Chapter 2
User Experience Evaluation Methods for Games in Serious Contexts

Sotiris Kirginas

Abstract User experience in digital games can be influenced by many factors such as flow [Csikszentmihalyi (Flow: the psychology of optimal experience. Harper Collins, 1990), Sweetser and Wyeth (Computers in Entertainment 3(3):1–24, 2005)], immersion [Brown and Cairns (ACM Conference on Human Factors in Computing Systems, CHI 2004, ACM Press, 2004), Ermi and Mayra (Proceedings of Chancing Views – Worlds in Play. Digital Games Research Association's Second International Conference, 2005)], frustration or tension [Gilleade and Dix (Proceedings of the 2004 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology – ACE '04, 2004)], psychological absorption [Funk et al. (Proceedings of the Second International Conference on Entertainment Computing Pittsburgh, Carnegie Mellon University, 2003)], and social game context [Bracken et al. (Online video games and gamers' sensations of spatial, social, and copresence. FuturePlay 2005, 2005)]. Most of these factors should be present in a digital game in order to provide the optimal gaming experience [Kirginas (Contemporary Educational Technology 14(2):ep351, 2022), Kirginas et al. (International Journal of Child-Computer Interaction 28, 2021), Kirginas and Gouscos (The International Journal of Serious Games 4:53–69, 2017; International Journal of Serious Games 3:29–45, 2016)]. As there are many different game genres, sub-genres, and game types, user experience needs to be explored in more detail in research studies. This need is even greater when we talk about serious games. User experience is a multifactorial concept that is difficult to measure. This chapter aims to present a range of quantitative and qualitative/objective and subjective/short-term and long-term/formative and summative methods that can be used to evaluate users' experience in serious games during and after the development process. It is also intended to provide insight into when the different user experience assessment methodologies should be employed in the development cycle.

S. Kirginas (✉)
National and Kapodistrian University of Athens, Athens, Greece
e-mail: skirginas@media.uoa.gr

K. M. L. Cooper, A. Bucchiarone (eds.), *Software Engineering for Games in Serious Contexts*, https://doi.org/10.1007/978-3-031-33338-5_2

Keywords Digital games · Serious games · Assessment · Experience · Quantitative · Qualitative

2.1 Introduction

Nowadays, digital games compete with traditional activities like reading books, watching movies, listening to music, surfing the Internet, or playing sports [1]. Digital games regularly attract billions of players online and offline, generating huge revenue. However, digital games also present new research challenges for many traditional and new scientific areas [1]. With recent advances in the field of human-computer interaction [2], new methods are available to precisely measure how people interact with entertainment technologies [3, 4]. With new evaluation methods of player interaction, we aim to support the traditional digital serious games development and improve game design process [5].

Game developers increasingly employ user testing with playability evaluation in the development of digital games [6–9]. Unlike other software, digital games often offer a unique experience that contains elements that are difficult to be evaluated. User experience in digital games can be influenced by many factors, such as flow [10, 11], immersion [12, 13], frustration or tension [14], psychological absorption [15], and social game context [16].

A component of a game design process is observation of players in response to mechanics. Since it is very time-consuming to gain such individual knowledge of game design, it is necessary to gain a more rapid understanding of the complex behavior of players in response to game mechanics. To gain a more complete view of user experience, several recent solutions have combined event logging with objective and subjective player feedback [3, 17]. Similarly, player behavior is modeled to find "optimal spots in the game and level design" [4].

This chapter aims (a) to present a range of quantitative and qualitative/objective and subjective/short-term and long-term/formative and summative methods that can be used to evaluate users' experience in digital serious games during and after the development process and (b) to provide insight into when the different user experience evaluation methodologies should be employed in the development cycle.

The structure of this chapter is as follows: First, we outline methods for evaluating the user experience of digital serious games based on the body of the literature. In the next section, we explain how users' experiences are measured in digital serious games. Last but not least, we discuss when, how, and why to use all main methodologies proposed to measure the effectiveness of games in serious contexts.

2.2 Defining User Experience

According to Almeida et al. [18], experience is both the process and the outcome of a user's engagement with the environment at a given moment. It is both an interactive (the process of playing the game) and an emotional (the consequence of playing) experience—a feeling (or combination of emotions) that occurs as a result of playing [19]. The interaction process is the way players interact when playing; it is how the player interacts with other playable and non-playable characters and objects in the game environment [20] and how they make decisions. The game limits this process, which is influenced by the players' background, motivations, expectations, and current emotional experiences, which can change during the game [19]. Almeida [20] argues that in many cases, the emotional state of the players also influences the interaction processes: If they are anxious, they may be less attentive, which could affect their ability to play and win, while if they are relaxed, they could be in a flow state according to Csikszentmihalyi [10]. This is still a fairly open field in the game industry, as horror games, a prominent video game genre, is dedicated to keeping players in flow through anxiety or fear.

This approach has an impact on the outcome of the game. If the emotional experience is positive, games can trigger positive emotions (e.g., satisfaction, happiness, and excitement); if the emotional experience is unpleasant, games can trigger negative emotions (anger, sadness, boredom). Positive or negative effects can influence the interaction process by changing players' motivations and engagement [19, 21, 22]. This bidirectional interaction can explain why players can sometimes experience both pleasure and frustration during the course of a game [20].

According to Roto [23], there are three phases of the game experience: (a) the expected game experience (before a player interacts with a game), (b) the game experience during interaction (experience that occurs while interacting with the game), and (c) the overall player experience (experience that occurs while interacting with the game) (experience after the game ends). The player experience during interaction is the most important of the three phases of player experience mentioned above. Examining the player experience during interaction is critical to improving a game, as this phase can identify features and components that provide a positive experience as well as those that do not. According to Lallemand [24], three factors should be considered in order to understand the game experience during the interaction phase: the human aspect (dispositions, expectations, needs, motivation, mood, etc.), the system aspect (e.g., complexity, purpose, usability, functionality, etc.), and the contextual aspect (or environment) in which the interaction takes place (e.g., organizational/social environment, meaningfulness of the activity, voluntariness of use, etc.).

2.3 Methods to Evaluate UX in Serious Games

While game developers should construct games that are rewarding, entertaining, and appealing to consumers in order to enhance game reviews and sales, designing and developing digital games is a demanding and difficult process [25]. Therefore, it is important to understand how different players behave and interact with games. Understanding target players and their game experiences during game development is critical to create a better user experience and perhaps improve game ratings and financial success.

A survey by the Entertainment Software Association (ESA) found that digital games have become an important part of the games industry in recent decades. Due to a number of variables such as rapidly growing market, broader player demographics, and unique controller interfaces and platforms, digital games are an important area of research [25].

Consequently, the opportunity is broader; however, a deeper understanding of player demographics and platforms is required to address this market. According to Mirza-Babaei [25], stereotypes of the single player (e.g., the image of a teenager addicted to digital games) are generally disappearing in the industry in favor of a new image of multiple players playing simultaneously on multiple devices. In modern digital games, there are different types of interaction that offer more opportunities for player interaction.

