electronics

MDPI

Article

Development of the Engagement Playability and User eXperience (EPUX) Metric for 2D-Screen and VR Serious Games: A Case-Study Validation of Hellblade: Senua's Sacrifice

Kim Martinez [1], David Checa [2] and Andres Bustillo [2,*]

[1] Department of History, Geography and Communication, Universidad de Burgos, 09001 Burgos, Spain; kmartinez@ubu.es
[2] Department of Computer Engineering, Universidad de Burgos, 09006 Burgos, Spain; dcheca@ubu.es
* Correspondence: abustillo@ubu.es

Abstract: Research into the design of serious games still lacks metrics to evaluate engagement with the experience so that users can achieve the learning aims. This study presents the new EPUX metric, based on playability and User eXperience (UX) elements, to measure the capability of any serious game to maintain the attention of players. The metric includes (1) playability aspects: game items that affect the emotions of users and that constitute the different layers of the game, i.e., mechanics, dynamics and aesthetics; and (2) UX features: motivation, meaningful choices, usability, aesthetics and balance both in the short and in the long term. The metric is also adapted to evaluate virtual reality serious games (VR-SGs), so that changes may be considered to features linked to playability and UX. The case study for the assessment of the EPUX metric is Hellblade, developed in two versions: one for 2D-screens and the other for VR devices. The comparison of the EPUX metric scores for both versions showed that (1) some VR dynamics augmented the impact of gameplay and, in consequence, engagement capacity; and (2) some game design flaws were linked to much lower scores. Among those flaws were low numbers of levels, missions, and items; no tutorial to enhance usability; and lack of strategies and rewards to increase motivation in the long term.

Keywords: serious games; game design; game evaluation; game engagement; virtual reality

Citation: Martinez, K.; Checa, D.; Bustillo, A. Development of the Engagement Playability and User eXperience (EPUX) Metric for 2D-Screen and VR Serious Games: A Case-Study Validation of Hellblade: Senua's Sacrifice. *Electronics* **2024**, *13*, 281. https://doi.org/10.3390/electronics13020281

Academic Editors: Gudrun Klinker and David A. Plecher

Received: 20 November 2023
Revised: 22 December 2023
Accepted: 3 January 2024
Published: 8 January 2024

1. Introduction

Following Djaouti et al.'s work [1], serious games are tools that use video game structures both for a broad range of non-entertainment purposes and for varied audiences. The desired goals are either to convey a message, to teach a lesson, or to provide an experience, regardless of whether commercial objectives are at the same time achieved [2]. Developing serious games requires many types of expertise, so the process often lacks structural unification [3]. Different theories and models have been proposed for this task from various perspectives: educators, engineers, designers, players, etc. [4]. Designing and developing serious games is therefore challenging because there is no unified criterion to achieve an engaging experience that helps with the learning objective [5].

Engagement has been defined in many ways in serious game research [6]. When studying the definition of engagement during an educational experience, some authors focus on aspects such as behavior, motivation, cognitive effort and attention [7]. Other authors define it as an affect or emotion related to a pleasurable experience [8]. Finally, there are also studies that define engagement by referring to associated concepts: immersion, presence and flow [9]. Although there is no consensus on the definition of engagement, it can be identified as "a multidimensional construct with behavioral, affective and cognitive components that affect the intensity and emotional quality of a user's involvement in initiating and performing activities" [10].

In tests reported in recent studies of some relevance [11,12], engagement was found to be essential for achieving the goals of serious games. Furthermore, it is not only researchers who understand its importance, as players also believe that it helps them with their learning [13–15]. This research aims to identify the playability and User eXperience (UX) features that affect player's engagement for serious game design. The first objective of this paper is to propose a theory applicable to all purposes and scopes within the categories that Göbel [16] advanced, so multiple features are considered. The elements of the game design process that affect engaging gameplay must be identified to develop this framework [17].

From the study of these elements, the second and main objective was the design of a metric to evaluate engagement in serious games gameplay. The challenge of creating tools to assess serious game design was identified a decade ago in the research work of Bellotti [18]. It is a difficult task, considering that the serious games literature is still developing, besides which there are many different genres and application areas. Recent works have created tools that encompass different dimensions, which Göbel categorized in his literature review [16], dividing the evaluation into learning, gaming, and UX aspects. A review of these frameworks [19] shows that each one may be more or less complete, but they usually lack one of the three aspects to be assessed. Moreover, they offer only a qualitative analysis, so no score is attached to the design result. Furthermore, no study has been exclusively focused on engagement, which would be a mix of the mentioned gaming and UX aspects [20].

Although there are other studies, such as Reese's research [21], that develop metrics on gameplay aspects of serious games, they do not focus on engagement. They address flow, the cognitive state that provides a sense of deep enjoyment so rewarding that it motivates engagement in the learning environment [22]. These are two closely related terms, but engagement can be considered a component of flow so it will be studied more specifically in this study. There is, therefore, still a gap in this field, which this paper seeks to fill: a metric for the analysis of serious games that researchers and educators may use to score engagement elements.

However, this study must adapt to the experiences of virtual reality serious games (VR-SGs), which show great educational capabilities [23,24]. This technology creates immersive environments and very attractive interactions that affect engagement sensations [25]. Playability features in VR change gameplay perceptions and engagement [26–30]. Furthermore, UX is unlike other devices because of the headset screen and controllers, which affect the design needs [31,32]. Even people used to playing video games may find it difficult to adapt to VR interactions and environments [33]. Therefore, VR opens up a completely different way of engaging with serious games, so the metric has to be modified for the evaluation of VR-SGs.

In conclusion, this research tries to solve (1) the absence of focus on playability and UX elements that create engagement; (2) the lack of its quantitative evaluation in serious games; and (3) the adaptation of the developed metric to 2D and VR serious games. It draws from serious-game design models for researchers, considering playability elements and UX-related perceptions. Hence, our first objective was to pinpoint the main features that affect player engagement for 2D and VR serious games, as a theoretical guide that can be used in the design of a serious game. The second and the main objective was to develop two versions of the Engagement Playability and User eXperience (EPUX) metric (one for 2D-screen devices and another for VR devices) with which to evaluate aspects of playability and UX. In this study, its characteristics are summarized in two tables (one for each version of the metric), so that it is fully operative, which might explain the different elements. Additionally, these two tables are extended in File S1 as a fully operative metric. Finally, a case study of the serious game Hellblade: Senua's Sacrifice was conducted to validate the metric. Hellblade is a serious indie game that was developed to raise awareness of people suffering from psychotic episodes and the difficulties that they face. It was chosen because it is a quality game that adapts its playability and UX to 2D and VR versions and both versions of the game could therefore be tested with the metric. The strengths and the

weaknesses of the engagement design could be identified from the results. Additionally, as a sub-objective, levels of engagement in both versions were compared to show which visualization device (VR headsets or 2D-screens) was more engaging.

2. Materials and Methods

In this section, the main engagement elements from previous theories of serious games design are discussed. The elements applicable to 2D-screen serious games are also described, as are any necessary variations for VR experiences.

2.1. Serious Games Engagement Layers

Different game theories were studied in order to identify the elements of the game design process that affect engagement. The criterion was that these studies have to involve all the agents in the creative process of a serious game: developers, content specialists, and other stakeholders. The reason is because this metric is designed for specialized researchers who understand and collaborate in these projects. The five models identified were as follows: the Mechanics, Dynamics and Aesthetics (MDA), the Design, Play and Experience (DPE), the Design, Dynamics and Experience (DDE), the Learning Games Design Model (LGDM) and the Educational Digital Games (EDG) frameworks. A comparative and qualitative analysis of these theories was performed to identify similarities, differences and gaps between the layers and agents that affect the engagement aspects.

The MDA framework [34] with its 3 layers (mechanics, dynamics and aesthetics) formed the basis of engagement features that represent the main playability elements. Each category influences the others in order to create emotional responses, so that the player actively wants to play and to achieve the educational aims. But since the 3 layers of the MDA framework did not cover the entire gameplay, other studies such as the DPE model expanded them [3]. This framework includes other layers in the game design regarding learning content, storytelling and UX, as shown in Figure 1. On the one hand, the role of MDA layers in playability engagement is reaffirmed in DPE theory and UX is another added component that affects it. On the other hand, the importance of narrative and the inclusion of learning content as appealing for the player is also identified in DPE theory and included in the following theories. However, their effect on engagement is not studied in this paper as they are beyond the scope of playability features.

