2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)

Player Ready? A Model for Evaluating VR Game Interfaces

Ralf Bierig
Department of Computer Science
Maynooth University
Maynooth, Ireland
ralf.bierig@mu.ie

Damien Mac Namara
Fury VR Ltd.
Limerick, Ireland
damienmacnamara@virtualfury.com

***Abstract*—We report on a study investigating the experience and usability of 3D interfaces in virtual reality games. To aid analysis of player behaviour performing simple tasks, a model is proposed which categorises player actions as discover, interact or move. For certain interfaces, movement caused unnecessary overhead as players must travel to the interface before discovery or interaction can take place. Additionally, menu items outside the players field of view caused task failures. The analysis converges with five heuristics for menu design in virtual reality games which we believe will aid developers in the field.**

***Index Terms*—virtual reality, menu interaction model, usability, user study**

I. Introduction

Virtual Reality (VR) applications have inherent usability issues which need to be addressed before mainstream adoption of VR can take place. By extension, VR games can suffer from similar usability issues which may deter gamers who currently play on traditional desktop environments. Tanriverdi et al. suggest that such interfaces tend to be three-dimensional in nature and more complex when compared to their desktop counterparts [1]. Indeed, previous studies evaluating game-playing in virtual versus desktop environments have shown that players tend to perform better in non-immersive (desktop) environments. [2], [3]. Interestingly, work by Pallavicini et al. suggests that players find VR games more appealing than their desktop counterparts [4]. In order to better the VR gaming experience, we believe a robust model is needed to understand how players perceive VR interfaces. Previous work by [5], had suggested a three-category model for describing VR actions but with the recent improvements in head-mounted displays, a deeper analysis of VR players is required for improved game design. In our DIM model we also present three categories of player action but further measure player efficiencies in each category.

The motivation for this research is twofold. Firstly, we observed real world usability issues in many VR games from simply watching players in our VR arcade (Fury VR) over a three and a half year period. During this time, we have administered thousands of gaming sessions. These sessions are mostly multiplayer in nature and span a number of genres but are predominately first-person-shooter (FPS). Players in our arcade are given remote instruction via TeamSpeak on how to play the game and host private servers. Although certain (but not all) games have onboarding mechanisms such as tutorials, these often require additional menu interactions to activate, or worse, still can be activated accidentally when trying to complete another task. We found significant issues with the discoverability (and general practicality) of the in-game menu systems present in these games, sometimes taking several of the initial minutes to instruct the customers on how to perform basic game tasks. We were unclear regarding the practicability of letting customers 'self-discover' the game mechanics themselves but found that this often lead to frustration when several interactions failed to produce the expected results. Our resulting experiment is an initial attempt to formalize the process we carry out on a daily basis in order to better understand the shortcomings of VR game interfaces. We knew that overcoming such usability issues would benefit casual players of VR games and help these players translate from a desktop gaming environment to a fully immersive virtual environment (VE). Even in well developed VR games with excellent game-play, we noted that cumbersome menu interfaces can detract from the gaming experience. Secondly, researchers and practitioners in VR usability would benefit from a formal way of transcribing VR gaming sessions for further analysis. We therefore present our DIM model in section III (also see figure 1) as a means of categorizing player actions and therefore better the understanding of VR gaming sessions. We specifically define three categories of player action: discovery, interaction and movement. For each category we analyze the players efficiency culminating in five heuristics to aid developers.

II. Related Work

Previous studies have shown notable differences between virtual and traditional desktop environments. Users in a virtual environment for example may not be able to see objects in the physical world, may suffer fatigue or become disorientated. [6], [7]. In a recent study comparing desktop and VR interfaces for product selection, [8] suggests that users feel a greater sense of immersion and presence in VR but interface improvements are still needed. Some key differences between virtual