Through the growing field of games user research (GUR), developers are evaluating their games for usability and user experience to improve the gaming experience. Games user research borrows user research techniques from human-computer interaction (HCI) and psychology, such as behavioral observation, interviews, questionnaires, and heuristic evaluation. Despite advances in applying user research methods to understand the usability of productivity applications, researchers and practitioners still face challenges in applying these methods to digital games. Digital games have unique characteristics that prevent the application of most conventional user research methods to the evaluation of the game experience [25].

As a result, user research methodological approaches have been modified and improved to better meet the goals of game development. These methods aim to provide players with a combination of qualitative and quantitative/objective and subjective/formative and summative/short-term and long-term methods to choose from depending on their research context and the needs of their participants. One of the main issues facing user experience and game usability evaluation is determining the optimal combination of different methods and combining the data from each method into a relevant report for game developers.

2.4 Analysis of Methodologies

Users' experiences in digital games can be measured and evaluated using different methods. These methods are classified in various ways in the following sections.

2.4.1 Quantitative vs Qualitative Assessment

Qualitative methods are used to explore and understand players' perceptions and interactions. Users' experiences are usually recorded in non-numerical data. In contrast, quantitative methods use numerical data [26]. Quantitative approaches show levels of engagement and interest by providing statistics, while qualitative approaches capture players' experiences during play. There are times when players lack emotional expression and do not speak freely when evaluating verbally or nonverbally. It is difficult for players to concentrate and talk about their experiences at the same time while playing a game. When evaluating a project, both methods should be used to achieve objective and comprehensive results.

In any research, researchers have to make a primary but basic methodological choice between the quantitative and the qualitative approach (or their combination) to investigate their topic. With the quantitative approach, they can find out "what happens," while with the qualitative approach, they investigate "why it happens." The aim of qualitative research is to "discover the views of the research population by focusing on the perspectives from which individuals experience and feel about events" [27]. In summary, qualitative assessment involves categorizing and evaluating qualitative data to help researchers analyze and interpret game events, user behavior, and player experiences. Collecting qualitative data can lead us down such paths, whereas collecting quantitative data cannot, especially when it comes to user experience.

2.4.2 Subjective vs Objective Assessment

Instruments for measuring players experience fall into two categories depending on their reliability: objective and subjective.

Objective assessment instruments provide accurate data that are objective and free from any subjective judgment of the participants because they are accurately recorded by machines [28]. Objective data are recorded automatically and continuously without disturbing the participants or interfering them.

In contrast, subjective instruments, are not precisely because they are completed by the users themselves, contain subjectivity, so they have lower reliability compared to objective instruments. An objective assessment tool measures the expressive or psychophysiological aspect of the user's experience using facial

expressions and collected psychophysiological data, while a subjective tool assesses the subjective feeling of the user's experience using self-reports, rating scales, and verbal protocols.

2.4.3 Short-Term vs Long-Term Assessment

In the early stages of game development, measuring users' initial and momentary experiences is important to obtain feedback [29]. It is also known that users' experiences change over time [30]. Therefore, it is necessary to use instruments that measure the experience over time to get more reliable information about game playability. In this way, a game developer can gain insight into how a player interacts with their game. Currently, user experience research mostly focuses on short-term evaluations. However, the relationship between a user and a game evolves over time, so long-term user evaluation is critical to a game's success.

These different categorizations are important because the reasons we want to measure user experience may vary from research to research. In some cases, we may want to measure qualitative attributes derived from the player's experience, while in other cases, we may want to measure quantitative attributes. Similarly, we may want to measure the player experience at a particular point in the game, such as when the player wins a significant player, or we may want to assess it over a longer period of time.

2.4.4 Formative vs Summative Evaluation

There are two types of evaluation in user experience, formative and summative. Which type of evaluation we should use depends on where we are in the process of developing a digital game.

Formative evaluations focus on identifying aspects of the design that work well or not well and why. These evaluations are conducted during the redesign of a game and provide information to gradually improve the game. Considering the case of designing a new digital game for mobile phones, as part of the design process, a prototype is created for this game and tested with (usually a few) users to see how easy it is to use and how players experience it. The research may reveal several weaknesses in the prototype, which are then addressed with a new design. This research is an example of a formative evaluation—it helps the designers determine what needs to be changed to improve the game. Formative evaluations involve testing and modifying the game, usually many times, and are therefore appropriate when developing a new game or redesigning an existing game. In both cases, the prototyping and testing steps are repeated until the game is ready for mass production [31].

Summative evaluation describes how well a game performs, often compared to a benchmark, such as a previous version of the game or a competitive game. Unlike formative evaluations, whose goals are to inform the design process, summative evaluations involve getting the big picture and evaluating the overall experience of a completed game. Summative evaluations are done less frequently than formative evaluations, usually immediately before or immediately after a redesign. Assume the redesign of the mobile phone game is complete, and now it is time to evaluate how well it performs compared to the previous version of the game. After the data from the survey is collected, it is then compared to the data obtained from the previous version of the game to see if there has been any improvement. This type of survey is a summative evaluation as it evaluates the product shipped with the goal of tracking performance over time and ultimately calculating our return on investment. However, during this study, we may uncover some usability issues. These issues should be noted and addressed during the next game design. Alternatively, another type of summative evaluations could compare results to those obtained from one or more competitive games or to known data across the gaming industry. All summary ratings give an overview of a game's usability. They are meant to serve as benchmarks so we can determine whether our own games have been improved over time. The final summative evaluation is the go/no-go decision on whether to release a product [31].

2.5 Overview of the Main Methodologies

There are a variety of tools and methods to uncover the quality of the experience generated by a game, either to improve it or to use the game for the purposes of education, training, awareness raising, and behavior change of subjects. Table 2.1 summarizes all UX assessment methods together with their assignment to one or more of the categories mentioned above.

2.5.1 Think-Aloud Protocol

The think-aloud protocol is a qualitative method of collecting data in which players describe their playing experiences to an expert facilitator. The facilitator pays attention to both verbal and nonverbal (e.g., behaviors, body language) players' responses to gain insights into the player experience [32]. Think-aloud protocol asks participants to spontaneously report any thoughts they have while they interact with a game without interpreting or analyzing what they have thought about [33]. The think-aloud protocol consists of two components: (a) the technique for collecting verbal data (think-aloud interview) and (b) the technique for predicting and analyzing verbal data (protocol analysis). The method is useful for researchers interested in observing, exploring, and understanding the thoughts and opinions of

Table 2.1 Overview of the main instruments and methods

Methods/instruments	Quantitative	Qualitative	Subjective	Objective	Short-term	Long-term	Formative	Summative
Think-aloud protocol		✓	✓		✓	✓	✓	✓
Surveys		✓	✓		✓	✓	✓	✓
Expert evaluation		✓	✓		✓	✓	✓	✓
Playtesting		✓	✓		✓	✓	✓	✓
Observation		✓	✓		✓	✓	✓	✓
Interviews		✓	✓		✓	✓	✓	✓
Focus groups		✓	✓		✓	✓	✓	✓
Psychophysiological measurements	✓	✓		✓	✓	✓	✓	✓
Self-assessment measurements	✓		✓		✓	✓	✓	✓

players, which can be a challenging endeavor [34, 35]. Depending on the interaction with the game, it can generate reports during the interaction or afterward [36].