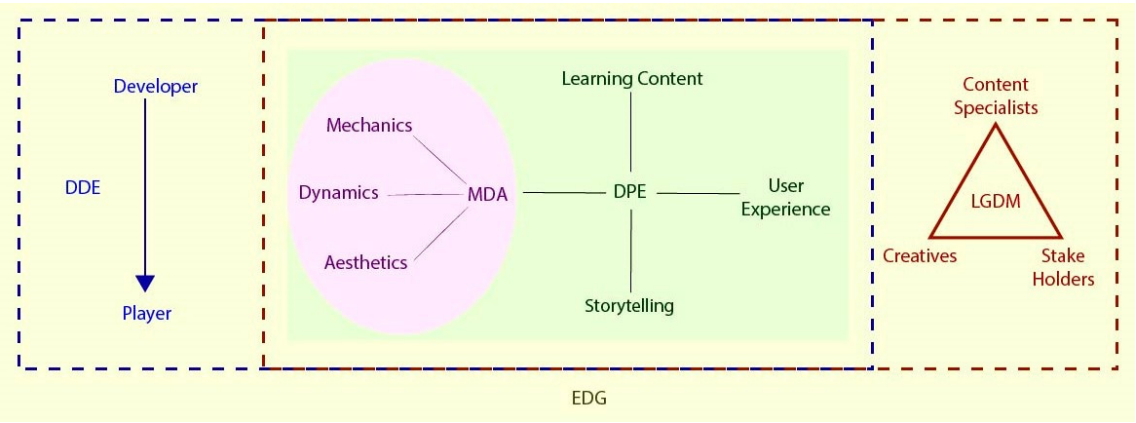

Figure 1. Scheme of the game design theories' interrelation.

The following 3 theories include the MDA and DPE layers that are summarized in Figure 1, as well as focusing on other particular agents of the game design process. The LGDM framework [6,25] is centered on the need for collaborative development between content specialists, creative experts, and stakeholders from design to dissemination. While content specialists show greater interest in learning content, stakeholders decide on the devices and resources that affect UX. In addition, creative experts are in charge of designing the MDA and the UX layers, which in a serious game design corresponds to the expert

researchers. Additionally, there is the DDE framework [35], which covers the development process of the MDA and the DPE layers, from game production to the player's eventual journey. The DDE framework is centered on mechanics game coding more so than on dynamics due to the unpredictability of the different type of players and their various behaviors. This theory emphasizes the importance of developers and their resources in the variety of mechanics that affect the engagement of players. Finally, the overall perspective of the previous 4 theories has been adapted to the EDG framework [4]. This framework follows the future line of research that should include both the MDA and the DPE layers, together with stakeholders and researchers, as with the LGDM model, and with development teams as with DDE theory.

The analysis of these serious games design theories allows us to identify the essential categories for games development. Regarding playability features, it has been shown how MDA layers and UX gather together the main elements of engagement. In addition, the variables that most affect development are (1) device choice due to its effect on UX and (2) capabilities of the researchers and development team to create attractive MDA layers.

2.2. Serious Games Engagement Elements

Engagement with a serious game should create an intrinsic motivation in the players, so that they wish to complete the experience. By doing so, it will be more likely that the games achieve the learning goals related to the intended message, lesson, or experience. The Key Characteristics of a Learning Game (KCLG) [36] define the emotions that should be induced to make a game engaging. The first one is challenge, which is defined by relevant, clear, and fixed goals in the game. Another is curiosity, understood in a sensory and cognitive way, and fantasy, which evokes mental images related to the player. Later, control was added as a feeling of self-determination and command of the learner [37]. There are various studies where the effect of these emotions on the serious purpose are discussed; however, these characteristics turn out to be fundamental when focusing on engagement.

More recently, Kim and Lee [38] developed the Dynamical Model of Gamification of Learning (DMGL). It is a framework that relates each emotion to the 3 layers of the MDA framework. The definition of each layer is: (1) mechanics are the components that are represented and the range of actions before the player; (2) dynamics are the predictable runtime behaviors that emerge from mechanics; (3) aesthetics are the audiovisual responses that evoke emotions in the player [39]. Regarding DMGL connections, game mechanics create difficulties for the players so that they feel the challenge. Elements of mechanics that build this feeling are levels, points, goals, and quests, which represent challenges for the player to overcome. Moreover, other elements such as rewards, badges, virtual items, and feedback (from either the game or other players) help to construct the fantasy elements of the game, as they reward the gameplay results. The category of dynamics is focused on creating challenge and awakening curiosity. Behaviors that produce rewards stimulate curiosity in the player, leading to progressive unlocking of the reward system and the appointments while continuing the game. Additionally, dynamic behavior patterns and systems based on time and vanquishing opponents encourage their desire to challenge themselves and others. Finally, game aesthetics are responsible for evoking a strong fantasy sensation when they evoke the virtual world. In this way, the audiovisual effects produce positive emotions such as love, beauty, delight, and surprise. Apart from these feelings, aesthetics also creates thrills, envy, connection, and comedy, as the player negotiates the dynamics, which in turn stimulate curiosity.

Figure 2 summarizes these connections explained by DMGL theory. As the MDA framework encompasses most of the main game elements, DMGL factors can be transformed into evaluation criteria. Consequently, the appearance of game components will be quantified in every action as the game progresses, as highlighted in the DDE framework [30]. In this way, the capacity of serious games to engage players with the objective can be measured [40,41].

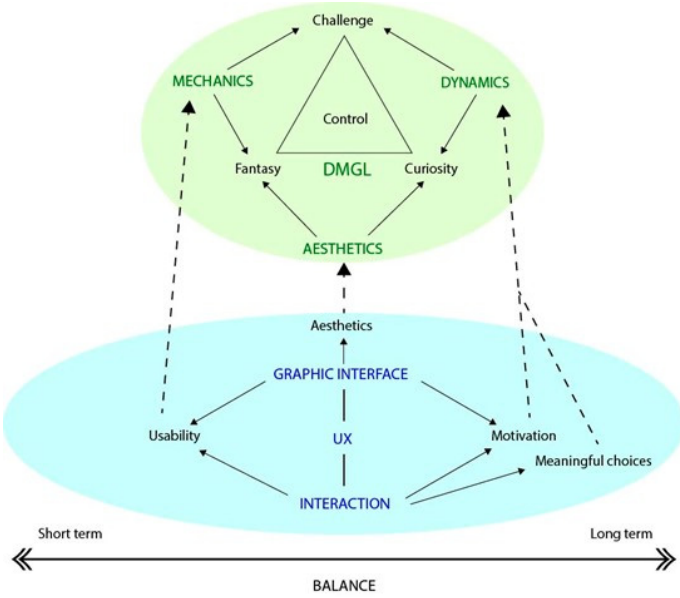

Figure 2. Scheme of DMGL theory and UX interrelation.

Nevertheless, the MDA framework lacks an essential aspect to allow the player to interact and engage with the gameplay. The UX layer that the DPE model identified was previously mentioned in the Introduction [3]. UX can be defined as the general effect of interactions with the game that are produced in the perceptions of the player [42]. UX design is based on two main elements: (a) graphic interface, interactive through a 2D-screen and peripherals, and (b) interaction, the player's experience and impact in the game [23]. Both features should be carefully planned to offer positive gameplay so that the user can enjoy the experience and fulfill the game objective.

Ferrara [43] identified the following UX features: motivation, meaningful choices, usability, aesthetics and the balance between these variables. Moreover, both long-term and short-term interactions should be considered for each feature. Motivation corresponds to the attraction that the initial user feels toward game interactions and the available rewards which maintain a long-term interest in achieving the goal. The second layer, meaningful choices, includes the decisions and interactions of each player that influence the game results. These are short-term tactics and long-term strategies that can be performed in individual or collaborative gameplay. The usability layer has to plan both interface and interaction elements to facilitate the perceptions of players of their own actions and the proposed goals. Interfaces should offer a sense of control from the start that can be progressively mastered. Aesthetic UX features relate to a pleasant direct sensory experience, which has to continue while the narrative unfolds and with it the game. Finally, balance measures how challenging and fair the game system is. UX must rate how much the basic game interactions are quickly learnt and perceived. Moreover, long-term aspects are measured by the player's effect throughout the game due to participation and progress.

The MDA framework defines playability aspects that are interrelated with UX elements, as shown in Figure 2. On the one hand, the most direct relationship is that of graphic interface aesthetics with MDA aesthetics, as both layers refer to audiovisual representation and the sensations it provokes. On the other hand, usability evokes a sense of control as the player masters the mechanics that the game offers. Finally, motivation related to the interactions and the meaningful choices about them depend on dynamic behaviors. All these aspects, and their balance and relationships, influence player engagement, so these are quantified in the EPUX metric. Likewise, the capacity of the development team referred to in the LGDM framework [6,25] must be taken into account in the evaluation. Their results are not comparable with the results of large studios with an in-house capacity to create commercial games.

2.3. Playability and UX Necessities for VR Experiences

Serious games developed for VR devices need an adaption of the EPUX metric. VR technology substantially changes player interactions, the control a player can exercise and emotional responses [44]. Consequently, playability and UX components must be individually considered for the advancement of VR serious-game design.

First, UX elements have to be examined because there were some issues found in earlier studies that had designed VR-SGs [33]. Those referring to UX development were as follows:

- Interactive experiences are preferred due to a "balance between costs, nowadays-technological development, immersion feeling and the possibilities that users have". On the contrary, passive experiences are usually limited and rarely achieve the proposed objectives.
- Immersion and usability are highly correlated with user satisfaction which is necessary for the game objective.
- Most studies show that users enjoy the experience, but the unfamiliar interface of VR technology limits the full potential for learning and training. In consequence, every VR-SG should include an extensive pre-training stage so users gain confidence and make the most benefit of its interaction with the VR environment.