978-1-7281-5965-2/20/$31.00 ©2020 IEEE

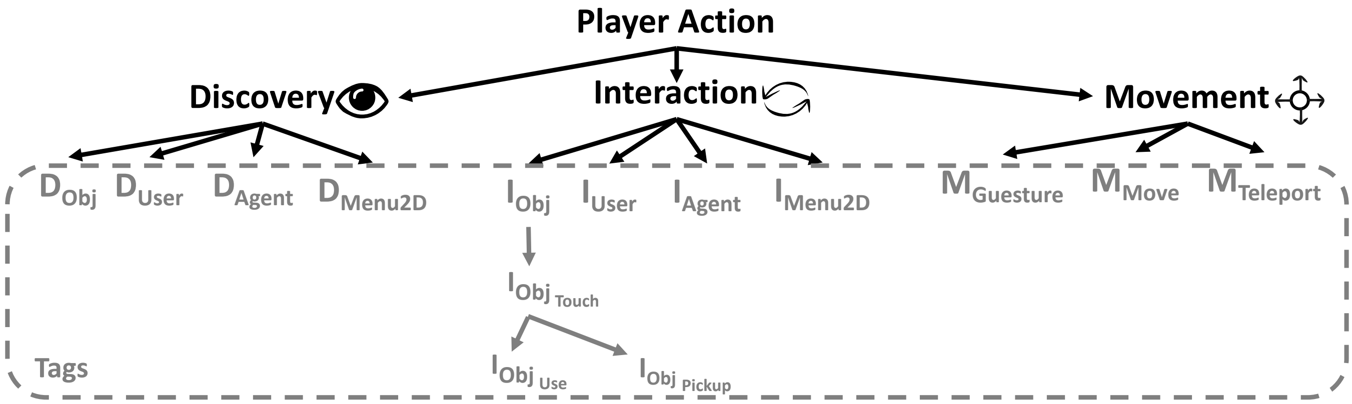

Fig. 1. The DIM model organizes VR player actions into the abstract categories of discovery (D), interaction (I) and movement (M) while providing optional, specific tags for refinement.

and desktop interfaces are highlighted by [1], most notably that VR interfaces tend to be 3D in contrast to 2D conventional screens. This may seem an obvious contrast, however 2D interfaces are often represented in VE's and utilized through well known approaches such as ray-casting [9]. With recent advances in home entertainment based Head Mounted Display (HMD) technology, some researchers have focused on natural user interfaces (NUI's) as a depth-based solution for multi-level menu selection and manipulation. In a study using the Leap Motion controller, [10] found varying results, as the hand interaction occluded the vision of participants when trying to read the menu. With the recent availability of the HTC Vive [1] and Oculus Rift [2], developers have again favored the use of a wand-type device over NUI's. In addition, with the advent of Room-Scale tracking, remoteness of menu components may be overcome with in-game teleportation, however older ray-casting techniques are still being used in parallel with newer approaches.

A. Game Interfaces

While VR applications continue to grow in niche areas, we have chosen to focus on VR game interfaces for our study, which tend to have a broader audience and greater appeal when compared to their traditional desktop counterparts [4]. Adams and Ernest describe genre delineation of video games as centered on the game play interaction [11]. It is widely acknowledged that a single game may belong to several genres, with varying degree [12]. In previous work [13] chose the first-person shooter genre due to the complex performance-intensive context of their study while evaluating the interaction fidelity of VR interfaces. Unlike [13] however, who aimed on locomotion aspects, our focus is on menu item interaction which is common across a wider variety of genres. In an attempt to abstract away from individual games we decided to instead investigate popular genres using the Steam VR game database.

[1]https://www.vive.com; accessed: April 27, 2020
[2]https://www.oculus.com/rift; accessed: April 27, 2020

B. Evaluation of Virtual Environments

In this paper we focus solely on 3D interfaces which we believe are better aligned with the immersive nature of VE's. In addition, 3D interfaces are likely more complex than 2D counterparts, requiring more study. In evaluating virtual environments (VE), modern evaluation techniques may not be sufficient for game interfaces. For example, playability has been a suggested extension to UX by measuring stimulation and entertainment value [14]. Some researchers suggest that traditional evaluation methods such as Nielson's heuristics are not entirely suitable for VE's and have put forward their own set of heuristics or extended the original [15]–[17]. Qualitative methods of evaluation are often employed [18]. In a review of literature pertaining to VE's, [3] found that several methods developed for 2D evaluation have been extended to support VE such as walkthrough's. observations, interviews and post-hoc questionnaires.