In order to implement the think-aloud protocol, the following steps are taken: (a) users are assigned tasks, (b) users are asked to speak aloud their thoughts during the performance of the tasks, (c) users' thoughts are recorded as they are performing the tasks, and (d) the material is analyzed and commented on by the researcher(s). Based on Avouris [37], the think-aloud protocol can be divided into the following variations: (a) critical response protocols, in which the user is required to speak aloud only during a predetermined task, and (b) periodic report protocols, in which the user explains his/her thoughts only after completing a particular task so that the task is not disturbed.

The advantage of think-aloud protocol is that researchers are able to identify players' main misconceptions, since it allows them to understand how players view a game. Think aloud also enables them to obtain a rapid and high-quality response from a small number of participants [38]. A number of researchers have criticized the method for disrupting user concentration [39] and claiming that self-observation would interfere with thought process and, as a result, wouldn't show real thought processes [33].

2.5.2 Expert Evaluation

Expert evaluation refers to an overview of the game conducted by an expert or a team of experts. It is a formative or summative evaluation conducted by designers and user experience experts to identify potential problems and improve the design [40]. Expert evaluation can be conducted for an existing game to identify problems that can be fixed by redesigning the game. Expert evaluation of games under development can identify new problems before a prototype is created. Klas [41] describes two types of expert evaluation: In the first, the experts themselves act as evaluators, conduct the evaluation, and report on the results. In the second, the evaluators supervise the experts, lead the evaluation, and assess their performance. In comparison, expert evaluations provide quick and cost-effective results, in contrast to more expensive types of qualitative user studies, such as playtesting, which require more evaluators for a representative result [41].

In addition, expert evaluation can be used at different stages of the development process to identify usability issues early in the process [40]. Expert evaluation can be made more efficient through the use of heuristic analysis. A heuristic is a set of guidelines that help ensure design is consistent with best practices within an industry, and it is often used by researchers to support their evaluations [42]. The evaluators then come together to produce the results report.

Typical findings include:

(a) Which features of the game may cause usability problems and need to be improved

(b) Which features are likely to be successful and should be retained
(c) Which features should be tested with real players

2.5.3 Cognitive Walk-Through (CW)

Cognitive walk-through (CW) is a user interface design method that allows designers to model how a particular type of user will understand a user interface through exploration [43, 44] and to evaluate the learnability of a digital serious game [45]. It is an expert-based evaluation method that is therefore relatively inexpensive to implement and can be used to identify usability issues in a system effortlessly, quickly and economically [46]. As in expert evaluation, a team of reviewers walks through a task and evaluates the interface from a new user's perspective. As our main interest is in serious games, we propose that the cognitive walk-through is an appropriate and effective method to evaluate the learning potential of serious games, as both the design and evaluation practices of serious games can benefit from the cognitive walk-through method.

A cognitive walk-through cannot be conducted until the design of the game, the task scenario, the user assumptions, the scope of the game, and the sequence of actions that players must perform to successfully complete a given task are accurately described [43, 46]. Then, an evaluator or group of evaluators (2–6 expert evaluators) simulates a series of cognitive processes that users go through when completing a set of tasks. By understanding the behavior of the interface and its influence on players, evaluators are able to choose actions that are difficult for ordinary players. It would therefore be useful to use this evaluation method in the early stages of system development to ensure that users' needs are met.

2.5.4 Playtesting

The term playtesting refers to the use of traditional user testing methods for games [47]. The game design literature argues that playtesting is the most popular and most important method for game developers to evaluate their game designs. It is important for game developers to use playtesters to give feedback on unintended challenges in their games, to collect data on the way players prioritize tasks and goals, and to understand how players understand the mechanics of the game [31]. During playtesting, testers who have characteristics similar to those of the expected end users (e.g., age, education level, professional similarities, gaming experience) test the first and subsequent versions of a game and provide feedback to the game developers, which is then incorporated into the game design [48].

Playtesting can be formal (or open), informal (or closed), or beta. Formal playtesting can be conducted with non-design group members according to Korhonen [49]. Participants are usually required to fill out a questionnaire or provide

contact information in order to be considered for participation. Several members of the design group can conduct informal playtesting. Finally, beta playtesting relates to the final phases of testing, before releasing a product to the public, and is sometimes conducted semi-formally with a limited version of the game to identify any last-minute issues.

2.5.5 Interviews

Interviews are an essential element of a qualitative evaluation session with users [50]. They provide one of the few ways of validating observations, discovering issues, gathering opinions, and determining the sources of challenges encountered by players [50]. Interviews can be used with other methodologies to enhance the gathered data and give a holistic perspective of the user's attitudes and emotions, and they are an essential element in identifying and understanding usability issues and obstacles in the player's experience [51]. Therefore, interviews seem to be the right choice for specific study aims and knowledge [52]. Nacke et al., for example, suggest using interviews to measure the PX and capture the context and social influences on the individual player's experience with serious games [5].

2.5.6 Focus Groups

Focus groups are a form of qualitative and subjective research. In a focus group, a group of people gather in a room to discuss a topic under guidance. It is a semi-structured interview process in which a small group of people, usually six or eight, discuss a specific study topic [19]. Krueger and Casey [53] describe the focus group method as a means of obtaining perceptions about a particular area of interest in a permissive, non-threatening environment (p. 5). To obtain qualitative data about the research topic, the moderator steers the discussion more or less according to its structure. Take a research project on user experience with a digital game. A more in-depth interview with the players might be necessary, but before we do that, we want to see what kinds of questions work and whether the players might raise issues we are not considering so that we can include them in our questions.

In a focus group, participants are selected based on their relevance and relationship to the topic. Therefore, they are not considered statistically representative of a significant population because they are not selected using strict probability sampling methods. Instead, participants are selected through random sampling, advertising, or snowballing, depending on the type of person and the characteristics the researcher wants to consider. There are several advantages of focus groups: It is a socially oriented research method that collects real-life data in a social setting, is flexible, has high validity, provides rapid results, and costs nothing to conduct. There are also some disadvantages of focus groups: the researchers have less control

than with individual interviews, the data can sometimes be difficult to analyze, the moderators need certain skills, and the discussion needs to take place in a conducive environment.

2.5.7 Observation

Observation is a deeply qualitative research methodology that can be integrated into a variety of qualitative and quantitative research projects. Researchers can gain a great deal of data and information from their observations by watching users engage in a particular activity and then analyzing it. When observation is combined with other methods and techniques, it is possible to gather valuable data to interpret the topic the researcher is exploring.

