A review of the use of psychophysiological methods in game research

KIVIKANGAS, J. Matias, et al.

Abstract

This article reviews the psychophysiological method in game research. Psychophysiological measurements provide an objective, continuous, real-time, noninvasive, precise and sensitive way to assess the game experience. However, the best results require controlled experiments with careful monitoring of variables, large enough sample sizes and expertise in electrical signal processing. We briefly explain the theory behind the method and present the most useful measures: electromyography (EMG), electrodermal activity (EDA), electroencephalography (EEG) and cardiac measures. We review previous studies that have used psychophysiological measures in game research and illustrate some future directions. Our article covers several research lines using the psychophysiological method in game studies, and offers a comprehensive list of references for those interested in the field.

Reference


DOI: 10.1386/jgvw.3.3.181_1
Review on psychophysiological methods in game research

J. Matias Kivikangas, Inger Ekman, Guillaume Chanel, Simo Järvelä, Ben Cowley, Mikko Salminen, Pentti Henttonen, Niklas Ravaja
Center for Knowledge and Innovation Research, Aalto University
PO BOX 21255
00076 Aalto, Finland
+358 50 3120 922
matias.kivikangas@aalto.fi, first.last@aalto.fi, sjarvela@aalto.fi

ABSTRACT
This paper reviews the psychophysiological method in game research. The use of psychophysiological measurements provides an objective, continuous, real-time, non-invasive, precise, and sensitive way to assess the game experience, but for best results it requires carefully controlled experiments, large participant samples and specialized equipment. We briefly explain the theory behind the method and present the most useful measures. We review previous studies that have used psychophysiological measures in game research, and provide future directions.

Keywords
Psychophysiology, review, game experience, facial EMG, EDA, EEG, ECG, HR, digital games

INTRODUCTION
Lately, psychophysiological measures (e.g., ECG, EMG, skin conductance) have gained some attention in (digital) game research and interest is rapidly growing. However, the various studies using psychophysiological measures do not yet form a common field: instead they have been from different scientific backgrounds and with different motivations. Studies have attempted to capture the game experience or demonstrate the psychological effects of gaming with physiological evidence, they have used real-time measures for adapting game features to the players’ physiology as well as utilized various sorts of physiological indicators to support the evaluation of design choices. Thus, we have a number of separate results for many separate research questions, but very little accumulated knowledge that could be used for answering more precise research questions or for creating theoretical syntheses. This critical review aims to contribute to creating such a body of knowledge.

In the next two sections, we very briefly introduce the theory behind the psychophysiological method and the practical use of it. After that, the previous research section details the most pertinent and recent work, conveniently subdivided into areas of special research interest. Finally, a section on future directions discusses the authors’ views on the most promising lines of inquiry.

PSYCHOPHYSIOLOGICAL RESEARCH
Psychophysiological research is defined as using physiological signals to study psychological phenomena [7]. However, physiological processes are typically not related to psychological phenomena with a one-to-one relationship, which sometimes makes interpreting the signals challenging. Digital games are a difficult stimuli in this sense - they typically provide sensory output in at least two modalities and sometimes employ physically complicated input devices; they also typically require complicated cognitive processing on different interpretative levels and operate on time scales from fractions of a second to several hours [34]. The motivations for playing also vary from person to person and from time to time [30]. As most of the psychophysiological reasoning employed originates from much simpler experimental situations, it is not self-evident that the same associations hold for experiences with digital games. For instance, a consistent association (between physiology and psychology) found when the participants look at pictures standardized for emotional stimulation, might not occur equally when the participants experience the same emotions spontaneously during a game. Another complication is the lack of a commonly accepted theory on how digital game experience arises, that could be used in psychophysiological game research.

