C. W. (1964). Psychophysiology, yesterday, today, and tomorrow. *Psychophysiology, 1*(1), 4–7. doi:10.1111/j.1469-8986.1964.tb02615.x.

Davidson, R. J., Scherer, K. R., & Goldsmith, H. H. (2003). *Handbook of affective sciences*. Oxford: Oxford University Press.

Döveling, K., Von Scheve, C., & Konijn, E. A. (2010). *The Routledge handbook of emotions and mass media*. London: Routledge.

Duchowski, A. T. (2007). *Eye tracking methodology: Theory and practice* (2nd ed.). New York: Springer.

Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In J. Cole (Ed.), *Nebraska symposium on motivation, 1971* (pp. 207–283). Lincoln: University of Nebraska Press.

Ekman, P. (1992a). Are there basic emotions? *Psychological Review, 99*(3), 550–553.

Ekman, P. (1992b). An argument for basic emotions. *Cognition and Emotion, 6*(3/4), 169–200. doi:10.1080/02699939208411068.

Ekman, P., & Davidson, R. J. (1994). *The nature of emotion: Fundamental questions*. New York: Oxford University Press.

Engl, S., & Nacke, L. E. (2013, February). Contextual influences on mobile player experience – A game user experience model. *Entertainment Computing 4*(1), 83–91, ISSN 1875-9521. doi:10.1016/j.entcom.2012.06.001.

Fairclough, S. H. (2011a). Biometrics and evaluation of gaming experience part two: A thought experiment. http://www.physiologicalcomputing.net/?p=1760. Last accessed 19 Mar 2013.

Fairclough, S. H. (2011b). Biometrics, game evaluation, and user XP: Approach with caution. http://www.physiologicalcomputing.net/?p=1698. Last accessed 19 Mar 2013.

Fridlund, A. J., & Cacioppo, J. T. (1986). Guidelines for human electromyographic research. *Psychophysiology, 23*(5), 567–589. doi:10.1111/j.1469-8986.1986.tb00676.x.

Frijda, N. H. (1986). *The emotions*. Cambridge, UK: Cambridge University Press.

Gokcay, D., & Yildirim, G. (2010). *Affective computing and interaction: Psychological, cognitive, and neuroscientific perspectives*. Hershey: IGI Global.

Gualeni, S., Janssen, D., & Calvi, L. (2012). How psychophysiology can aid the design process of casual games: A tale of stress, facial muscles, and paper beasts. In *Proceedings of the international conference on the Foundations of Digital Games (FDG '12)* (pp. 149–155). Raleigh: ACM. doi:10.1145/2282338.2282369.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 239–250). Amsterdam: North Holland Press.

Hazlett, R. L. (2006). Measuring emotional valence during interactive experiences: Boys at video game play. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (*CHI '06*) (pp. 1023–1026). Montréal: ACM. doi:10.1145/1124772.1124925.

Hazlett, R. L. (2008). Using biometric measurement to help develop emotionally compelling games. In K. Isbister & N. Schaffer (Eds.), *Game usability: Advancing the player experience* (pp. 187–205). Burlington: Morgan Kaufmann.

Hugdahl, K. (1995). *Psychophysiology: The mind-body perspective*. Cambridge: Harvard University Press.

Izard, C. E. (1972). *Patterns of emotions: A new analysis of anxiety and depression*. New York: Academic.

James, W. (1884). What is an emotion? *Mind, 9*(34), 188–205.

Jasper, H. H. (1958). Report of the committee on methods of clinical examination in electroencephalography. *Electroencephalography and Clinical Neurophysiology, 10*, 370–375.

Kivikangas, J. M., Chanel, G., Cowley, B., Ekman, I., Salminen, M., Järvelä, S., & Ravaja, N. (2011a). A review of the use of psychophysiological methods in game research. *Journal of Gaming & Virtual Worlds, 3*(3), 181–199. doi:10.1386/jgvw.3.3.181_1

Kivikangas, J. M., Nacke, L., & Ravaja, N. (2011b). Developing a triangulation system for digital game events, observational video, and psychophysiological data to study emotional responses to a virtual character. *Entertainment Computing, 2*(1), 11–16. doi:10.1016/j.entcom.2011.03.006.

Kleinginna, P. R., & Kleinginna, A. M. (1981). A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and Emotion, 5*(4), 345–379.