The researcher must have specific skills, and the observation procedure involves some methodological risks, especially in terms of its validity and reliability, as the question of objectivity and impartiality is always present. Therefore, it is usually better for inexperienced researchers to combine this technique with another one, such as an interview, in order to collect all the data needed, to shed light on certain aspects of the study or to triangulate the information.

2.5.8 Surveys

Surveys may be used in research to examine player-game interactions and, depending on the results, improve the gaming experience [54]. The goal of surveys is to collect data on a subset of the population being studied by the researcher [55]. The survey results can then be extrapolated to the full population. Surveys are a quick, simple, and low-cost technique to collect a big amount of data that tells more about the subjective experience of playing a game [54, 56]. This may give the impression that creating a survey is simple, yet seemingly slight oversights can dramatically restrict the utility of your survey data.

Surveys can help researchers collect objective and subjective data. Objective data are directly observable and can be verified by others, such as demographic characteristics and the number of hours spent playing games. In contrast, subjective data are not objectively verifiable, such as attitudes and emotions. Overall, surveys can be used to assess player attitudes and experiences, motives, player characteristics, differences between groups of players, or different iterations of a design [54]. Their advantages include ease of use, use in many situations, minimal cost, access to large population, absence of interviewer bias, and fast transmission/response times [54, 56]. Surveys become even more effective when combined with other methods [54]. For example, while game analytics may indicate that players are more likely to succeed in a game, survey data may show that players were less challenged and bored [57]. In addition, survey data can be combined

with physiological measures, such as facial recognition and electrodermal activity measurements [21]. Researchers can create their own questionnaires to measure outcomes or use existing, validated questionnaires to compare the results of their own studies with those of other studies.

Below are some of the most commonly used questionnaires:

2.5.8.1 The Player Experience of Need Satisfaction

According to Rigby and Ryan [58], people have three universal needs: competence (perception of a challenge), autonomy (voluntary aspects of an activity), and relatedness (connection to others). These are the main components of what we call Player Experience of Need Satisfaction (PENS) method. The PENS evaluation includes two additional factors, presence (the experience of being in the game world) and intuitive control, both of which are considered key features of games [59]. Using 7-point Likert scales, the PENS assesses these needs as well as the additional factors. When games meet these motivational criteria, the game experience and game success improve significantly. The PENS method is methodologically easy to apply as it successfully targets specific experiences related to need satisfaction and provides practically rapid feedback. These measurements can be easily applied to specific design or game concepts as well as to games that already have established.

2.5.8.2 Challenge Originating from Recent Gameplay Interaction Scale

The challenge originating from recent gameplay interaction scale (CORGIS) is a psychometric instrument developed by Denisova et al. [60]. This instrument is used to assess perceived challenge in digital games. The questionnaire assesses four types of perceived challenge in games:

Cognitive challenge: it stems from the need to plan ahead, memorize, exert effort, prepare, and multitask.

Performative challenge: it arises from the fact that the game requires the player to act quickly and accurately.

Emotional challenge: it arises from the emotions evoked in the player, which can also affect the things he thinks about outside the game.

Decision-making challenge: it arises from having to make decisions that are difficult or can lead to unfortunate outcomes.

2.5.9 Immersive Experience Questionnaire

Jennett et al. [61] developed the Immersion Experience Questionnaire (IEQ) to measure the level of immersion of players. It measures the user experience using a 5-

point Likert scale but focuses primarily on the concept of immersion. The IEQ uses positively and negatively worded questions. For every positively worded question, there is a negatively worded question, which adds accuracy to the questionnaire. The total score is the sum of the scores of the positively and negatively worded questions. When the IEQ was developed, it was assumed that immersion was based on five components. In practice, however, immersion is considered as a single dimension, with the components influencing the interpretation of the results.

2.5.9.1 Sensual Evaluation Instrument

The sensual evaluation instrument (SEI) was developed by Isbister et al. [62]. This is a nonverbal, body-based tool that can be used to capture shared responses more directly, saving designers time and energy and in turn increasing the likelihood that users will engage early in the design process. The SEI consists of eight sculptural objects that represent the range of emotions one would expect to experience when interacting with a digital game. The objects are not one-to-one with specific emotions. Rather, they are meant to serve as a starting point so that everyone can develop their own expressive taxonomy of the objects. People share their feelings as they engage in the experience. They arrange the objects as they wish or show in some way that they feel comfortable with the object or objects that correspond to their current feelings. In the end, the researcher watching the video in conjunction with SEI can better understand how the player felt during the game [63].

2.5.9.2 Game Experience Questionnaire

It is a tool designed specifically for young children (8–12 years old) to assess their gaming experiences. The game experience questionnaire (GEQ) [64] assesses seven different dimensions of gaming experience (immersion, flow, effectiveness, intensity, challenge, positive emotion, negative emotion) Each of the seven dimensions is distinguished into five sub-themes rated on a 5-point Likert scale. The game experience questionnaire is divided into three separate modules, each of which deals with a different experience: (1) core module, which evaluates the user's experience while playing the game; (2) social presence module, which evaluates the user experience while playing a game with others; and (3) post-game module, which evaluates the user's experience after completing the game. It has the advantage of measuring different aspects of the game experience (immersion, flow, effectiveness, intensity, challenge, positive emotions, and negative emotions), assessing the experience during and after the game, and assessing social presence as well. As it covers such a large area, it can be difficult to complete by all the researchers, so many researchers only use some of the modules.

2.5.10 Psychophysiological Measurements

Quantitative and qualitative researches both use psychophysiological measurements to assess users' experiences. As users' experiences during gameplay can have a significant impact on the playability of digital games, physiological data can be very useful to assess players' emotional state and performance, especially when correlated with subjective measurements [21]. So far, results have only been reported for first-person shooters games [65, 66]. The question arises whether physiological and subjective measurements might prove equally reliable for other types of digital games. The main methods for assessing user experience using physiological methods are as follows:

Electrodermal activity (EDA): perhaps the most commonly used physiological measurement. It is often referred to in the literature as galvanic skin response or skin conductance. Sweat gland secretions during play are indicators of positive arousal and mental activity [67, 68].

Cardiovascular activity measurement: an important physiological measure of human activity. Cardiovascular activity measures heart rate and heart rate variability [69, 70].

Electromyography (EMG): provides measurements of the electrical muscles. When a person is excessively anxious, skeletal movements are observed as a sign of involuntary muscle contractions during intense mental activity, intense emotions, and cognitive stress [46, 71, 72].

Facial expression: analyses human facial expressions during activity and measures basic human emotional states such as happiness, sadness, anger, surprise, disgust, etc. [73].

Electroencephalography (EEG) is performed with special electrodes that are attached to the participant's head during the test. Brain activity is then measured using frequency wave patterns that represent different mental activities [74, 75]. Since electrodes are used in electroencephalography, it is purely a laboratory measurement.