III. The DIM Model

The DIM model was created to help analyze player actions in virtual reality games. For each player we observed, we placed their actions into one of three categories: *Discovery*, *Interaction* or *Movement*. We drew on our expertise of observing players in our virtual reality arcade (Fury VR) over a period of three and a half years to create the three categories. Players in our arcade are monitored remotely and guided through their games step by step using spoken instruction. We know from thousands of gaming sessions that players could not always find an item or menu, were not always able to interact with them once discovered or did not know how to teleport in a VE. Discovery actions are therefore exploratory actions such as moving the HMD to look around a room. The interaction category relates to object use or manipulation while movement actions relate to physical or virtual movement.

We also wanted to be more specific when coding our gaming sessions so we created 'tags' in each category which describe the player action in greater detail. Effectively, each gaming session in our study was mapped to a corresponding table of tags. For each unit of time (1 second) we selected the tag which best fit the observed player action. The model therefore distinguishes between differing subsequent interactions by assigning a minimum unit of time (1 second) to each identifiable interaction. The gameplay analysis in section V is therefore based on a second-by-second transcription, even though the model will also work with high-frequency data collection and automatic tagging (e.g. through a logging feature).

We found that players could discover and interact with objects, users, agents (artificial players controlled by an AI), and two-dimensional menus (Menu2D). For movement, the player can gesture, physically move or teleport (move in the VE without physically moving). The DIM model and the tags are shown in figure 1. We found these fourteen tags were sufficient for the purposes of our study.

A primary requirement is that our model is not hardware dependent: The tags do not express actions in terms of button-pressing. Abstracting away from hardware allows for a greater scope across varying HMD's. It is also possible that each tag can be chained together to describe a particular sequence or pattern of actions. We explore such patterns using a Markov model in Section V.

While the primary aim of the model is to organise our understanding of complex menus that occupy a VE, the actions that are described by the model also cover a wide range of player actions that are generally possible in a VE. This means, the model can also be used to describe other activities, such as player interactions (e.g. team communication in a collaborative game) or player-agent interactions (e.g. combat situations).

IV. Methodology and Analysis

We applied a 2x2 within-subject full-factorial study design with tasks and game genres as its two independent variables. At the time of the study the most popular genres on Steam VR were 'action' (991 games) and 'casual' (967 games). Since games may be associated with more than one genre, we used Steam's ordering of the genre tags to determine its primary genre. For this study, we selected two candidate games for each genre — four games in total — based on every-day experience from Fury VR, the currently largest VR arcade in Ireland.

The games were predominately FPS titles that appear on the top-ten list in the Steam library and span VR paintball, VR dodge ball, VR archery and a laboratory environment. Each game had a highly developed 3D menu system as a key requirement. In total we were able to study six different types of 3D menu interface from the pool of games selected.

A. VR Tasks

We applied two abstract tasks that are frequently performed in VR gaming environments irrespective of game or game genre:

- *Start-a-Game:* Participants need to start either a single or a multi-player game on any game level. They may perform any interactions required to get from the initial Steam[3] starting screen to the initial game level.
- *Exit-a-Game:* Participants need to leave the game level and return to the main lobby by means of any combination of interactions they choose.

These general tasks allowed us to learn as much as possible about how players interacted with the menu interface without being too specific about any individual game.

B. Participants

For our study, we sampled participants (players) with prior 2D gaming experience but no experience on room-scale VR headsets. We recruited eight people — 6 male and 2 female, average age 27.8 years — with the incentive of a one-hour voucher for the VR arcade. Based on a self-reporting questionnaire, players never used an HTC Vive nor an Oculus Rift before but may have had other mobile VR experiences. They also reported their gaming expertise [19], on a 10-point Likert-scale where they achieved an average score of 6.9.