The practical challenge is to identify research questions that can be answered even when the game is complex and the psychological processes numerous, and then to create an experimental design with the proper and necessary controls so that no confounding variables will affect the results (cf. e.g. [43]). This is no minor task, as demonstrated by the large number of methodological problems in the current literature (including our own research; see section on previous research, below). If the game used in an experiment is not extremely simple and as such easily
controllable, the phenomenon of interest must be very strong or the sample size large enough that the reactions of interest are not confounded among noise. Statisticians advise [19] that a sample size of at least 28 is needed to reliably detect a large effect size, even when the assumptions of the population are met—and they rarely are. Thus, although measurements from a few participants can be informative in the sense that they may demonstrate a point or provide directions for further inspection, it is statistically implausible that such an experiment would reliably confirm anything beyond the sample.

**Practical pros and cons**

Physical reactions are part of the processes that underlie the (player’s) game experience. Hence, it is tempting to claim that the psychophysiological method provides a tool for measuring the game experience itself. In truth, it is only possible to tap into those parts of the experience that have recognized measurable physical concomitants. Nevertheless, in those areas, physiological measures can provide more objective and precise information of the player’s emotional and cognitive processes than is available by subjective methods.

The physiological processes measured are mostly involuntary. Therefore measurements are not contaminated by participant answering style, social desirability, interpretations of questionnaire item wording or limits of participant memory, nor by observer bias. Moreover, for studying game experiences the main benefit is that measurements can be recorded automatically and continuously (in real-time), without disturbing the participant’s natural behavior. Another benefit is the sensitivity of the psychophysiological method: measures are sensitive enough to pick up responses smaller than what the human eye can detect. Combined with other methods (e.g., self-report and observational data), psychophysiological methods add significant precision to studying the gaming experience.

As practical limitations, the data acquisition devices are typically expensive, and sufficient attention and time should be given to personnel training and device maintenance. Also, experimental preparations and procedures during testing (setting up the equipment, placing electrodes, testing the signals) take considerably more of the participants’ time than when using, for instance, questionnaires.

**THEORY AND MEASURES**

As there are more comprehensive looks at the theory and practice of psychophysiological methods elsewhere, in this paper we repeat only the very basics. *Handbook of Psychophysiology* [7] is strongly recommended for the definitive review of the methodology. There are also other papers that present introductions to the methods specifically in gaming context [40, 15]. Unfortunately there is no single, commonly accepted theory for game experience, so most of the theoretical framework used in game research is borrowed from other fields of study (for useful psychological theories in media research, see [53]).

**Valence and arousal**

A significant part of the game experience arises from emotional reactions [29]. According to the dimensional model of emotions (see [61, 37, 72] for different views), all emotions can be located within a few basic dimensions, typically valence (hedonic tone) and arousal (bodily activation). For example, joy and anger are not fundamentally different in quality (cf. basic emotions theory [17]): on the visceral level of automatic physical reactions they are created by the same two systems [50, 76].

Facial electromyography (EMG), which measures the electrical activity of facial muscles (see [71] for details), can be used for assessing positive and negative emotional valence (on the importance of facial expressions in emotional processing, see [6, 36]). The benefits of EMG compared to coding from a video are the automation, objectivity (no observer biases), temporal precision (milliseconds), and detection of even minuscule responses [5]. On the other hand, facial EMG measurements are sensitive to noise, both of technical origin (e.g., bad contact between electrode and skin) and from confounding sources of muscle activity, such as speaking and other social communication (see behavioral ecology view [17]).

Electrodermal activity (EDA) or skin conductance (often misleadingly called galvanic skin response, GSR, see below) is associated with emotional arousal [12, 36]. EDA data can be analyzed as skin conductance level (SCL), an aggregate over a certain period of time, or number of amplitude of discrete skin conductance responses (SCR, also electrodermal response, or GSR) to a specific phase or event. Electrodermal responses are slow (delay of one to four seconds), but in general, EDA is less sensitive to noise and less ambiguous than facial muscle and heart activity.

Cardiac activity (e.g., heart rate, HR; measured with electrocardiograph, ECG, or peripheral pulse oximeter) is among the most widely used physiological research methods, but because the heart and circulatory system is regulated by many different bodily processes, interpreting the signal’s relevance to the game context can be challenging. In different studies cardiac activity has been interpreted as an index of both valence and arousal, but also of attention, cognitive effort, stress, and orientation reflex during media viewing [53]. Still, cardiac measures have been used successfully in some game studies (see section on previous research, below).