2.5.10.1 Biofeedback Measuring Device

The biofeedback measuring device is a device designed and built in the Laboratory of New technologies of the Department of Communication and Media Studies, University of Athens. This device consists of a sensor part housed on a typical computer mouse, an analogue electronic circuit that transmits the processed signal to a typical home computer, and finally a software component that converts the measurements into a suitable format. The STC is seamlessly detected by the contact of the thumb and ring finger with the Al-Si ring sensors, located on the left and right sides of the computer mouse, respectively (Fig. 2.1).

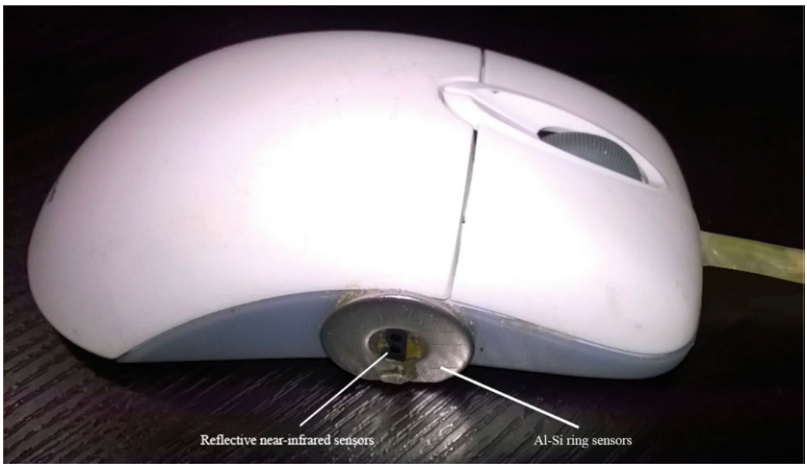

Fig. 2.1 The biofeedback measuring device

Heart rate is also detected by reflective near-infrared sensors in the center of the ring sensors (Fig. 2.1), based on the principle of reflective absorption that occurs during changes in skin coloration caused by the pulsation of blood in the tissue.

2.5.10.2 FaceReader

A software application called FaceReader was developed by Noldus Information Technology. The FaceReader software uses algorithms to rate facial images according to seven basic emotional states—happy, sad, angry, surprised, scared, disgusted, and "neutral emotional state." These seven emotions are rated from 0 (not at all) to 100 (perfect match). FaceReader "is an effective tool for measuring emotional experience during human-computer interaction, as it strongly suggests that more effective and well-designed systems elicit more positive emotions and fewer arousing falls than less effective applications" [21].

2.5.10.3 Self-Assessment Methods

Self-assessment methods are subjective, most often quantitative, and either short or long term. They provide players with the ability to self-evaluate or make judgments about their experience and the games they play based on specific self-assessment tools. Their great advantages are ease of use and the use in many situations. However, their disadvantage lies in the subjectivity of the judgments, which can be affected by a number of factors, including bias, differences in age and gender, economic and social status, and past experiences, among others.

2.5.10.4 Fun Toolkit

The Fun Toolkit was developed by Read and MacFarlane [76]. It consists of three separate questionnaires:

(a) Smileyometer: It is a measurement scale based on a 5-point Likert scale, with ratings from 1 "Poor" to 5 "Excellent." The Smileyometer can be used both before and after the child's experience with a digital application, be it an educational software or a website or a digital game. By using it before engaging with the application, we can gather information about the children's expectations from the game. Using it latter, we can collect information about the fun of the game or the emotional experience of the players.
(b) Fun sorter table: A fun sorter table generally compares a set of products, whether they are educational software or digital games, as in our case. For a survey on children's ratings of digital games, children compare and rank them from best to worst or from easiest to hardest or from what they intend to play again to what they intend to play less.
(c) Again and again table: The questionnaire consists of a table in which children mark whether they experienced each activity with a "Yes," "Maybe," or "No." The idea for this tool comes from the field of psychology where it is argued that we are more likely to return to an activity we liked again and again if we like it. In the present study, children were asked, "Would you like to play with the toy again?", and they had to answer accordingly.

2.5.10.5 Self-Assessment Manikin (SAM)

The Self-Assessment Manikin (SAM) is a system for evaluating three dimensions of gaming experience: valence, arousal, and dominance [77]. It uses three pictorial scales, illustrating cartoon creatures. All three scales are 9-point and take values from 1 to 9, with 5 representing the middle of the scale. Although it is stated that it is a weighted method, there are insufficient studies that support this claim. Its advantages include ease of completion and its ability to be used in different circumstances. The disadvantages are what all objective assessment tools suffer from: objectivity of judgment and difficulty in matching experience with graphic.

2.5.10.6 UX Curve

The UX Curve is a tool for retrospectively evaluating user experiences. There is a timeline and a horizontal area in which the user can graph his positive and negative experiences. The advantage of UX Curve is that it allows the user to design the most immersive game experience. Nevertheless, its disadvantage is that it relies on retrospective memory from the game rather than reality for its completion [78].

2.5.10.7 MemoLine

The MemoLine is actually a timeline that can be used for retrospective evaluations. There are as many frames as there are time periods in which the user plays a game. As the tool is intended for children, the experiences they have are represented by three different colors: green represents positive experiences, red represents negative experiences, and gray represents times when the game is not played, e.g., weekends. Users are given questionnaires for each of these game scenarios: usability, challenge, quantity, and general impression [79].

The above questionnaires are certainly not the only ones. There are a large number of other relative questionnaires such as Emo-watch, EGameFlow, Gameful Experience Questionnaire, Model for the Evaluation of Educational Games (MEEGA+), Game User Experience Satisfaction Scale (GUESS), iScale CORPUS (Change Oriented analysis of the Relationship between Product and USer), and many others.

2.6 Discussion

The aim of this chapter is to provide an overview of evaluation methods to game developers and researchers whose research interests are related to digital serious games. This process is extremely important, considering that serious games differ from games whose goal is to entertain players, rather than teach or train them. It is also very important not only to describe these methods but also to highlight the advantages and disadvantages of each method, as well as to explain when, how, and why it makes sense to use each of these evaluation instruments. As it has been discussed in this chapter, the tools for evaluating player experience can be divided into four groups: objective-subjective, quantitative-qualitative, formative-summative, and short term-long term.

Beginning with the objective and subjective instruments for evaluating players' experiences, things are plain. Objective evaluation instruments provide objective data, free from any subjective judgment. Data are accurately recorded by machines and software, without disturbing or interfering participants. In contrast, subjective instruments are not accurate, as they are completed by the users themselves and therefore have lower reliability than objective instruments. Each of these evaluation methods has its own advantages and disadvantages. On the one hand, objective evaluation provides reliable results but is difficult to be applied as it requires expensive equipment and is a purely laboratory procedure. In contrast, subjective evaluation is easier to be applied, since it only requires finding suitable subjects, whether they are players or experts, but the data collected is less reliable due to the subjectivity of the participants. An evaluation system that uses both objective (e.g., a skin conductance measurement) and subjective methods (e.g., a self-reported questionnaire) to evaluate players' experiences is proposed to overcome the disadvantages and benefit from both forms of evaluation. Therefore, researchers

are able to collect data that is free of users' biases, while at the same time they can interpret it based on users' perceptions and opinions.