C. Procedure

We installed an observation laboratory at the arcade using a HTC Vive VR headset with two controllers in a room-scale configuration. Players were welcomed and signed both an information sheet and an ethical consent form. They were informed that they could cancel the session at any time. Players, wearing a HTC Vive VR headset, then took part in a short 5-10 min training session to learn the handling of controllers and the most common navigational gestures. Based on the 2x2 within-subject grouping, each player performed two tasks (one from each type) in two genres (each with a different game, selected in random order). Each game in each genre pool had an equal chance to be selected. The ordering of genres and tasks was counterbalanced to limit possible learning and fatiguing effects. Tasks were dictated to the player by a researcher via the audio headset when situated in a game. We employed a think-aloud protocol by encouraging players to verbalize their intentions and interactions. We recorded their think-aloud audio and their point-of-view video for both study tasks from a remote monitor. After the second task, the VR headset was removed and players were briefly interviewed about their satisfaction with the menu interfaces in each of the tasks using an SUS questionnaire. [20]. Afterwards, the players received a one-hour VR arcade voucher.

D. Menu Types

We analyzed six VR menu interfaces of varying composition: Hidden, unreachable, door, telportable, wearable and floating. The hidden menu interface was not visible to the player until they pressed a button on the controller. The unreachable menu interface could not be reached by the player and instead a projectile was fired into the menu to trigger its event. The door menu interface was a standard three dimensional door which could be pushed or pulled open. The teleportable interface is the only menu interface which could be teleported into in order to activate it, therefore keeping the player well within the flow of the game. The wearable interface was worn by the player on either hand but was

[3]https://store.steampowered.com, last access: April 27, 2020

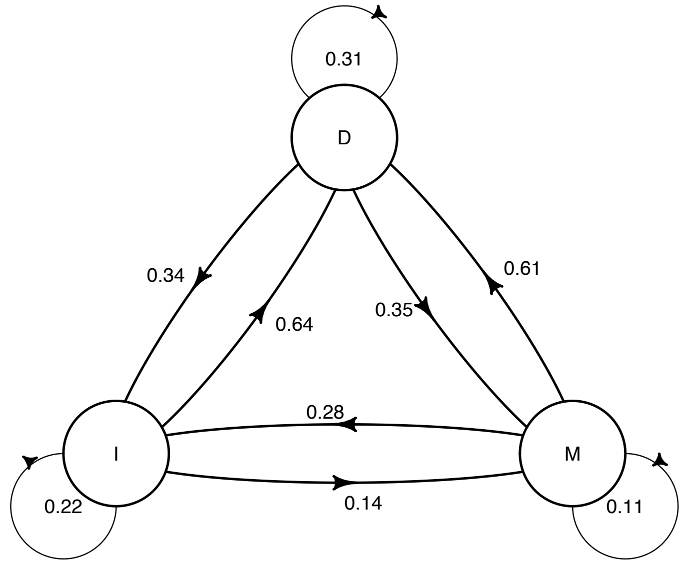

Fig. 2. Probability transitions between discovery (D), interaction (I) and movement (M).

generally outside their field of view. Finally, the floating interface required the player to physically move or teleport before interaction was possible. As players could choose between menu interfaces in some games, certain interfaces were selected more than others by our 8 players.

V. Results

A. Analysis of DIM Categories

We recorded approximately two hours of combined video and audio as raw observational data from the sixteen VR sessions — two sessions per player. Recordings were carefully annotated by two researchers following the DIM model, as described in the previous section, and additionally cross-verified by a research student. Each player action was therefore mapped to one of the fourteen tags together with its duration in full seconds. We only transcribed player actions that were a) at least one second long and b) singular (no parallel actions at any time). This was necessary to make manual transcription possible in the absence of automatic VR logging tools. This transcribed dataset consisted of a total of 612 annotations — an average of $\sim$40 annotated player actions per session.