**Attention**

Attention, or allocation of mental resources to a specific stimulus, causes physiological changes such as the orienting reflex, or parasympathetic activation (in certain circumstances detectable by EMG, EDA, or HR [53, 35, 75]). However, a game stimulus typically requires active coping, resulting in increased arousal that may mask such
subtle changes. This often makes it difficult to study attention (in the psychological sense) in games.

Electroencephalography (EEG) provides data about the brain’s electrical activity with millisecond accuracy. The signal is examined for event-related potentials (ERP) evoked by specific events, or for changes in the power of different frequency bands. Certain features of the signals have been shown to be associated with drowsiness and vigilant attention [11] or to reflect inactivity in the brain regions (smaller use of mental resources [67]). EEG has also been used to study the processing of visual emotional stimuli [1]. However, to this date, the use of EEG in game research has been sparse, perhaps due to the complicated nature of the signal, which combined with a complex stimulus produces a range of methodological challenges.

Other methods
More marginal methods, but still with some potential, include measurements of cortisol levels from participant saliva to investigate participant stress; measuring respiration for studying emotions or attention, or for providing control data when measuring cardiac activity; using eye gaze tracking and pupil size measurements for investigating arousal, cognitive effort, or attention level and its direction; and examining brain activity with magnetoencephalography (MEG) or functional magnetic resonance (fMRI) (see [7] for details of each method). To extend from psychophysiological measurements, there is some evidence that body movement and position (measured by acceleration sensors or position cameras) might be associated with attention, interest, and emotions [21, 27].

PREVIOUS RESEARCH
As mentioned, previous research has rarely been systematic on a specific issue or question, but mostly separate studies on separate issues. The following sections outline the most prominent lines of interest within this area of study.

Studies validating the method in gaming context
Although the corpus of physiological game studies is growing, there have been few attempts at validating proven psychophysiological results in the context of digital games; that is, finding out if the evidence from studies using pictures, short auditory or video clips and others (e.g. [36]) apply when studying a multimodal game experience. Earliest are from Hazlett [24] and Mandryk et al [40], who have presented studies, although with small sample sizes, supporting the use of psychophysiological measures in game research. More recently, Nacke [45] and others published studies as an attempt at a common methodology for a design-oriented approach. Their papers have a look at EEG [48], EDA and HR [47, 14], and facial EMG [47] and conclude with a recommendation for the methodology in a game context.

The FUGA project1 examined the construct validity, reliability, and predictive validity of facial EMG, EDA and EEG, but also fMRI, eye tracking, and physical activity and behavioral indicators in this regard [21, 65, 33, 27]. In general the conclusions supported the use of these methods in game context, but clearly showed the necessity of proper experimental design and that care must be taken in interpretation of the signals: otherwise, for instance, self-reported and physiologically indexed emotions may turn out to be significantly different things.

A strong sign of validity is good agreement between prediction and observation. Findings from the FUGA project show that EDA and facial EMG activity measured during an experiment predicted actual play in the three-week following period [manuscript under preparation]. Mandryk and others [41], modeled five emotions using an input of EDA, facial EMG, and cardiac measurements, to predict self-reports with tentatice success. Yannakakis and Hallam [80] successfully used a similar approach to create a model of children’s entertainment preferences, measuring cardiac indices and EDA.

Social game experience
Studying the experience of social interaction has provided some particularly consistent results, albeit concentrating on very precise questions. Several studies have reported that both arousal and positive valence are higher when playing against a friend, compared to playing against a stranger (EDA and facial EMG [59], and self-reports [22]), regardless whether they are in the same room or not [54]. A similar effect has been observed when playing against a human, compared to playing against computer ([59], and [42] with regrettably small sample), although it is not clear whether this is because human-vs-computer play lacks the social aspect or because the human-vs-computer game might simply involve easier challenges and/or be otherwise functionally different to a human-vs-human game.