A lot of information can be gained from both formative and summative evaluation, which can be used by developers to improve their games. In formative evaluation, developers and experts or players have a dialogue about game play, which helps gather information for game design. An evaluation of this kind identifies what aspects of the design work well and what aspects don't. As a game is being redesigned, these evaluations provide information that can be used to improve the game gradually. As opposed to formative evaluation, summative evaluation discusses how well a game performs, usually in comparison with a benchmark, such as a previous version or a competitive game. A summative evaluation takes a step back from formative evaluations, which aim to inform the design process, and instead looks at the big picture and evaluates the overall experience. In most cases, summative evaluations are conducted just before or just after a redesign, and they are less frequent than formative evaluations. A developer can thus use formative or summative evaluation based on what they want to measure and the stage at which it is being developed. Formative and summative evaluations can be implemented with most of the tools described and suggested in this chapter, and the development team can decide which types to use. As a general rule, formative evaluations produce qualitative data, and summative evaluations produce quantitative data. To conduct a formative evaluation, developers should rely on instruments such as a think-aloud protocol, cognitive walk-throughs, observation, focus groups, interviews, etc. To conduct a summative evaluation, developers should rely on instruments such as psychophysiological and self-assessment measurements.

Serious game evaluation is essential for any developer, as it is an important function at every stage of game development. A comprehensive evaluation of players' experience is beneficial to a developer in many ways. It is a well-known method to assess the strengths and weaknesses of the game experience, which further serves as a basis for working and improving the overall game experience. Usually, the evaluation is done at the end when the game is ready for use. However, some developers also evaluate player experiences in the short term (during development). This has its own advantages, because if the game development is not going in the desired direction, the developer can correct it, instead of waiting for the end of the development and then making corrections.

Lastly, qualitative methods provide statistics about player engagement and interest, while quantitative approaches help developers and researchers study players' perceptions and interactions. The researcher must choose a methodological approach (or a combination of both) when researching any topic (either quantitative or qualitative). In a quantitative approach, developers discover "what happens," while in a qualitative approach, they discover "why it happens." In summary, qualitative assessment involves categorizing and evaluating qualitative data to help us analyze and interpret game events, user behavior, and player experiences. Collecting qualitative data can lead us down such paths, whereas collecting quantitative data cannot, especially when it comes to user experience.

2.7 Conclusions

This chapter is intended to serve as a guide for serious game developers and researchers who wish to evaluate existing games, to improve the players' experience and reach an optimal level. A player experience evaluation should record and interpret players' experiences of interacting with a digital game, and it is important that these records are accurate and reliable in order to produce meaningful and useful results. It is also important that an evaluation can identify the situations and factors that impact the player experience and make it more or less positive. In this case, we can make the necessary adjustments and changes to improve the player experience. According to what has been discussed in this chapter, the tools for evaluating the player experience can be divided into four groups: objective-subjective, quantitative-qualitative, formative-summative, and short term-long term.

Since the game experience is multidimensional and difficult to measure, it is important to use methods with different characteristics. Measurement and evaluation of player experience should be done using instruments derived from different methods, e.g., quantitative instruments and qualitative evaluation instruments or objective instruments and qualitative evaluation instruments. It is possible to negatively impact our evaluation efforts if we only use instruments from a single methodology.

Last but not least, the methodology we use to evaluate the user experience is crucial for understanding and interpreting the experience of playing a digital serious game. Future research should evaluate digital games using different evaluation methods and instruments. These studies should ultimately aim to find the most effective combination of tools and methods to measure the potential of a game.

References

1. Nacke, L.E.: Affective ludology: scientific measurement of user experience, interactive entertainment. Blekinge Institute of Technology (2009)
2. Mandryk, R.L., Atkins, M.S., Inkpen, K.M.A.: Continuous and objective evaluation of emotional experience with interactive play environments. In: Proceedings of CHI 2006, Montréal, Québec, Canada, April 2006, pp. 1027–1036. ACM (2006)
3. Drachen, A., Canossa, A.: Towards gameplay analysis via gameplay metrics. In: Proceedings of MindTrek, Tampere, Finland, October 1–2. ACM (2009)
4. Kim, J.H., Gunn, D.V., Schuh, E., Phillips, B., Pagulayan, R.J., Wixon, D.: Tracking real-time user experience (TRUE): a comprehensive instrumentation solution for complex systems. In: Proceedings of CHI 2008, pp. 443–452. ACM, Florence, Italy (2008)
5. Nacke, L., Niesenhaus, J., Engl, S., Canossa, A., Kuikkaniemi, K., Immich, T.: Bringing digital games to user research and user experience. In: Proceedings of the Entertainment Interfaces Track 2010 at Interaktive Kulturen 2010 ceur workshop proceedings, 12–15 September (2010)
6. Pagulayan, R., Keeker, K., Wixon, D., Romero, R.L., Fuller, T.: User-centered design in games. In: The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications. L, pp. 883–906. Erlbaum Associates, New York, NY (2003)