Together with the LGDM framework, the importance of considering the development team and its limitations has been highlighted in this study. There are fundamental VR-related factors that will determine whether the experience is passive or active: VR devices, development possibilities, and team budget [29]. The researcher using the EPUX metric must try to be aware of these sorts of restrictions for the development of VR-SGs. Another no less important result is user satisfaction and how the interface is controlled and learnt.

First of all, the specific headset for which the game is developed must be considered. As one review underlined [33], VR users prefer experiences in which there is interaction with game components, following UX layers of motivation and meaningful choices. Therefore, devices such as Oculus Rift or HTC Vive will be better for serious games [45], because these headsets include interactive controllers and have 6 degrees of freedom (DoF). This characteristic means that the device tracks rotational and translational motion, so that there is interaction through the environment that creates more flow and presence. On the other hand, 3 DoF headsets such as Oculus Go, which only have rotational motion, will be less preferable due to their restrictive interaction [45].

A probable future problem was also detected in connection with development because of budget limitations. Designing high-visual-quality VR environments and explorative interaction experiences could be impracticable due to its high costs [33]. Hence, evaluating SG-VRs should consider the outcomes of gaming with those constraints in mind, without blindly penalizing game components that are not of high quality. On the basis of this interactive experience, some inconveniences such as VR sickness and headaches [22,46] must be avoided so that the interaction leads to user engagement. In addition, no highly complex actions using controllers or body movements within the VR environment should be implemented. When users are unfamiliar with VR controllers, this novel interaction causes difficulties in their physical and virtual movements [32]. Any of these mishaps could disturb the perceptions of users and affect their motivation to continue with the experience.

Another relevant aspect is UX possibilities of learning how to interact. SG-VRs should design a training tutorial to make users comfortable with movement and controls in the virtual environment. This usability feature will increase the perceptions of players with regard to their own actions and the game goals [31]. Additionally, UX design should achieve more difficult interactions according to the user's meaningful choices, thereby leaving more space for flow sensation. Finally, even if the game development is restricted, the interface should be audio–visually pleasant with regard to the quality of its aesthetics. This layer is a constant in every user interaction, so it has to be attractive, accessible, and easily understood. The interface may use controller buttons or movement, although any

changes that the player makes should be accompanied by visual or audio effects. These new metrics are collected in Figure 3, defining their connections with VR UX elements.

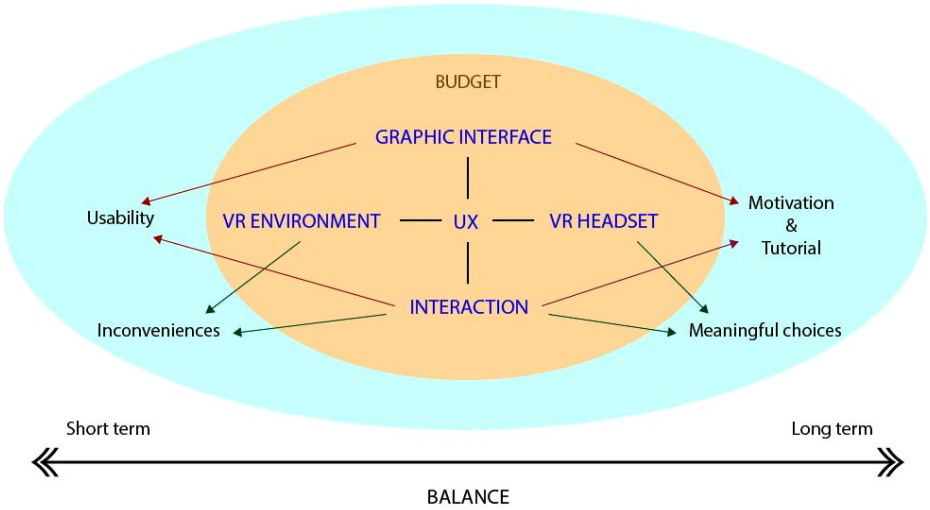

Figure 3. Scheme of proposed UX elements for VR.

Regarding playability features, mechanics and dynamics elements which involve challenge are exactly the same for VR-SGs [31]. The player likewise has to overcome the mechanics' quests, goals, and levels, as well as using strategies to cope with patterns and progression in the dynamics. In contrast, the way that fantasy mechanics features are perceived by VR users is radically different [5]. The experience changes from watching videos, items and characters on a 2D-screen to experiencing them within a VR surrounding. In addition, the design of the experience varies greatly depending on the type of interaction and the headset. When the game has 3 DoF, its players only have the freedom to look around, which lowers their engagement. At another level, a 6 DoF experience lets the players walk around and interact with objects and characters at the scene [27]. The possibility of being able to move in space and change the orientation and the perspective in which the virtual elements are displayed improves engagement and, therefore, the outcomes of the game [26].

Curiosity aspects are quite different, because of the specific flow within the VR environment. Options for exploration and interaction with characters, items, and venues will strongly affect user engagement. Once again, these metrics have to be reconsidered to assess the extent to which the process of unlocking and progressing is immersive for the player. Finally, the aesthetic layer which includes every visual and sound effect is quite different in a VR environment due to the surrounding sensations that it generates [27]. The emotions of fantasy and curiosity evoked within a VR environment will be stronger than 2D-screen experiences as long as the user perceives them as well integrated; otherwise, the emotional effect could be negative [30]. Audiovisual metrics must therefore be closely examined to ascertain whether positive instead of negative sensations are induced in the player when assuming the role of a spectator. Having adapted these aspects well, user involvement with playability should be assured.

3. Results

In this section, the EPUX metric is presented to evaluate engagement characteristics in serious games and to examine its strengths and flaws. A second version is also introduced to apply to VR-SGs with the necessary changes. The application of the EPUX metrics to the serious indie game Hellblade: Senua's Sacrifice is then discussed, both for the 2D and the VR versions. Finally, the results of testing the metrics with two study group assessments are presented to evaluate its independence with the user, once again for 2D-screen and VR devices.

3.1. EPUX Metric for 2D-Screen Serious Games

In this paper, the EPUX metric is proposed in order to apply the theories that are described in Section 2 to evaluate 2D-screen serious-game engagement. Table 1 reflects the scheme of the features that are evaluated and shows the maximum marks, while File S1 collects a detailed description of the proposed marks and levels for each category that is evaluated. This metric introduces a complete novelty when grouping the playability and UX aspects, as well as assigning scores in their evaluation. The total score of 100 will rate the playability and interaction possibilities. According to the result, it will be possible to estimate the capacity to engage the player during the game. A high score will mean a higher probability of the user achieving the expected goal. On the other hand, a low score will mean poorly engaged players, likely to jeopardize the serious game objective.

Table 1. EPUX metric for 2D-screen serious games (green for Playability emotions and blue for UX elements).

Playability Emotion	MDA Layer	Features	Marks
Challenge	Mechanics	Levels, points, goals, quests.	/10
	Dynamics	Time and enemy patterns, opponent players.	/10
Fantasy	Mechanics	Rewards, badges, virtual items, feedback.	/10
	Aesthetics	Audio and visual effects relating to love, beauty, delight, and surprise.	/10
Curiosity	Dynamics	Reward scheduling, progressive unlock, and appointments.	/10
	Aesthetics	Audio and visual effects regarding thrill, envy, connection, and comedy.	/10
SUBTOTAL			**/60**
UX element	**Term-time**	**Features**	**Marks**
Motivation	Short-term	Initial attraction to play.	/4
	Long-term	Available rewards during the game.	/4
Meaningful choices	Short-term	Possible tactics for player.	/4
	Long-term	Possible strategies for player.	/4
Usability	Short-term	Sense of control from the start.	/4
	Long-term	Possibility of mastering the interface.	/4
Aesthetics	Short-term	Pleasant direct sensory experience.	/4
	Long-term	Pleasant experience throughout the game.	/4
Balance	Short-term	Basic game interactions are quickly learnt.	/4
	Long-term	Progress of player affects the interface.	/4
SUBTOTAL			**/40**

Score distribution is based on the importance that is attributed to each feature in the theories described in the previous sections, from which the engagement aspects were extracted. MDA playability was identified as more important for engagement design than UX in the following frameworks: EDG [4], DDE [35], LGDM [6,25] and DPE [3]. The UX effect depends on the interaction with playability elements [42] so it has a lower score than MDA layers with a ratio of 1:1.5. In this way, playability is more fundamental, but a game difficulty could pass without a good UX. Within playability, Challenge, Fantasy, and Curiosity are each worth 20%, because the KCLG framework [36] establishes equity between its emotions. At the same time, each of the three emotions are, in DMLG theory [38], equally divided between two MDA layers involved in the feeling, as they are also equated in that theory. The appearance of the elements described above in each section must be taken into account to evaluate these features in a serious game. Each one will be valued at over 10 points for inclusion throughout the actions developed in the experience. Regarding UX

elements, the score is equally divided between the five UX features (Motivation, Meaningful choices, Usability, Aesthetics, and Balance), because these characteristics, as Ferrara [43] determined in his study, are of the same importance. The 8 points that each is worth are equally split between short- and long-term applications, as the author also established similar importance to motivate the gameplay from the beginning of the game to the end. The score will be determined according to the features available in the game interface from among all those described in Table 1.