We noted that the distribution of DIM categories in the dataset suggests a possible priority. Our participants spent most of their time discovering their surroundings and objects (53.2%), followed by interacting with the environment (27.9%) followed by moving in the environment (18.9%). We then looked at the behavior of players and how they moved in between the three DIM categories across all game genres and all tasks. Figure 2 shows this in the form of transition probabilities. Again, we found discovery to be a state where participants most likely remained (probability of 0.31), followed by interaction (probability of 0.22) and very little in movement (probability of 0.11). Discovery was also the category to which players most likely returned to (0.64 probability to return from interaction and 0.61 probability to return from movement). This has to be seen in the context that the model currently does not support a 'nothing-state' detecting players' idleness. Such a state could be tackled with the addition of eye tracking or intervention during think-aloud. The most interesting finding in the transitional behavior of players is that movement actions were hardly maintained (0.11 probability) and players frequently reverted back to the other two categories — discovery with 0.61 probability and interaction with 0.28 probability. We affirm in the next section that the transitory nature of movement actions may also be an impediment to menu interaction.

B. Measurements

We evaluated the efficiencies of six VR menu interfaces by applying our DIM model to each think aloud session. We measure the total task time and the time spent discovering, interacting and moving around the VE. We allocated a threshold time to each category based on our best estimations from the arcade. Considering our participants were new to VR gaming, we allocated a one minute threshold to discover the interface and a further thirty-second threshold for interaction. If a player took longer than these times, we made an observation on what we believed was the reason for this excess time. Three of the interfaces required either physical or teleport movement in order to reach the interface. Players opted to teleport in six of our sessions. In these cases we give the player one minute to learn how to teleport before flagging for further analysis. Satisfaction was measured between each game using a short Likert scale questionnaire, which was a miniaturized version of the popular Systems Usability Scale (SUS) [20]. We also note a binary measure of effectiveness: either the player completes a task or they did not.

C. Usability

With regards to the effectiveness, only the hidden menu had failures, as two players did not complete their tasks. The reason was twofold, firstly, the players were not able to discover the hidden menu and secondly, the players were distracted by nearby objects which they believed to be interfaces. These reasons are expanded on in our heuristics. Menu interface efficiency was generally poor with only one player completing an entire task in under one minute. For the reminder of the tasks, nine took between one and five minutes, four took more than five minutes to complete, while two tasks took longer than ten minutes to complete. Finally, satisfaction was again poor with a reported SUS score of 57.8, this is well below the SUS average of 68. Out of the sixteen sessions, players exceeded their discovery and interaction thresholds four times respectively while movement thresholds were exceeded three times. These formed the basis of our heuristics (see Section 6) A summary of observations issues and failures are shown in Table I and Table II.

VI. Suggested Heuristics

We present five heuristics based on the efficiencies discovered in each DIM category and the task failures we observed.

TABLE I
SUMMARY OF OBSERVATIONS

Type	Observation	Category
Hidden	Menu button not found (O1)	Discovery
Unreachable	Could not fire projectile (O2)	Interaction
	Accidental Menu Trigger (O3)	Interaction
Door	Poor placement of Menu (O4)	Discovery
	Could not move to (O5)	Movement
	Open gesture not complete (O6)	Interaction
Teleportable	Distracted by nearby menu (O7)	Discovery
	Could not move to (O8)	Movement
Wearable	Wearable not noticed (09)	Discovery
	Got lost in sub menus (O10)	Interaction
Floating	Could not move to (O11)	Movement

TABLE II
SUMMARY OF FAILURES

Type	Observation	Category
Hidden	Menu button not found (F1)	Discovery
	Weak menu affordance (F2)	Interaction

We omit rules which are more akin to generic usability. For instance we do not include our observation on the weaknesses of a poor affordance as this is a given for successful user interface design.