7. Pagulayan, R., Steury, K.R., Fulton, B., Romero, R.L.: Designing for fun: user-testing case studies. In: Funology: From Usability to Enjoyment, pp. 137–150. Kluwer Academic Publishers, Norwell, MA (2004)
8. Desurvire, H., Caplan, M., Toth, J.A.: Using heuristics to evaluate the playability of games. In: CHI '04 Extended Abstracts, pp. 1509–1512. ACM, Vienna (2004)
9. Korhonen, H., Koivisto, E.M.I.: Playability heuristics for mobile games. In: Proceedings of Conference on HCI with mobile devices and services, pp. 9–16. ACM, Espoo, Finland (2006)
10. Csikszentmihalyi, M.: Flow: the psychology of optimal experience, vol. 39, 1st edn. Harper Collins, New York (1990)
11. Sweetser, P., Wyeth, P.: GameFlow: a model for evaluating player enjoyment in games. Comput. Entertain. **3**(3), 1–24 (2005)
12. Brown, E., Cairns, P.: A grounded investigation of immersion in games. In: ACM Conference on Human Factors in Computing Systems, CHI 2004, pp. 1297–1300. ACM Press (2004)
13. Ermi, L., Mayra, F.: Fundamental components of the gameplay experience: analysing immersion. In: de Castell, S., Jenson, J. (eds.) Proceedings of Chancing Views – Worlds in Play. Digital Games Research Association's Second International Conference, Vancouver (2005)
14. Gilleade, K.M., Dix, A.: Using frustration in the design of adaptive videogames. In: Proceedings of the 2004 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology – ACE '04, pp. 228–232. (2004)
15. Funk, J.B., Pasold, T., Baumgardner, J.: How children experience playing video games. In: Proceedings of the Second International Conference on Entertainment Computing Pittsburgh, Carnegie Mellon University, pp. 1–14 (2003)
16. Bracken, C., Lange, R.L., Denny, J.: Online video games and gamers' sensations of spatial, social, and copresence. FuturePlay 2005. East Lansing (2005)
17. Nacke, L., Lindley, C., Stellmach, S.: Log who's playing: psychophysiological game analysis made easy through event logging. In: Proceedings of Fun and Games, 2nd International Conference, Eindhoven, The Netherlands, October 20–21, pp. 150–157. Springer (2008)
18. Almeida, S.: The player and video game interplay in the gameplay experience construct. PhD, Universidade de Aveiro (2013)
19. Kirginas, S., Gouscos, D.: Development and validation of a questionnaire to measure perceptions of freedom of choice in digital games. Int. J. Serious Games (IJSG). **3**(2), 29–45 (2016)
20. Almeida, S., Veloso, A., Roque, L., Mealha, O., Moura, A.: The video game and player in a gameplay experience model proposal. In: Proceedings of Videojogos 2013 – 6th Annual Conference in the Science and Art of Video Games. University of Coimbra, Coimbra, Portugal (2013)
21. Kirginas, S., Psaltis, A., Gouscos, D., Mourlas, C.: Studying children's experience during free-form and formally structured gameplay. Int. J. Child-Comput. Interact. **28** (2021)
22. Kirginas, S., Gouscos, D.: Exploring the impact of free-form and structured digital games on the player experience of kindergarten and primary school students. In: Russell, D., Laffey, J. (eds.) Handbook of Research on Gaming Trends in P-12 Education, pp. 394–420. Hershey, PA, Information Science Reference (2016)
23. Roto, V.: User Experience from Product Creation Perspective. Towards a UX Manifesto, pp. 31–34 (2007)
24. Lallemand, C.: Towards consolidated methods for the design and evaluation of user experience. Doctoral dissertation, University of Luxembourg (2015)
25. Mirza-Babaei, P.: Getting ahead of the game: challenges and methods in games user research. User Exp. Magaz. **15**(2) (2015) https://uxpamagazine.org/getting-ahead-of-the-game/
26. Neill, J.: Qualitative & Quantitative Research. http://wilderdom.com/research/QualitativeVersusQuantitativeResearch.html (2009)
27. Bird, M., Hammersley, M., Gomm, R., Woods, P.: Educational Research in Action/ Fragkou E: Translate in Greek. Educational Research in Practice-Study Manual, Patras (1999)
28. Cacioppo, J., Tassinary, L., Berntson, G.: Handbook of Psychophysiology, 3rd edn. Cambridge University Press, New York (2007)

29. Vermeeren, A., Lai-Chong Law, E., Roto, V., Obrist, M., Hoonhaut, J., Väänänen-Vainio-Mattila, K.: User experience evaluation methods: current state and development needs. In: Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries: NordiCHI '10 (2010)
30. Fenko, A., Schifferstein, H.N.J., Hekkert, P.: Shifts in sensory dominance between various stages of user-product interactions. Appl. Ergon. **41**, 34–40 (2010)
31. Joyce, A.: Formative vs. Summative Evaluations [blog]. https://www.nngroup.com/articles/formative-vs-summative-evaluations/ (2019, July 28)
32. Pellicone, A., Weintrop, D., Ketelhut, D.J., Shokeen, E., Cukier, M., Plane, J.D., Rahimian, F.: Playing aloud: leveraging game commentary culture for playtesting. Int. J. Gaming Comput.-Mediat. Simulat. (IJGCMS). **14**(1), 1–16 (2022)
33. Guss, C.D.: What is going through your mind? Thinking Aloud as a method in cross-cultural psychology. Front. Psychol. **9**, 1292 (2018)
34. Johnson, W.R., Artino, A.R., Jr., Durning, S.J.: Using the think aloud protocol in health professions education: an interview method for exploring thought processes: AMEE Guide No. 151. Medical teacher, 1–12. Advance online publication (2022)
35. Lundgren-Laine, H., Salantera, S.: Think-aloud technique and protocol analysis in clinical decision-making research. Qual. Health Res. **20**(4), 565–575 (2010)
36. Zhang, X., Simeone, A.L.: Using the think aloud protocol in an immersive virtual reality evaluation of a virtual twin. In: Proceedings of the 2022 ACM Symposium on Spatial User Interaction (SUI '22), pp. 1–8. Association for Computing Machinery, New York, NY (2022)
37. Avouris, N., Katsanos, C., Tselios, N., Moustakas, K.: Introduction to Human-Computer Interaction [Undergraduate textbook]. Chapter 11. Kallipos, Open Academic Editions (2015)
38. Jordan, P.W.: An introduction to usability. Taylor & Francis, London (1998)
39. Hashemi Farzaneh, H., Neuner, L.: Usability evaluation of software tools for engineering design. In: Proceedings of the 22nd International Conference on Engineering Design (ICED19), Delft, The Netherlands, 5–8 August (2019)
40. Nova, A., Sansalone, S., Robinson, R., Mirza-Babaei, P.: Charting the uncharted with GUR: how AI playtesting can supplement expert evaluation. In: Proceedings of the 17th International Conference on the Foundations of Digital Games (FDG '22). Association for Computing Machinery, New York, NY, Article 28, pp. 1–12 (2022)
41. Klas, C.: Expert evaluation methods. In: Dobreva, M., O'Dwyer, A., Feliciati, P. (eds.) User Studies for Digital Library Development, pp. 75–84. Facet (2012)
42. Rajanen, M., Rajanen, D.: Heuristic evaluation in game and gamification development. In: Proceedings of GamiFin 2018 Conference, Pori (2018)
43. Allendoerfer, K., Aluker, S., Panjwani, G., Proctor, J., Sturtz, D., Vukovic, M., Chen, C.: Adapting the cognitive walkthrough method to assess the usability of a knowledge domain visualization. In: IEEE Symposium on Information Visualization, 2005. INFOVIS 2005, pp. 195–202. IEEE (2005)
44. Farrell, D., Moffat, D.C.: Adapting cognitive walkthrough to support game based learning design. Int. J. Game-Based Learn. (IJGBL). **4**(3), 23–34 (2014)
45. Salazar, K.: Evaluate Interface Learnability with Cognitive Walkthroughs. [online]. https://www.nngroup.com/articles/cognitive-walkthroughs/ (2022)
46. Farzandipour, M., Nabovati, E., Sadeqi Jabali, M.: Comparison of usability evaluation methods for a health information system: heuristic evaluation versus cognitive walkthrough method. BMC Med. Inf. Decis. Making. **22**(1), 1–1 (2022)
47. Yanez-Gomez, R., Cascado-Caballero, D., Sevillano, J.L.: Academic methods for usability evaluation of serious games: a systematic review. Multimed. Tools Appl. **76**, 5755–5784 (2017)
48. Koutsabasis, P., Gardeli, A., Partheniadis, K., Vogiatzidakis, P., Nikolakopoulou, V., Chatzigrigoriou, P., Vosinakis, S.: Field playtesting with experts' constructive interaction: an evaluation method for mobile games for cultural heritage. In: 2021 IEEE Conference on Games (CoG), pp. 1–9 (2021)