3.2. EPUX Metric Modification for VR

Following the modifications under Section 2.3, this study proposes a variation of the EPUX metric to evaluate playability and UX elements in VR, as described in File S1 (a detailed description of the proposed marks and levels for each category for evaluation). The same total ratings will be used: 60% for playability and 40% for UX features. Regardless of the device for which the serious game is designed, the gameplay is prioritized over the UX upon which it is based. Nevertheless, some sections are separated or have more added features to take into account the distinct VR characteristics. This new structure is shown below in Table 2 as a scheme of the complete metric, with VR differences indicated in italics.

Playability scores are the ones that change the most in the VR version, as explained under Section 2.3, except for the challenge emotion. These challenge-related mechanics and dynamics have to be equally designed for every device, so they each have a maximum score of 10. On the other hand, mechanics for fantasy and dynamics for curiosity have to be evaluated in different ways depending on their interactivity [10,25–30]. In the same way, the aesthetic layer depends on the surrounding VR environment to generate emotions [27]. Therefore, these scores of 10 are divided equally into 5 points. One part will continue to measure the appearance of playability elements and the emotions that the aesthetics are intended to provoke, but another is meant to score their integration within the VR environment. If it is not properly implemented, the engagement emotions that are sought will not be evoked [27,30,33]. If the gameplay in VR is not perceptibly interactive, the player becomes a spectator, feeling neither fantasy nor curiosity. Unless the aesthetics are adapted to the virtual environment, emphasizing immersion and interactions, a sense of beauty or thrill will not be evoked in the player.

Regarding UX elements, the division and the ratings of features are exactly the same as for the 2D metric. The engagement of these experiences depends on interaction with the devices and how well they are designed, which should be evaluated in the same way in the 2D-screen as in the VR. What changes when evaluating SG-VRs is the aspects that have been explained under Section 2.2 and that are included in other sections. The chosen device and the controllers for the game affect motivation and meaningful choices. If these permit interaction with the virtual environment and items, the player will be driven to continue in the long term. Likewise, when these interactions are permitted, the range of tactics that the player can apply in the short term will increase. One of the main features to consider is the interface design in VR to facilitate user control. The inclusion of a tutorial will be valued in the short-term balance section, as it is based on general adaptation to the game. Continuing with this sense of control, short-term usability is used to evaluate whether there are complex actions within VR which could frustrate the player. The section on long-term usability will also take into account that some tools have been applied to avoid dizziness or headaches while playing. Finally, aesthetics always have to offer a pleasant experience, but in the long term, the budget limitations to develop the game should be considered.

Table 2. EPUX metric for Virtual Reality (green for Playability emotions and blue for UX elements; special features for VR are marked in italics).

Playability Emotion	MDA Layer	VR Features	Marks
Challenge	Mechanics	Levels, points, goals, quests.	/10
	Dynamics	Time and enemy patterns, opponent players.	/10
Fantasy	Mechanics	Rewards, badges, virtual items, feedback.	/5
		These are interactive and may or may not be integrated in the VR environment.	/5
	Aesthetics	Audio and visual effects regarding love, beauty, delight, and surprise.	/5
		The sense of surrounding intensifies the emotion.	/5
Curiosity	Dynamics	Reward scheduling, progressive unlock and appointments.	/5
		Interactive in nature and may or may not be explored.	/5
	Aesthetics	Audio and visual effects regarding thrill, envy, connection, and comedy.	/5
		A sense of interaction intensifies the emotion.	/5
SUBTOTAL			**/60**
UX element	**Term-time**	**VR Features**	**Marks**
Motivation	Short-term	Initial attraction to play.	/4
	Long-term	Available rewards during the game. *The device and the controllers permit interaction.*	/4
Meaningful choices	Short-term	Possible tactics of players. *Interaction dependent on device and controllers.*	/4
	Long-term	Possible player strategies.	/4
Usability	Short-term	Sense of control since the beginning. *There are no complex actions necessary to play.*	/4
	Long-term	Possibility of mastering the interface. *No dizziness or headaches caused.*	/4
Aesthetics	Short-term	Pleasant direct sensory experience.	/4
	Long-term	Pleasant experience throughout the game. *Budget limitations taken into account.*	/4
Balance	Short-term	Basic game interactions are quickly learnt. *There is a tutorial.*	/4
	Long-term	Progress of player affects the interface.	/4
SUBTOTAL			**/40**

3.3. Study Case: The Evaluation of Hellblade: Senua's Sacrifice

In this sub-section, the EPUX metrics presented in File S1 are tested on Hellblade: Senua's Sacrifice, shown in Figure 4. This is an independent AAA game (a high-quality low-budget experience) developed as an adult fantasy experience at Ninja Theory studios and released in 2017 [47]. Hellblade is a commercial game that addresses a social and educational objective in its development: to raise awareness of people living with mental illness. Specifically, the story is focused on Senua, the lead character, who suffers from psychotic episodes, lives with her illness and faces the effects on her path through life. The studio worked with expert neuroscientists and psychologists, mental illnesses associations and people affected by psychosis, so that the illness was not represented as mere fantasy. It is in fact based on scientific studies and real experiences. Hellblade is a commercial project that may not be academically classified as a serious game due to its lack of quantifiable knowledge outcomes. Nevertheless, the conscious objective and the application of the gameplay and educational content mean that the game is classified as a social awareness and impact game and can be studied as such. Both the narrative and the story that this video

game recounts have been analyzed in some papers in the context of an empathetic approach towards mental health awareness [48,49]. Particularly, the importance of collaboration between researchers and commercial developers was highlighted in the study of Fordham and Ball [50] so that players could engage with this important issue through a ludic framework. The game was adapted a year later to VR, which permitted the comparison between both versions in this case study.

Figure 4. Screen captures of Hellbalde: Senua's Sacrifice gameplay.

3.3.1. Evaluation of the 2D-Screen Device Edition

The application of the EPUX metric to Hellblade is extensively presented in File S2 and gives a total result of 59 out of 100. This score means that the design of both playability and the UX are acceptable for game engagement. Furthermore, the scores of both sections pass the minimum, 35/60 for playability aspects and 24/40 for UX elements. Regarding playability emotions, both challenge sections score 5 points. The mechanics of Hellblade offer basic goals to progress through the plot and enemies with dynamic patterns. At important story points, there are some goals and missions that have to be accomplished to move on. However, there are few changes to the basic goals that offer little variety throughout the story so defeating enemies becomes simple and there are no items to gain experience points or levels. Regarding the evocation of fantasy, there are only some Norse stone totems (as Figure 4 shows) that can be found, so this section only scored 3 for mechanics. On the other hand, aesthetics scored 7, because the effects are of a certain quality, creating an environment in which the player immerses and connects with Senua's situation. Curiosity is the playability emotion with the highest score: 15 of 20. The game is based on dynamics in which the player, acting as Senua, has to constantly solve puzzles and unlock the next test such as, for example, aligning scenario items to form a symbol (as Figure 4 shows). Once again, the aesthetics are quite good, using the sets and the extraneous voices to convey chilling emotions.

Every UX element rates 5 or 4 points of 8, meaning that the interface is acceptable, so the user will have no expected problem when playing. Regarding motivation, Hellblade presents an unknown and interesting environment with a narrative and effects that invite the player to explore the game and its simple interface facilitates gameplay. So, in the short term, there is high attraction, though the focus decreases in the long term due to the lack of interactive rewards. Likewise, the game is based on short-term tactics, based on how to unlock puzzles or defeat enemies, which depend on meaningful user choices. However, there is no kind of long-term strategy as the game has no experience progression or defined economy. Initially, the game shows little usability, because it includes no tutorial, so the player may at times lose track of the situation. As the game progresses, the controls

have to be mastered to fight and to solve the puzzles. The interface aesthetics are simple and attractive from the beginning as the interactive items are differentiated, encouraging a following. Nevertheless, the gameplay becomes somewhat repetitive throughout the experience. Finally, balance is quite good in the short term because it is easy to learn the necessary controls quite quickly. In contrast, some interactions increase in difficulty as the player progresses with the story.