Previous studies were often conducted with competitive games, but the difference between competitive and cooperative modes has also been investigated. One study [38] examined a simple singular cooperative or competitive event with other character presented as either player controlled, or computer controlled, but the results (higher arousal in EDA and HR for competitive and human-controlled conditions) may be explained by the two different game-play operations, trading (cooperative) vs. dueling (competitive), inducing different activity levels.

In one of our own experiments [32] participants played a simple action game, either cooperatively in teams of two against a team of computer-controlled characters, or competitively against each other with teams of one human and one computer-controlled character. Positive valence (facial EMG) tended to be higher in competitive mode for

---

1 FUGA - Fun of Gaming, see: http://project.hkkk.fi/fuga/
males, and in cooperative mode for females. Unfortunately, this experiment did not adequately control the opponent type, potentially a strong confound, as described above.

**Studying game features**

The research of virtual environments has also produced some studies which may be relevant to game experience. For example, it has been reported that people respond to virtual characters the same way they do to other people (facial EMG [78], pupil size, eye tracking [44]) and that this is (at least to some extent) dependent on how human-like the virtual character looks [10]. One interesting study [74] demonstrated that as the sole stimulus, darkness on a desktop virtual environment does not induce anxiety or stress (measured by HR and cortisol level), contrary to darkness in the real world and in immersive virtual environments. However, it is unclear to which extent games can be likened to virtual environments, which lack the intrinsic motivation provided by game mechanics (cf. magic circle [62]).

A few studies have focused on audio features. Findings include higher physiological arousal and better performance in an FPS with background music, as opposed to no music (using tonic cardiovascular measures [70]) and increased stress (cortisol level) in presence of music [25]. Similar effects (increased tonic HR as index of arousal) were found [79] only with interaction of game sounds and background color, but with a very simple game. A recent study did not find a difference in analysis of tonic physiological measures (facial EMGs, EDA) when testing effects of sounds and music in an FPS game [46]. However, the researchers question whether tonic measurements are in fact suitable for this kind of comparison. Due to large differences in games, experimental designs and analyses, the only conclusive outcome from these studies is an expected tendency of heightened arousal in response to game audio.

**Studying game events**

Although a major benefit of psychophysiological methods is the temporal precision and continuous measurement (see [41] for a good example), it has been exploited regrettably rarely. A powerful analysis technique is to extract only those parts of the data that correspond to a defined repetitive event, and statistically testing the trend of all reactions to it. In a pioneering work, Mandryk and Inkpen [42] presented responses to events, but almost in an anecdotal manner: five participants showed an expected rise of arousal (EDA) to scoring a goal and three to fighting in a hockey game. Our own results demonstrate that collection of points in a platform racing game was associated with increased arousal (HR, SCL, EEG) and positive affect (facial EMG), whereas reaching the goal was associated with decreased arousal and increased positive affect [56, 63]. Reasonably enough, getting points and winning seems to be experienced positively, and success and fighting is arousing.

More interestingly, and contrary to what could be expected, we found both in the racing game and in an FPS game that the death of the player’s own character seemed to cause a positive affect, whereas killing an opposing character in an FPS seemed to elicit a negative reaction [56, 57, 63, 58]. We have subsequently re-confirmed this effect in a separate study [31], using a different game, and comparing the events based on the opponent (friend, stranger, computer). We found that the death of the player’s character elicits a positive reaction regardless of the opponent, whereas the response to a kill is positive only when the opponent is human.

**Studying game effects**

A notable portion of gaming-related psychophysiological research has focused on the effects that games have on the players, and the medical or societal implications of such effects (see [3] and [18] for two reviews on effects of game violence using some physiological evidence, with opposite conclusions). Unfortunately the results of the game effect studies tend to be of limited use for game researchers. For example, it is unclear if the association with higher arousal and violent games (as compared to non-violent) tell much about the actual experience of game violence, as most studies compare reactions to different games or even different types of games, which makes it questionable whether other differences are adequately controlled (see e.g. [2, 20]; but cf. [68, 4]). This limitation is not unique to game effect studies, however: some otherwise plausible findings suffer from the same uncertainty (e.g., [64, 28]), showing that the presence of a story or screen size increases physiological arousal (tonic SCL) (but see [55]).