49. Korhonen, H.: Comparison of playtesting and expert review methods in mobile game evaluation. In: Fun and Games '10: Proceedings of the 3rd International Conference on Fun and Games, pp. 18–27 (2010)
50. Drachen, A., Mirza-Babaei, P., Nacke, L.E.: Games User Research. Oxford University Press (2018)
51. Carneiro, N., Darin, T., Pinheiro, M., Viana, W.: Using interviews to evaluate location-based games: lessons and challenges. J. Interact. Syst. **11**(1), 125–138 (2020)
52. Isbister, K., Schaffer, N.: Game usability: advice from the experts for advancing the player experience. Morgan Kaufmann Publishers, Burlington, MA (2008)
53. Krueger, R.A., Casey, M.A.: Focus groups A practical guide for applied research. Sage Publications, Thousand Oaks (2000)
54. Bruhlmann, F., Mekler, E.D.: Surveys in games user research. In: Drachen, A., Mirza-Babaei, P., Nacke, L. (eds.) Games User Research, pp. 141–162. Oxford University Press, Oxford (2018)
55. Rahman, M.M., Tabash, M.I., Salamzadeh, A., Abduli, S., Rahaman, M.S.: Sampling techniques (probability) for quantitative social science researchers: a conceptual guidelines with examples. Seeu Rev. **17**(1), 42–51 (2022)
56. Story, D.A., Tait, A.R.: Survey research. Anesthesiology. **130**(2), 192–202 (2019)
57. Hazan, E.: Contextualizing data. In: El-Nasr, M.S., et al. (eds.) Game analytics, pp. 477–496. Springer, London (2013)
58. Rigby, S., Ryan, R.: The Player Experience of Need Satisfaction (PENS) Model, pp. 1–22. Immersyve Inc (2007)
59. Ryan, R., Rigby, S., Przybylski, A.: The motivational pull of video games: a self-determination theory approach. Motiv. Emot. (2006)
60. Denisova, A., Cairns, P., Guckelsberger, C., Zendle, D.: Measuring perceived challenge in digital games: development & validation of the challenge originating from recent gameplay interaction scale (CORGIS). Int. J. Hum.-Comput. Stud., 137 (2020)
61. Jennett, C., Cox, A.L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., Walton, A.: Measuring and defining the experience of immersion in games. Int. J. Hum.-Comput. Stud. **66**(9), 641–661 (2008)
62. Isbister, K., Höök, K., Sharp, M., Laaksolahti, J.: The sensual evaluation instrument: developing an affective evaluation tool. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1163–1172 (2006)
63. Laaksolahti, J., Isbister, K., Höök, K.: Using the sensual evaluation instrument. Digit. Creativ. **20**(3), 165–175 (2009)
64. Ijsselsteijn, W.A., Poels, K., de Kort, Y.A.W.: The game experience questionnaire: development of a self-report measure to assess player experiences of digital games, deliverable 3.3. FUGA technical report, TU, Eindhoven, The Netherlands (2008)
65. Nacke, L.E., Grimshaw, M.N., Lindley, C.A.: More than a feeling: measurement of sonic user experience and psychophysiology in a first-person shooter game. Interact. Comput. **22**(5), 336–343 (2010)
66. Drachen, A., Nacke, L.E., Yannakakis, G., Lee Pedersen, A.: Correlation between heart rate, electrodermal activity and player experience in First-Person Shooter games. In: Spencer, S.N. (ed.) Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games, pp. 54–2010. ACM, Los Angeles, CA (49)
67. Ortega, M.C., Bruno, E., Richardson, M.P.: Electrodermal activity response during seizures: a systematic review and meta-analysis. Epilepsy Behav. **134**, 108864 (2022)
68. Paloniemi, S., Penttonen, M., Eteläpelto, A., Hökkä, P., Vähäsantanen, K.: Integrating self-reports and electrodermal activity (EDA) measurement in studying emotions in professional learning. In: Methods for Researching Professional Learning and Development, pp. 87–109. Springer, Cham (2022)
69. Richter, H., Krukewitt, L., Müller-Graf, F., Zitzmann, A., Merz, J., Böhm, S., Kühn, V.: High resolution EIT based heart rate detection using Synchrosqueezing (2022)

70. Thissen, B.A., Schlotz, W., Abel, C., Scharinger, M., Frieler, K., Merrill, J., Haider, T., Menninghaus, W.: At the heart of optimal reading experiences: cardiovascular activity and flow experiences in fiction reading. Read. Res. Q. **57**(3), 831–845 (2022)
71. Fuentes del Toro, S., Wei, Y., Olmeda, E., Ren, L., Guowu, W., Díaz, V.: Validation of a low-cost electromyography (EMG) system via a commercial and accurate EMG device: pilot study. Sensors. **19**(23), 5214 (2019)
72. Tortora, G., Derrickson, B.: Principles of Anatomy and Physiology, 12th edn. Biological Sciences Textbook (2009)
73. Ramos, A.L.A., Dadiz, B.G., Santos, A.B.G.: Classifying emotion based on facial expression analysis using Gabor filter: a basis for adaptive effective teaching strategy. In: Computational Science and Technology, pp. 469–479. Springer, Singapore (2020)
74. Westover, M.B., Gururangan, K., Markert, M.S., Blond, B.N., Lai, S., Benard, S., et al.: Diagnostic value of electroencephalography with ten electrodes in critically ill patients. Neurocrit. Care. **33**(2), 479–490 (2020)
75. Masood, K., Alghamdi, M.A.: Modeling mental stress using a deep learning framework. IEEE Access. **7**, 68446–68454 (2019)
76. Read, J.C., MacFarlane, S.J.: Using the Fun Toolkit and other survey methods to gather opinions in child computer interaction. In: Interaction Design and Children, IDC2006. ACM Press, Tampere (2006)
77. Lang, P.J.: The cognitive psychophysiology of emotion: fear and anxiety. In: Tuma, A.H., Maser, J.D. (eds.) Anxiety and the Anxiety Disorders, pp. 131–170. Lawrence Erlbaum Associates (1985)
78. Kujala, S., Roto, V., Mattila, K., Karapanos, E., Sinnela, A.: UX curve: a method for evaluating long-term user experience. Interact. Comput. **23**, 473–483 (2011)
79. Vissers, J., De Bot, L., Zaman, B.: MemoLine: evaluating long-term UX with children. In: Proceedings of the 12th International Conference on Interaction Design and Children, New York, pp. 285–288 (2013)