3.3.2. Evaluation of the VR Edition

The virtual reality version of Hellblade tells the same story as the previous version. The player has to go through identical environments and face the same mechanics and dynamics to solve the puzzles. Nevertheless, some adaptations have been introduced in the gameplay for VR devices. In this edition, the control of Senua is still in the third person, but the camera through which the player observes the game changes positions according to the obstacles she encounters. These movements respond to the need to adapt to a 360° vision and prevent the user from becoming dizzy. Furthermore, at certain points within the game, the controls are locked and cinematics are played. In most instances, the perspective is changed to a screen where the video is played. In others, the VR environment is used, so that Senua speaks to the player directly, simulating very close contact. These different adjustments are shown in Figure 5.

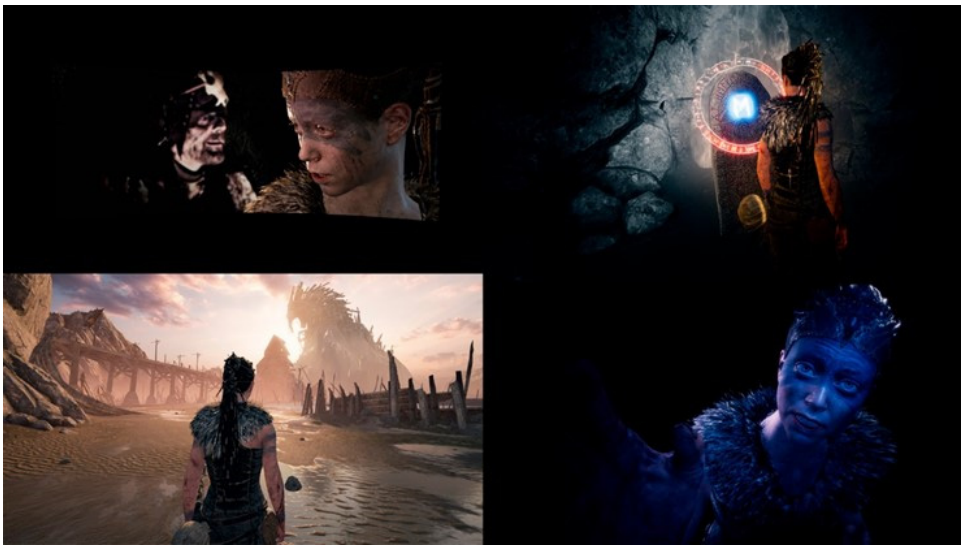

Figure 5. Screen captures of Hellbalde VR gameplay.

When the metric was applied to this version of Hellblade, it scored 67 out of 100, as shown in File S2. This score is the result of adding the 41 points of the playability part to the score of 26 for the UX features. All playability emotions had a minimum acceptable rating, but challenge had the lowest one in this table. Once again, the mechanics and dynamics of the game progress lacked sufficient variety, each of which only scored 5. When evaluating fantasy, the Norse stone totems, which are the mechanics items, were too few and not directly interactive, but were merely integrated in the environment, so this section received 3 points. On the other hand, fantasy aesthetics achieved the highest score of 10. The visual and sound effects were of great quality in VR. In fact, the player is completely surrounded and lives a delightful experience as Figure 5 shows. With regard to curiosity, the dynamics scored 8 points as they promote the progressive unlocking of gaming features, while the VR environment permits more direct interaction and exploration when solving the puzzles. Finally, the aesthetics once again scored very highly because of their quality and the interactive ability to intensify emotions.

The evaluation of the UX elements resulted in all of them achieving at least half of the full score. Starting with motivation, the initial attraction to play the game is very high, but

that attraction is lost as the experience continues. There are only a few totems throughout the story, and these have few interaction possibilities. The same situation repeats for the meaningful choices of the game. Hellblade includes a lot of tactics with which the player can interact with the environment and unlock the puzzles, yet is lacking in long-term strategies. Concerning usability, there is no tutorial to begin the control of Senua, so the user could be a little confused. However, the player can master the interface as the game progresses. In addition to this section, there are no development issues which could cause typical VR inconveniences, such as dizziness or headaches. The evaluation of aesthetics concluded that the game offered a satisfactory experience due to the audio–visual quality when starting playing, although this impression is not as strongly felt throughout the game. Finally, the short-term balance was considered to be good, because the player can easily learn to play, even without a tutorial. In the long term, some of the interactions become more difficult as both the player and Senua advance through the scenarios of the game.

3.4. The Evaluation by Study Groups

The study was performed with volunteers from the University of Burgos community who tried one of the versions. All were master's degree students, with similar socio-cultural characteristics and knowledge of video game design and multimedia experiences. The volunteers were explained the objective of this study and the metric features so that they would have a solid foundation for using the EPUX metric. The 2D-screen devices group was formed of twenty-four people and the VR group of twenty-three people, both groups being almost equal. Each group was informed of the educational purpose of the game and the EPUX metrics for its evaluation. They played the game for an hour and were then asked to fill in the metric and a survey including personal data, in addition to their interest in video games and former experience with VR devices. Likewise, they were asked to rate their experience and describe the video game with five adjectives. This information was sufficient to cluster the surveys into two types of evaluators: non-usual players and usual players. The correlation between the results and age and gender was studied but did not yield any results relating these factors to the engagement evaluation. The results, the average scores and standard deviations, along with the individual results of the researcher are available online [51], while the main outputs are summarized in Figures 6 and 7.

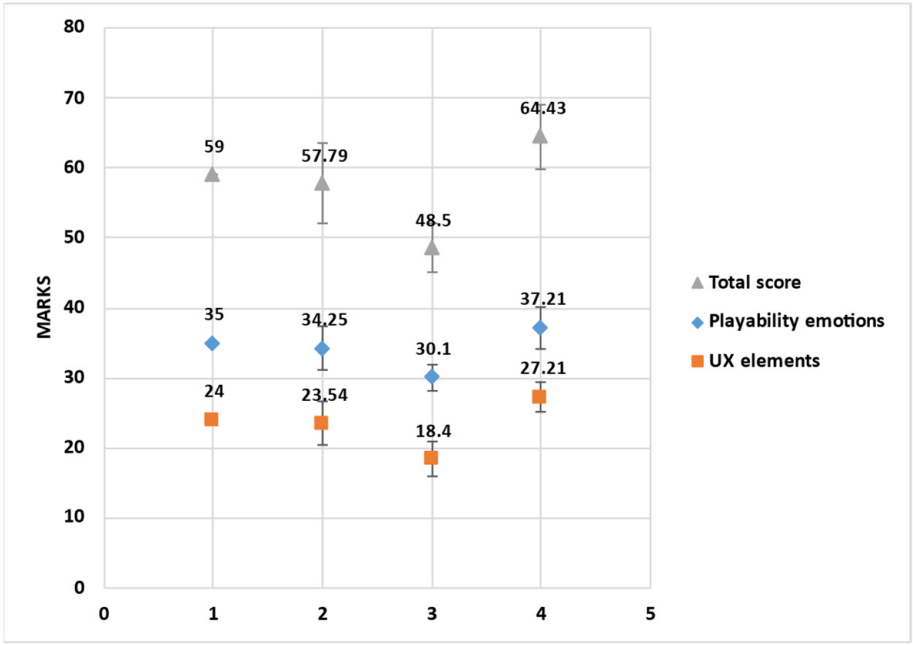

Figure 6. 2D-version scores for Hellblade: Senua's sacrifice.

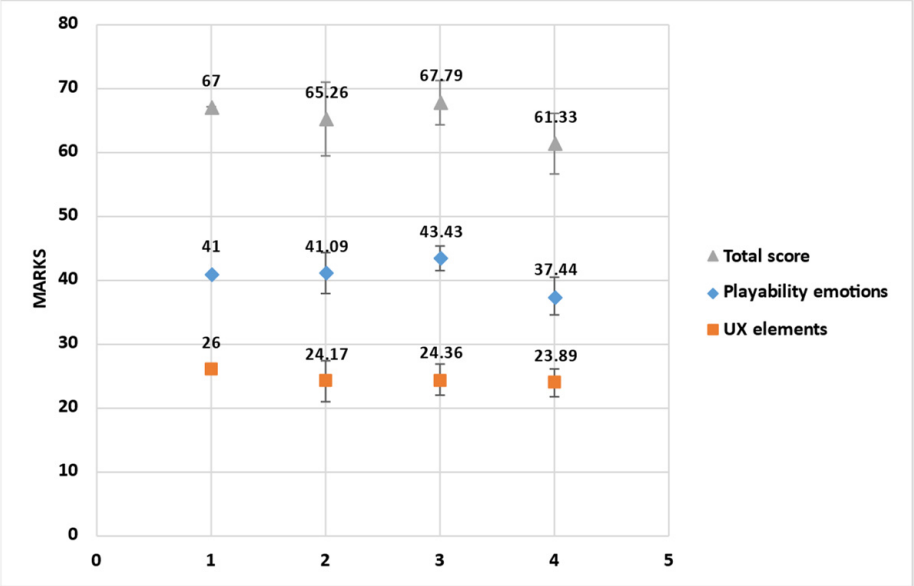

Figure 7. VR-version scores for Hellblade: Senua's sacrifice.