Lacey, J. I. (1959). Psychophysiological approaches to the evaluation of psychotherapeutic process and outcome. In E. A. Rubinstein & M. B. Parloff (Eds.), *Research in psychotherapy* (pp. 160–208). Washington, DC: American Psychological Association. doi:10.1037/10036-010.

Lane, R. D., & Nadel, L. (2002). *Cognitive neuroscience of emotion*. New York: Oxford University Press.

Lang, P. J. (1995). The emotion probe. Studies of motivation and attention. *American Psychologist, 50*, 372–385.

Lange, C. G. (1912). The mechanism of the emotions. In B. Rand (Ed.), *The classical psychologists* (pp. 672–684). Boston: Houghton Mifflin.

Larsen, J. T., McGraw, A. P., & Cacioppo, J. T. (2001). Can people feel happy and sad at the same time? *Journal of Personality and Social Psychology, 81*(4), 684–696. doi:10.1037/0022-3514.81.4.684.

Lazarus, R. S. (1968). Emotions and adaptation: Conceptual and empirical relations. In *Nebraska symposium on motivation*. Lincoln: University of Nebraska Press.

LeDoux, J. (1998). *The emotional brain*. London: Orion Publishing Group.

Lewis, M., Haviland-Jones, J. M., & Barrett, L. F. (2010). *Handbook of emotions* (3rd ed.). New York: Guilford Press.

Lykken, D. T., & Venables, P. H. (1971). Direct measurement of skin conductance: A proposal for standardization. *Psychophysiology, 8*(5), 656–672. doi:10.1111/j.1469-8986.1971.tb00501.x.

Mandryk, R. (2008). Physiological measures for game evaluation. In K. Isbister & N. Schaffer (Eds.), *Game usability: Advice from the experts for advancing the player experience* (pp. 207–235). Burlington: Morgan Kaufmann.

Mandryk, R. L., & Atkins, M. S. (2007). A fuzzy physiological approach for continuously modeling emotion during interaction with play environments. *International Journal of Human Computer Studies, 65*(4), 329–347. doi:10.1016/j.ijhcs.2006.11.011.

Mandryk, R. L., Inkpen, K. M., & Calvert, T. W. (2006). Using psychophysiological techniques to measure user experience with entertainment technologies. *Behavior & Information Technology, 25*(2), 141–158. doi:10.1080/01449290500331156.

Mirza-Babaei, P. (2011). Understanding the contribution of biometrics to games user research. In *Proceedings of "Think Design Play: The fifth international conference of the Digital Research Association" (DIGRA 2011)*, Hilversum. http://www.digra.org/dl/display_html?chid=11310.43254.pdf. Last accessed 19 Mar 2013.

Mirza-Babaei, P., Nacke, L. E., Fitzpatrick, G., White, G., McAllister, G., & Collins, N. (2012). Biometric storyboards: Visualising game user research data. In *Proceedings of the 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts (CHI EA '12)* (pp. 2315–2320). Austin: ACM. doi:10.1145/2223656.2223795.