1) Keep important menu items within the players field of view. (O1, O4, O9, F1). This applies particularly to hidden and wearable menu interfaces. As players tend to spend more time discovering their environment, the existence of an invisible menu may be an afterthought as other visible options are exhausted. When given a choice between a visible and a wearable menu, 3 out of 4 times the player used the visible menu even though the wearable menu had more functionality. Further, menus which rely on the use of hidden projectiles are dependent on the player discovering the projectile, which in our experiment were worn on the players back and outside their field of view.
2) Give full functionality to items with strong affordance. (O1, O6, F2). In the hidden menu task, the environment was surrounded by a curtain which had a strong perceived affordance. Unsurprisingly, players tried to pull apart the curtain in order to advance beyond the border. This resulted in the longest interaction time of any interface. Weak functionality is also a problem as we found the door menu would not trigger unless the player extended their arm in a certain way. However a slight pull action did open the door causing some players to believe the door only opened slightly and had no other function.
3) Prevent accidental event triggering with smart menu placement. (O3, O7, O10). Placing menu items too close to a player's initial spawn point may cause accidental event triggering. Likewise a nearby menu interface can detract from a more critical interface used to progress the game. We observed that movable objects placed too close to a players starting position can be touched / knocked over accidentally, thus triggering their events. For one task this constantly activated the tutorial and impeded the player's use of the unreachable menu. In addition, another player believed a nearby character customization menu would progress them through to the main game lobby and thus ignored the telportable menu interface nearby.
4) Movement can cause unnecessary overhead (O5, O8, 011). Placing a menu item outside of the player's walkable play-space may create an additional overhead since the player may have no knowledge of teleportation. The effectiveness of the menu interface will therefore depend on the teleportation method. This was evident for the door, telportable and floating types as some players could not reach the menu interface in order to interact with them.
5) Projectiles can cause unnecessary overhead (O2). As with movement, the dependency on firing projectiles in order to interact with a menu interface requires the player to have foreknowledge of how to fire projectiles. Again, the effectiveness of the menu interface is dependent on the projectile firing mechanism.

A. *Limitations*

The DIM model is more than adequate to represent the games we have included in this study, however additional studies are needed to test the robustness of the model. As mentioned, the model may improve with the use of eye-tracking to counter the nothing state and VR controller event logging to help with granularity. We observed that some players physically moved while attempting a software teleport because they thought that the teleportation beam needed to be active during physical movement. In this instance we must decide which action is more dominant within the one second frame. Finally, the thresholds we applied to each action category are more exploratory than hard and fast rules. A larger study with more participants would yield improved thresholds.

VII. CONCLUSIONS AND FUTURE WORK

Our study suggests serious deficiencies with three dimensional VR menu's in terms of their efficiency and player satisfaction. Using our DIM model we were able to analyze six menu types under three categories of player action. We found that players spend most of their time discovering their environment and that movement tends to be a transitory state between object interaction and discovery. By assigning thresholds to each action category, we further investigated issues that arose within each category, culminating in five heuristics for VR developers. In particular we caution developers on the use of interfaces dependent on either teleportation or projectiles. Finally, the implementation of hidden menu interfaces could pose discoverability issues. Surprisingly however, when hidden menus were discovered, efficiency was generally good.

Our next study will evaluate a larger pool of diverse menu types with 30+ players. The addition of a hardware logger would give us finer granularity and would help to streamline the video coding process thus allowing us to scale evaluation. We feel these improvements would help to further the robustness of the model.

ACKNOWLEDGEMENTS

We thank Dominic Carr from the Galway-Mayo Institute of Technology, Ireland for his contributions to the initial idea for this paper.

REFERENCES

[1] V. Tanriverdi and R. J. Jacob, "Vrid: a design model and methodology for developing virtual reality interfaces," in *Proceedings of the ACM symposium on Virtual reality software and technology*. ACM, 2001, pp. 175–182.

[2] C. Swindells, B. A. Po, I. Hajshirmohammadi, B. Corrie, J. C. Dill, B. D. Fisher, and K. S. Booth, "Comparing cave, wall, and desktop displays for navigation and wayfinding in complex 3d models," in *Computer Graphics International, 2004. Proceedings*. IEEE, 2004, pp. 420–427.