3.4.1. Group Results of the 2D-Screen Device Edition

This study group consisted of twenty-four people, ten women and fourteen men in their twenties up to their fifties. Fourteen of the participants were used to video games, while the other ten never usually played. Regarding the playability emotions metric, the average subtotal was 37.21, set alongside the result of 35 from the researcher, with a standard deviation of 6.31. Looking at the features' averages in Figure 6, the values were quite close to the original researcher outcome and the standard deviation was acceptable at lower than 2, as the total score was 10. The lowest variation, 1.48 corresponded to the fantasy emotion aesthetics, indicating that players agreed on the video game aesthetic's ability to recreate the environment. In fact, the study group scores reaffirmed those of the individual researcher, highlighting the audio–visual elements, compared to the more disappointing scores for mechanics and dynamics.

As regards the UX elements metric, the average subtotal score was 23.54, almost equal to the result of 24 from the researcher, and the standard deviation was 6.28, as is summarized in Figure 6. Both values indicated even closer assessments than in the playability section. Furthermore, the characteristic averages differed in quite a similar way to the previous section, in most cases distanced below one with respect to the scores of the experts, and the standard deviation was around 1, a relatively high number considering that 4 was the highest score. These marks also showed that the metric offered a comparable measurement of both the long- and the short-term characteristics of the game.

The total average score was 57.79, close to the score of 59 of the researcher, with a standard deviation between both scores of 11.54. The conclusion that can be drawn from these ratings is that, according to the participants, the game barely met the playable characteristics necessary for a serious game. This assessment corresponds to the Hellblade playability weaknesses analyzed in Section 3.3.1. However, a great difference may be observed within this group when considering the personal data of the group members. On the one hand, the participants who were accustomed to playing scored the video game with higher scores, similar to those of the researcher. In addition, the correlation coefficient between the EPUX metric total scores and their perception of the game as a fun experience is 0.37. On the other hand, those who hardly ever or never played video games tended to give lower ratings, especially in the UX section. In this case, the coefficient of relationship between total scores and fun perception is much higher at 0.70. A linear regression fit for both cases also provides some arguments in the same direction: although the correlation is mainly not linear between both variables (R^2 0.14 and 0.49,

respectively), it also shows how the total EPUX score does not strongly depend on the fun perception (EPUX = 2.00 × Fun + 38.17 and EPUX = 3.19 × Fun + 39.85, respectively), but the dependence is stronger in those who hardly ever or never played video games. This comparison reveals that an absence of both gaming experience and predisposition to play video games affected the perceptions of the evaluator with regard to motivation and possible game options. Usual players are able to evaluate game design features more objectively without being affected by the greater or lesser enjoyment of the experience.

3.4.2. Group Results of the VR Edition

The second group consisted of nine females and fourteen males, also aged between their twenties and fifties. Of these, fourteen people expressed less interest in video games and nine much higher levels of interest. Nevertheless, most participants barely had any experience with VR, which was therefore a completely new experience for them all. Concerning the playability metric, the subtotal average score was 41.9, a value almost identical to the score of 41.0 of the researcher, and the standard deviation was 6.66, as shown in Figure 7. Regarding the ratings for each characteristic, the results showed a standard deviation of around 1, which underlined the very similar scores from the participants in most cases. As with the evaluation of the researcher, the highest scores were matched with the perception of aesthetics and their integration in the virtual environment. On the contrary, the most distant average values corresponded to the mechanics that caused fantasy. This variation may be due to the fact that most participants have played VR for the first time; thus, they perceived virtual items in more immersive and suggestive ways.

In the UX element section, the sub-total average score was 23.89, once again close to the researcher marks of 26.0, and the standard deviation was 4.3, as is summarized in Figure 7. The standard deviation of each aspect was around 1.0. As with the previous study group, the short-term and the long-term characteristics once again showed similar deviations. These data reaffirm the stability of the metrics. On the one hand, the average values closest to the scores of the researcher were usability and aesthetics, as the students themselves described the game as "attractive, interesting and immersive". On the other hand, motivation, meaningful choices, and balance aspects were more subjective elements, dependent on the way they perceived their control over the game.

The total average score of the group members was 65.26, almost equal to the score of 67.0 of the researcher, with a standard deviation of 10.01. So, the scores of this study group also confirmed the evaluation of the researcher, awarding the VR version a higher score, and thereby confirming that this technology increased the playability and learning of this video game. Regarding the way in which the participants' level of familiarity with video games affected their scores, the group members with greater familiarity gave lower ratings to the playability aspects and the UX balance than the other group members. These participants also have a much lower correlation coefficient between total scores and fun perception (0.33) than non-usual players (0.78). The same occurs with the linear regression fits, which are lower with habitual players (EPUX = 4.38 × Fun + 30.15) than with non-habitual players (EPUX = 6.02 × Fun + 25.87). As in previous case, the linear regression fits for both cases just provide more some arguments in the same direction, although the correlation is mainly not linear between both variables (R^2 0.11 and 0.61, respectively). The higher expectations of regular players with regard to specifications and control over the game interface might explain these lower scores. And like the PC version, they are able to evaluate playability features more objectively.

4. Discussion

The EPUX metric has been developed for the evaluation of serious game engagement among expert designers. One of the researchers applied the metric to both versions of Hellblade: Senua's Sacrifice to test the capability of the metric to assess these elements and to compare 2D-screen and VR applications. The analysis was also conducted with two groups of university volunteers to reduce subjectivity in the case study. The sample

consisted of people with an interest in the experience and with a similar level of studies which could produce homogeneous results. A comparison of the scores of each study group with the scores of the researcher was intended to establish whether people with different levels of gaming experience could apply the metric with roughly similar results when looking at the weaknesses and strengths of different features of Hellblade. If the results were very different, it could be determined that the EPUX metric should be better explained or that some other aspects may need to be included.

The evaluation of the researcher of both versions of this game with the EPUX metric yielded a positive result. Therefore, Hellblade: Senua's Sacrifice has well-designed playability and UX for its learning aim to raise awareness of mentally ill people living through psychotic episodes. However, there are some differences between the ratings of the two editions. Starting with the final result, the first version scored 59 out of 100, and the VR version scored 67. Comparing both tables, the different scores showed how the same game using a VR headset offered certain advantages over traditional video game devices.

The playability results for the 2D-screen device and the VR version improved from 35 to 41. The main increase in this section was due to the difference in the perception of aesthetics. From the start, Hellblade has a high-quality audio–visual design adding an emotive attachment for the player with the storyline. When the game transfers these features to the VR environment in a way that surrounds the player, the sensations that the game conveys are maximized. Fantasy and curiosity aesthetics are therefore scored highest of all. As for dynamics, there was only a slight rise in the curiosity section. This change is because unlocking the puzzles was perceived in a more interactive and direct way in the VR game. The scores for the other features of dynamics and all the mechanics were exactly the same for both versions. All objectives, enemies, patterns, and items remain the same in the story. In conclusion, what VR enhances is the feeling of fantasy, bonding with the imaginary world, and curiosity to become familiar with the experience.

UX scores barely differed, ranging from 24 in the 2D-screen device version to 26 in the VR version. Motivation, meaningful choices, usability and balance sections were equally scored, as shown in both tables of File S2. The similar ratings should be pointed out, because it means that the game has been well adapted for VR. The experience neither caused control problems in the VR environment, nor generated dizziness, nor headaches for the players. The only section that changed was aesthetics, following the same trend as in the playability evaluation. VR UX aesthetics was awarded 2 more points, because the sensory experience was greatly increased on this device. However, as in the previous version, the long-term section also lost a point, because the environment was finally quite repetitive.

Furthermore, the results of both study groups confirmed the suitability of these metrics for use as an evaluation tool for more serious games, both in VR and in traditional devices. The total averages yielded very similar scores to the ratings of the researcher, which sustain the strengths and weaknesses found in Hellblade. From the case study, an ideal evaluator profile can also be defined: a person with some experience and liking for video games. This background might mean that the person is more critical with playability and more objective with the UX. It is also likely that in future experiences, the playing time will need to be further extended, so that the participants are able to assess the characteristics more accurately. Likewise, these tests point out that certain UX elements may be more subjective, such as motivation, meaningful choices, and balance aspects.

In summary, the Hellblade case study demonstrated that the metrics served to evaluate the engagement design of a serious indie game. The various features encompass all the important playability and UX components. Analyzing these characteristics provides a result to determine whether the experience is engaging. The proposed tables can be used to assign an equitable score to the importance of the elements that are identified and can be easily understood. In this way, observing the results can determine whether any MDA layer has flaws and how the design may be improved. Likewise, the problems that the user may have in the interaction with the game or devices will be identified and classified as short- or long-term issues.

Specifically, this study case has shown how a well-adapted experience to VR can increase the gameplay and, therefore, the engagement capacity. A comparison of the previous tables also showed that the low scores in some sections responded to flaws in game design. The change in these features and their application might be practically the same in both versions. Regarding playability, the game might need to include more varied levels, missions, items and rewards throughout its duration. These aspects could improve the scores for mechanics and dynamics to offer a greater challenge. With regard to UX elements, introducing a tutorial to explain the controls at the beginning of the game might enhance usability and balance. Furthermore, adding strategies and rewards in the long term might also increase motivation and scores for meaningful choices.