Perhaps the best attempt at understanding game effects has also provided an exceptionally interesting take at game experience and its research methodology: Weber and others [77] deconstructed the whole 50-min play session of thirteen participants, and analyzed player behavior in different phases (danger, safe, combat, etc.) and over time, as well as the most common event combinations and average heart rate during different events. More studies with this kind of systematic approach would significantly contribute to the basic understanding of the link between game structure and experience.

**Psychophysiology in game design**

Another important part of the psychophysiological game research literature is concerned with formative evaluation, for example using physiological methods to support design decisions [66]. A further application is the use of physiological signals for creating adaptive systems, such as dynamic difficulty adjustment (DDA), that provide a (supposedly) optimized game experience. For instance research [8, 51, 73] has demonstrated that it is possible to automatically assess emotional states in games from many physiological signals, and it was reported that affect-based DDA enhances gaming experience compared to traditional DDA (without physiology) [39]. Researchers have also
adaptively added effects and game content, with mixed results [14].

As adaptive technology advances, applications which are potentially viable for digital games in the near future have begun to emerge, for example, using EDA to sense player affect through the gamepad [69]. Some emotional reactions such as frustration can also be detected using no more than the pressure exerted by the player on the analog buttons of the gamepad ([23, 27], see also [60]). As the capacity for natural expression with game controllers increases, so do the possibilities for player modeling using only the controller as input. For instance, new controllers for all the major game consoles offer the potential for new methodologies utilizing their sophisticated motion-sensing technologies. Within the context of Uncommon User Interfaces, some researchers [49, 16, 26, 52] have utilized psychophysiological indicators as direct input, providing completely new interfaces. Recently, commercial applications have also started capitalizing on the potential of using signals such as EGG as novel input for games and toys (for example, Emotiv: http://www.emotiv.com/, Star wars science: Force trainer: http://unclemilton.com/starwarsscience/)

**FUTURE DIRECTIONS**

By continuously and synchronously measuring the physiological activity of many players it is possible to analyze concurrent physiological activations, referred to as physiological linkage [15]. In the case of game studies, such measures could allow better insights in the social experience of play.

Another potential future direction would be to analyze the reaction patterns of multiple psychophysiological signals together, and classifying typical reactions by interdependent reaction patterns between signals [41]. Individual intrasignal patterns could then further be linked to different game types and personality profiles, which would give us an explanation of how subjective gaming experiences are formed. In two similar approaches using non-psychophysiological source data [13, 9], the researchers presented separate methods for interpreting continuous data from the game engine. Such approaches offer another perspective on the rich source of data from psychophysiological recordings, with the possibility of integrating sources through a common framework.

**Conclusion**

A large number of studies have shown that psychophysiological measures can be used to index emotional, motivational, and cognitive responses to media messages (e.g. video, television, radio and textual messages), and similar evidence in the context of digital games is slowly growing. However, the emerging field is lacking useful and widely accepted game-specific theoretical background, systematic research, and accumulated results between studies. A first step towards developing the method is by validation of existing findings in a wider set of contexts: how do they hold for different types of games, game environments, modalities, social environments, etc.? As soon as there is a more substantial corpus of basic knowledge, the development of methodology guidelines for studying game experiences should be undertaken, including determining statistical requirements for experiment design, establishing intra-signal patterns of significance and better definition of the typical modes of interaction found in games. This would allow for tighter control of experiments and eventually more reliable and generalizable, ultimately more useful, results.

**ACKNOWLEDGMENTS**

This work was supported by the Finnish graduate school on User Centered Information Technology (UCIT).

**REFERENCES**


Symposium on Computational Intelligence and Games 2009, Milan, Italy, September 7-10.


responses to violent video game events. *Emotion, 8*, 1, 114—120.