Mirza-Babaei, P., Nacke, L., Gregory, J., Collins, N., & Fitzpatrick, G. (2013). How does it play better? Exploring user testing and biometric storyboards in games user research. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems (CHI 2013)*. To appear.
Nacke, L. E. (2009). *Affective ludology: Scientific measurement of user experience in interactive entertainment*. Ph.D. thesis, Blekinge Institute of Technology, Karlskrona, Sweden.
Nacke, L., & Lindley, C. A. (2008). Flow and immersion in first-person shooters: Measuring the player's gameplay experience. In *Proceedings of the 2008 conference on future play: Research, play, share (Future Play '08)* (pp. 81–88). Toronto: ACM. doi:10.1145/1496984.1496998.
Nacke, L., Lindley, C., & Stellmach, S. (2008) Log who's playing: Psychophysiological game analysis made easy through event logging. In P. Markopoulos, Bd. Ruyter, W. IJsselsteijn, D. Rowland (Eds.), *Proceedings of Fun and Games, Second International Conference*, Eindhoven, The Netherlands, October 20–21, 2008 (Lecture Notes in Computer Science) (pp. 150–157). Dordrecht: Springer. doi:10.1007/978-3-540-88322-7_15
Nacke, L. E., Grimshaw, M. N., & Lindley, C. A. (2010). More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game. *Interacting with Computers, 22*(5), 336–343. doi:10.1016/j.intcom.2010.04.005.
Nacke, L. E., Kalyn, M., Lough, C., & Mandryk, R. L. (2011). Biofeedback game design: Using direct and indirect physiological control to enhance game interaction. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems (CHI '11)* (pp. 103–112). Vancouver: ACM. doi:10.1145/1978942.1978958.
Nishimoto, S., Vu An, T., Naselaris, T., Benjamini, Y., Yu, B., & Gallant Jack, L. (2011). Reconstructing visual experiences from brain activity evoked by natural movies. *Current Biology: CB, 21*(19), 1641–1646.
Panksepp, J. (2004). *Affective neuroscience: The foundations of human and animal emotions*. Oxford, UK: Oxford University Press.
Pelachaud, C. (2012). *Emotion-oriented systems*. Hoboken: Wiley.
Picard, R. W. (1997). *Affective computing*. Cambridge, MA: MIT Press.
Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research, and experience* (Vol. 1, 3, pp. 3–31). New York: Academic.
Plutchik, R. (1991). *The emotions*. Lanham: University Press of America.
Prokasy, W. F., & Raskin, D. C. (1973). *Electrodermal activity in psychological research*. New York: Academic.
Ravaja, N. (2004). Contributions of psychophysiology to media research: Review and recommendations. *Media Psychology, 6*(2), 193–235. doi:10.1207/s1532785xmep0602_4.
Ravaja, N., Turpeinen, M., Saari, T., Puttonen, S., & Keltikangas-Järvinen, L. (2008). The psychophysiology of James Bond: Phasic emotional responses to violent video game events. *Emotion, 8*(1), 114–120. doi:10.1037/1528-3542.8.1.114.
Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161–1178. doi:10.1037/h0077714.
Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*(1), 145–172. doi:10.1037/0033-295x.110.1.145.
Sander, D., & Scherer, K. R. (2009). *The Oxford companion to emotion and the affective sciences*. Oxford/New York: Oxford University Press.
Schachter, S. (1964). The interaction of cognitive and physiological determinants of emotional state. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 1, pp. 49–80). New York: Academic.
Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review, 69*, 379–399.
Scherer, K. R. (1984). Emotion as a multicomponent process: A model and some cross-cultural data. *Review of Personality & Social Psychology, 5*, 37–63.
Scherer, K. R., Bänziger, T., & Roesch, E. B. (2010). *Blueprint for affective computing: A sourcebook*. Oxford/New York: Oxford University Press.
Smith, M. E., Gevins, A., Brown, H., Karnik, A., & Du, R. (2001). Monitoring task loading with multivariate EEG measures during complex forms of human-computer interaction. *Human*

Factors: The Journal of the Human Factors and Ergonomics Society, 43(3), 366–380. doi:10.1518/001872001775898287.

Stemmler, G., Aue, T., & Wacker, J. (2007). Anger and fear: Separable effects of emotion and motivational direction on somatovisceral responses. *International Journal of Psychophysiology, 66*(2), 141–153. doi:10.1016/j.ijpsycho.2007.03.019.

Stern, R. M., Ray, W. J., & Quigley, K. S. (2001). *Psychophysiological recording* (2nd ed.). New York: Oxford University Press.

Tassinary, L. G., Cacioppo, J. T., & Vanman, E. J. (2000). The skeletomotor system: Surface electromyography. *Handbook of Psychophysiology, 2*, 274–278.

Waterink, W., & van Boxtel, A. (1994). Facial and jaw-elevator EMG activity in relation to changes in performance level during a sustained information processing task. *Biological Psychology, 37*(3), 183–198. doi:10.1016/0301-0511(94)90001-9.

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology, 76*(5), 820–838.

Yannakakis, G. N., & Hallam, J. (2008). Entertainment modeling through physiology in physical play. *International Journal of Human Computer Studies, 66*(10), 741–755. doi:10.1016/j.ijhcs.2008.06.004.

Yannakakis, G. N., Hallam, J., & Lund, H. H. (2008). Entertainment capture through heart rate activity in physical interactive playgrounds. *User Modeling and User-Adapted Interaction, 18*(1), 207–243.