5. Conclusions

In this study, the main elements that affect serious games engagement and the proposed EPUX metric have been gathered together for their evaluation. This study can be applied to 2D-screen and VR experiences through two versions of the metric. All aspects under consideration fit into playability and UX characteristics. The VR metric needs some changes that are associated with its user interactions, unlike any other device experience. Both versions were presented in two tables, introducing the novelty of scored results to measure the assessment. The validity of the metric was tested in a case study of an indie serious game, pointing out its weaknesses and strengths and showing results that could be applied to future developments.

This work has also been focused on the capabilities of VR in serious game research due to its level of engagement. The virtual environment affects the player's feelings to commit with the game, which is essential for the learning goal. In the first instance, every VR-SG should be designed according to the VR headset and the development team budget limitations. From thereon, virtual items and interactions must be well adapted both to the issues that have been explained and to the narrative so as to ensure the learning aims of the game. Another notable concern in this paper is the adjustment of usability to VR experiences. Acquiring VR control in the game should be offered through interactive tutorials. Once these interactive matters are resolved, the game will prevent disturbing perceptions for VR users. VR-SG development could then be focused on applying VR capabilities to increase the emotive reactions of players through playability elements. These features will involve items that are integrated in the game environment, player interactions, and visual and sound effects.

Finally, as mentioned, the new metrics and evaluation tables have been tested on Hellblade: Senua's Sacrifice, a serious indie game developed to raise awareness of the difficulties of mentally ill people living through psychotic episodes. The case study has shown how the EPUX metric can serve a habitual video game player to evaluate the engagement design. The similar scores from both the researcher and the study group members, together with the correlation coefficients and linear regression fits, have demonstrated that the metric is well defined to understand the features being assessed. Nevertheless, the methodology followed in which usual and non-usual players have participated has its limitations since EPUX metric is intended for game design experts. It is expected that the evaluation of usual players will be similar to that of experts based on the results obtained, but future validations are necessary.

However, the scores help to determine playability flaws and UX that may cause short- or long-term problems. In particular, this game might need the inclusion of a tutorial at the beginning and more mechanics, dynamics, and potential strategies throughout the story [52]. Additionally, Hellblade's adaption shows how VR can increase engagement capacity using the available technology. As the VR environment is well implemented, the perceived emotions of the players are enhanced. It also offers better playability and UX, because its mechanics and dynamics produce an active experience involving movement around the scenario and interaction with puzzles and enemies [53].

Future works will apply the EPUX metric to different serious games with more evaluators to reinforce its utility for assessing these engagement experiences. Likewise, these future tests will address the characteristics of the evaluators and include a quantitative evaluation of their impressions of the metric. We will continue to explore the relationship of scores between those who are usual and non-usual players, as results have shown that regular players evaluate more objectively. We expect their assessment to be similar to that of game design experts, so we will seek the participation of game designers to verify this hypothesis and the usefulness of the EPUX metric in the evaluation of design engagement. In this process, some layers of the EPUX metric will be studied in more detail, such as aesthetics, where image quality, intensity, color harmony and saturation, quality of sound effects and music, and level of detail and realism can be quantified [54]. Additionally, VR characteristics will be further analyzed to synthetize each game element, influencing its intrinsic sensations [55]. From there, more developed and detailed metrics will be defined to create future VR-SG engagement frameworks. These developments will lead to comparisons between the improvements outlined in this study and subsequent game evaluations. Additionally, the value of serious games that may be built on the basis of these key factors and features will be tested and, likewise, their validity or need for further redefinition confirmed. Finally, not only the engagement aspects of playability and UX will be studied, but also how these aspects are related to the learning content to guarantee that the learning aims may be achieved [56].

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/electronics13020281/s1, File S1: Detailed rubric of the 2D and VR versions of EPUX metric.; File S2: Comprehensive evaluation of the 2D and VR versions of Hellblade: Senua's Sacrifice with EPUX metric.

Author Contributions: Conceptualization: K.M. and A.B.; Methodology: A.B.; Formal Analysis: A.B.; Funding Acquisition: K.M. and A.B.; Investigation: K.M.; Resources: K.M. and D.C.; Validation: K.M. and D.C.; Project Administration: A.B.; Supervision: A.B.; Writing and Original Draft: K.M. and D.C.; Writing, Review and Editing: A.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially supported by the GruaRV project (Reference Number INVESTUN/18/BU/0002) of the Consejería de Industria, Comercio y Empleo of the Junta de Castilla y León (Spain) and the Ministry of Science, Innovation and Universities (FPU18/04688).

Data Availability Statement: Data are contained within the article and Supplementary Materials.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Djaouti, D.; Alvarez, J.; Jessel, J.P. Classifying Serious Games. In *Advances in Game-Based Learning*; Routledge: London, UK, 2011; pp. 118–136. [CrossRef]
2. Nasution, N.K.G.; Jin, X.; Singgih, I.K. Classifying games in container terminal logistics field: A systematic review. *Entertain. Comput.* **2021**, *40*, 100465. [CrossRef]
3. Winn, B. The design, play, and experience framework. In *Handbook of Research on Effective Electronic Gaming in Education*; ISR: New York, NY, USA, 2009; Volume 3, pp. 1010–1024.
4. Pereira de Aguiar, M.; Winn, B.; Cezarotto, M.; Battaiola, A.L.; Varella Gomes, P. Educational Digital Games: A Theoretical Framework About Design Models, Learning Theories and User Experience. In *Design, User Experience, and Usability: Theory and Practice*; Springer International Publishing: Cham, Switzerland, 2018; pp. 165–184. [CrossRef]
5. Checa, D.; Bustillo, A. Advantages and limits of virtual reality in learning processes: Briviesca in the fifteenth century. *Virtual Real.* **2019**, *24*, 151–161. [CrossRef]
6. Hookham, G.; Nesbitt, K. A Systematic Review of the Definition and Measurement of Engagement in Serious Games. In Proceedings of the Australasian Computer Science Week Multiconference, ACM, Sydney, Australia, 29–31 January 2019; p. 42.
7. Lim, C.P. Global citizenship education, school curriculum and games: Learning Mathematics, English and Science as a global citizen. *Comput. Educ.* **2008**, *51*, 1073–1093. [CrossRef]
8. Xhembulla, J.; Rubino, I.; Malnati, G.; Barberis, C. Intrigue at the Museum: Facilitating Engagement and Learning through a Location-based Mobile Game. In Proceedings of the 10th International Conference on Mobile Learning, Madrid, Spain, 28 February–2 March 2014.