[3] B. S. Santos, P. Dias, A. Pimentel, J.-W. Baggerman, C. Ferreira, S. Silva, and J. Madeira, "Head-mounted display versus desktop for 3d navigation in virtual reality: a user study," *Multimedia Tools and Applications*, vol. 41, no. 1, p. 161, 2009.

[4] F. Pallavicini, A. Ferrari, A. Zini, G. Garcea, A. Zanacchi, G. Barone, and F. Mantovani, "What distinguishes a traditional gaming experience from one in virtual reality? an exploratory study," in *International Conference on Applied Human Factors and Ergonomics*. Springer, 2017, pp. 225–231.

[5] J. Kjeldskov, "Interaction: Full and partial immersive virtual reality displays," *Proceedings of IRIS24, University of Bergen, Bergen*, pp. 587–600, 2001.

[6] D. A. Bowman, E. Kruijff, J. J. LaViola Jr, and I. Poupyrev, "An introduction to 3-d user interface design," *Presence: Teleoperators and virtual environments*, vol. 10, no. 1, pp. 96–108, 2001.

[7] D. Bowman, E. Kruijff, J. J. LaViola Jr, and I. P. Poupyrev, *3D User Interfaces: Theory and Practice, CourseSmart eTextbook*. Addison-Wesley, 2004.

[8] A. Schnack, M. J. Wright, and J. L. Holdershaw, "Immersive virtual reality technology in a three-dimensional virtual simulated store: Investigating telepresence and usability," *Food Research International*, 2018.

[9] D. A. Bowman and L. F. Hodges, "An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments," in *Proceedings of the 1997 symposium on Interactive 3D graphics*. ACM, 1997, pp. 35–ff.

[10] M. M. Davis, J. L. Gabbard, D. A. Bowman, and D. Gracanin, "Depth-based 3d gesture multi-level radial menu for virtual object manipulation," in *Virtual Reality (VR), 2016 IEEE*. IEEE, 2016, pp. 169–170.

[11] E. Adams, *Fundamentals of game design*. Pearson Education, 2014.

[12] T. H. Apperley, "Genre and game studies: Toward a critical approach to video game genres," *Simulation & Gaming*, vol. 37, no. 1, pp. 6–23, 2006.

[13] R. P. McMahan, D. A. Bowman, D. J. Zielinski, and R. B. Brady, "Evaluating display fidelity and interaction fidelity in a virtual reality game," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 4, pp. 626–633, 2012.

[14] J. L. G. Sánchez, F. L. G. Vela, F. M. Simarro, and N. Padilla-Zea, "Playability: analysing user experience in video games," *Behaviour & Information Technology*, vol. 31, no. 10, pp. 1033–1054, 2012.

[15] A. G. Sutcliffe and K. D. Kaur, "Evaluating the usability of virtual reality user interfaces," *Behaviour & Information Technology*, vol. 19, no. 6, pp. 415–426, 2000.

[16] A. Sutcliffe and B. Gault, "Heuristic evaluation of virtual reality applications," *Interacting with computers*, vol. 16, no. 4, pp. 831–849, 2004.

[17] R. A. Blade and M. L. Padgett, "Virtual environments standards and terminology," *Handbook of virtual environments: Design, implementation, and applications*, pp. 15–27, 2002.

[18] D. Geszten, B. P. Hámornik, A. Komlodi, K. Hercegfi, B. Szabó, and A. Young, "Qualitative analysis of user experience in a 3d virtual environment," in *Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community*. American Society for Information Science, 2015, p. 124.

[19] M. J. Schuemie, P. Van Der Straaten, M. Krijn, and C. A. Van Der Mast, "Research on presence in virtual reality: A survey," *CyberPsychology & Behavior*, vol. 4, no. 2, pp. 183–201, 2001.

[20] J. Brooke, "SUS: A 'quick and dirty' usability scale," in *Usability Evaluation in Industry*, P. W. Jordan, B. Thomas, B. A. Weerdmeester, and A. L. McClelland, Eds. London: Taylor and Francis, 1996, ch. 21, pp. 266–290.