9. Kiili, K.; Lainema, T. Foundation for measuring engagement in educational games. *J. Interact. Learn. Res.* **2008**, *19*, 469–488.
10. Chamberlin, B.; Trespalacios, J.; Gallagher, R. Bridging research and game development: A learning games design model for multi-game projects. In *Gamification in Education: Breakthroughs in Research and Practice*; IGI Global: Hershey, PA, USA, 2018; pp. 66–88.
11. Chen, C.-C.; Tu, H.-Y. The Effect of Digital Game-Based Learning on Learning Motivation and Performance Under Social Cognitive Theory and Entrepreneurial Thinking. *Front. Psychol.* **2021**, *12*, 750711. [CrossRef] [PubMed]
12. Sanchez, D.R.; Nelson, T.; Kraiger, K.; Weiner, E.; Lu, Y.; Schnall, J. Defining motivation in video game-based training: Exploring the differences between measures of motivation. *Int. J. Train. Dev.* **2021**, *26*, 1–28. [CrossRef]
13. Rosenthal, S.; Ratan, R.A. Balancing learning and enjoyment in serious games: Kerbal Space Program and the communication mediation model. *Comput. Educ.* **2022**, *182*, 104480. [CrossRef]
14. Czauderna, A.; Guardiola, E. The Gameplay Loop Methodology as a Tool for Educational Game Design. *Electron. J. E-Learn.* **2019**, *17*, 207–221. [CrossRef]
15. Lu, Y.-L.; Lien, C.-J. Are They Learning or Playing? Students' Perception Traits and Their Learning Self-Efficacy in a Game-Based Learning Environment. *J. Educ. Comput. Res.* **2019**, *57*, 1879–1909. [CrossRef]
16. Göbel, S. Serious Games Application Examples. In *Serious Games*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 319–405.
17. Shoukry, L.; Gobel, S. Reasons and Responses: A Multimodal Serious Games Evaluation Framework. *IEEE Trans. Emerg. Top. Comput.* **2017**, *8*, 245–255. [CrossRef]
18. Bellotti, F.; Berta, R.; De Gloria, A. Designing Effective Serious Games: Opportunities and Challenges for Research. *Int. J. Emerg. Technol. Learn. (IJET)* **2010**, *5*, 22. [CrossRef]
19. De Almeida, J.L.F.; Machado, L.d.S. Design requirements for educational serious games with focus on player enjoyment. *Entertain. Comput.* **2021**, *38*, 100413. [CrossRef]
20. Cano, S.; Collazos, C.A.; Aristizábal, L.F.; Gonzalez, C.S.; Moreira, F. Assessing User Experience for Serious Games in Auditory-Verbal Therapy for Children with Cochlear Implant. In Proceedings of the World Conference on Information Systems and Technologies, Lecce, Italy, 4–5 June 2020; pp. 861–871.
21. Reese, D.D. CyGaMEs Selene player log dataset: Gameplay assessment, flow dimensions and non-gameplay assessments. *Br. J. Educ. Technol.* **2015**, *46*, 1005–1014. [CrossRef]
22. Menezes, F.M.; Silva, I.C.S.; Frosi, F.O. Game User Experience (UX): Explorando a Teoria da Diegese. SBC. In Proceedings of SBGames 2017, XVI SBGames, São Paulo, Brazil, 2–4 November 2017.
23. Martinez, K.; Menéndez-Menéndez, M.I.; Bustillo, A. Awareness, Prevention, Detection, and Therapy Applications for Depression and Anxiety in Serious Games for Children and Adolescents: Systematic Review. *JMIR Serious Games* **2021**, *9*, e30482. [CrossRef] [PubMed]
24. Eichhorn, C.; Plecher, D.; Klinker, G. VR enabling knowledge gain for the user (VENUS). In Proceedings of the 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct, Singapore, 17–21 October 2022.
25. Chamberlin, B.; Trespalacios, J.; Gallagher, R. The learning games design model: Immersion, collaboration, and out-comes-driven development. *Int. J. Game-Based Learn (IJGBL)* **2012**, *2*, 87–110. [CrossRef]
26. Downs, E.; Willemsen, P.; Leff, D.; Adigun, C.; Kothapalli, S.S.; Miller, L.; Boynton, K.; Berndt, M.M. Grave errors: Exploring the influence of motion mechanics on learning outcomes in a virtual cemetery. *Technol. Mind Behav.* **2022**, *3*, 1–12. [CrossRef]
27. Tong, L.; Lindeman, R.W.; Regenbrecht, H. Adaptive Playback Control: A Framework for Cinematic VR Creators to Embrace Viewer Interaction. *Front. Virtual Real.* **2022**, *2*, 798306. [CrossRef]
28. Ferguson, C.; Broek, E.L.v.D.; van Oostendorp, H. On the role of interaction mode and story structure in virtual reality serious games. *Comput. Educ.* **2019**, *143*, 103671. [CrossRef]
29. Radianti, J.; Majchrzak, T.A.; Fromm, J.; Wohlgenannt, I. A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Comput. Educ.* **2019**, *147*, 103778. [CrossRef]
30. Pallavicini, F.; Pepe, A.; Minissi, M.E. Gaming in Virtual Reality: What Changes in Terms of Usability, Emotional Response and Sense of Presence Compared to Non-Immersive Video Games? *Simul. Gaming* **2019**, *50*, 136–159. [CrossRef]
31. Irshad, S.; Perkis, A.; Azam, W. Wayfinding in Virtual Reality Serious Game: An Exploratory Study in the Context of User Perceived Experiences. *Appl. Sci.* **2021**, *11*, 7822. [CrossRef]
32. Wang, D.; Wang, X.; Zheng, Q.; Tao, B.; Zheng, G. How Interaction Paradigms Affect User Experience and Perceived Interactivity in Virtual Reality Environment. In *International Conference on Human-Computer Interaction*; Springer International Publishing: Cham, Switzerland, 2020; pp. 223–234.
33. Checa, D.; Bustillo, A. A review of immersive virtual reality serious games to enhance learning and training. *Multimed. Tools Appl.* **2020**, *79*, 5501–5527. [CrossRef]
34. Hunicke, R.; Leblanc, M.; Zubek, R. MDA: A formal approach to game design and game research. In Proceedings of the AAAI Workshop on Challenges in Game AI, San Jose, CA, USA, 25–26 July 2004.
35. Walk, W.; Görlich, D.; Barrett, M. *Design, Dynamics, Experience (DDE): An Advancement of the MDA Framework for Game Design*; Korn, O., Lee, N., Eds.; Game Dynamics; Springer: Cham, Switzerland, 2017; pp. 27–45. [CrossRef]
36. Malone, T.W. What makes things fun to learn? In *A Study of Intrinsically Motivating Computer Games*; Technical report; Xerox Palo Alto Research Center: Palo Alto, CA, USA, 1980.

37. Malone, T.W.; Lepper, M.R. *Making Learning Fun: A Taxonomy of Intrinsic Motivations for Learning*; Snow, R.E., Farr, M.J., Eds.; Aptitude, learning and instruction III; Routledge: London, UK, 1987.
38. Kim, J.T.; Lee, W.-H. Dynamical model for gamification of learning (DMGL). *Multimedia Tools Appl.* **2013**, *74*, 8483–8493. [CrossRef]
39. Junior, R.; Silva, F. Redefining the MDA Framework—The Pursuit of a Game Design Ontology. *Information* **2021**, *12*, 395. [CrossRef]
40. Alexiou, A.; Schippers, M.C. Digital game elements, user experience and learning: A conceptual framework. *Educ. Inf. Technol.* **2018**, *23*, 2545–2567. [CrossRef]
41. Leonardou, A.; Rigou, M.; Panagiotarou, A.; Garofalakis, J. Effect of OSLM features and gamification motivators on motivation in DGBL: Pupils' viewpoint. *Smart Learn. Environ.* **2022**, *9*, 14. [CrossRef]
42. Schell, J. *The Art of Game Design: A Book of Lenses*; Elsevier: Amsterdam, The Netherlands, 2008.
43. Ferrara, J. *Playful Design: Creating Game Experiences in Everyday Interfaces*; Rosenfeld Media: New York, NY, USA, 2012.
44. Stapleton, A.J. Serious Games: Serious Opportunities. In Proceedings of the Australian Game Developers' Conference, Washington, DC, USA, 8–10 July 2004.
45. Atsikpasi, P.; Fokides, E. A scoping review of the educational uses of 6DoF HMDs. *Virtual Real.* **2021**, *26*, 205–222. [CrossRef]
46. Yue, K.; Wang, D. EEG-Based 3D Visual Fatigue Evaluation Using CNN. *Electronics* **2019**, *8*, 1208. [CrossRef]
47. Antoniades, T. The Independent AAA Proposition–Tameem Antoniades. Medium. 25 May 2018. Available online: https://medium.com/@ninjatam/the-independent-aaa-proposition-3bc6dd657240 (accessed on 17 May 2022).
48. Firmino, J.G.A. Hellblade: Loucura e performance em videogames. *Rev. Fragm. Cult.–Rev. Interdiscip. Ciências Humanas* **2018**, *28*, 335. [CrossRef]
49. Ings, S. Hallucinations of Hellblade. *New Sci.* **2017**, *235*, 47. [CrossRef]
50. Fordham, J.; Ball, C. Framing Mental Health Within Digital Games: An Exploratory Case Study of Hellblade. *JMIR Ment. Health* **2019**, *6*, e12432. [CrossRef]
51. EPUX Surveys for Hellblade: Senua's Sacrifice Validation. Development of the Engagement Playability and User Experience (EPUX) Metric for 2D-Screen and VR Serious Games: A Case-Study Validation of "Hellblade: Senua's Sacrifice". 2023. Available online: https://shorturl.at/fDGQ8 (accessed on 20 December 2023).
52. Miguel-Alonso, I.; Rodriguez-Garcia, B.; Checa, D.; Bustillo, A. Countering the Novelty Effect: A Tutorial for Immersive Virtual Reality Learning Environments. *Appl. Sci.* **2023**, *13*, 593. [CrossRef]
53. Cao, Y.; Ng, G.-W.; Ye, S.-S. Design and Evaluation for Immersive Virtual Reality Learning Environment: A Systematic Literature Review. *Sustainability* **2023**, *15*, 1964. [CrossRef]
54. Lei, Z.; Xie, Y.; Ling, S.; Pastor, A.; Wang, J.; Dong, J.; Le Callet, P. Multi-Modal Aesthetic Assessment for Mobile Gaming Image. In Proceedings of the 2021 IEEE 23rd International Workshop on Multimedia Signal Processing (MMSP), Tampere, Finland, 6–8 October 2021; pp. 1–5.
55. Israel, K.; Zerres, C.; Tscheulin, D.K. Virtual Realit—Substitute for a Real Experience? The Role of User Motivation, Expectations and Experience Type. *Int. J. Innov. Technol. Manag.* **2022**, *20*, 2350018. [CrossRef]
56. Martinez, K.; Menéndez-Menéndez, M.I.; Bustillo, A. A New Measure for Serious Games Evaluation: Gaming Educational Balanced (GEB) Model. *Appl. Sci.* **2022**, *12*, 11757. [CrossRef]