

Article

# An Approach to Assess the Impact of Tutorials in Video Games

Dario Benvenuti <sup>\*</sup>, Lauren S. Ferro, Andrea Marrella  and Tiziana Catarci

Department of Computer, Control and Management Engineering, Sapienza University of Rome, Via Ariosto 25, 00185 Rome, Italy

<sup>\*</sup> Correspondence: d.benvenuti@diag.uniroma1.it

**Abstract:** Video games are an established medium that provides interactive entertainment beyond pure enjoyment in many contexts. Game designers create dedicated tutorials to teach players the game mechanisms and rules, such as the conventions for interaction, control schemes, core game mechanics, etc. While effective tutorial design is considered a crucial aspect to support this learning process, the existing literature approaches focus on designing ad hoc tutorials for specific game genres rather than investigating the impact of different tutorial styles on game learnability and player engagement. In this paper, we tackle this challenge by presenting a general-purpose approach aimed at supporting game designers in the identification of the most suitable tutorial style for a specific genre of video games. The approach is evaluated in the context of a simple first-person shooter (FPS) mainstream video game built by the authors through a controlled comparative user experiment involving 46 players.

**Keywords:** video game; tutorial; player psychology



**Citation:** Benvenuti, D.; Ferro, L.S.; Marrella, A.; Catarci, T. An Approach to Assess the Impact of Tutorials in Video Games. *Informatics* **2023**, *10*, 6. <https://doi.org/10.3390/informatics10010006>

Academic Editor: Antony Bryant

Received: 21 October 2022

Revised: 4 January 2023

Accepted: 6 January 2023

Published: 11 January 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Nowadays, video games are accepted as a part of mainstream culture and are considered an established medium that affords many different interactive experiences, often pushing the boundaries of traditional (and often linear) interaction design [1]. Moreover, since video games have become easier to develop, they are increasingly adopted in domains beyond entertainment, such as cybersecurity [2,3] and education [4].

Unlike application software, video games provide complex interactive elements, such as branching narratives, environmental exploration, enemies and quests, which are often intended to be experienced together and in real-time. Combining these elements allows players to determine their own stories at their own pace, experiencing the consequences of their actions. By doing so, players can reflect, understand where they went wrong and try again [5]. Consequently, video games train a systematic way of thinking that allows players to learn by gaming [6]. Considering this, the ways in which players must learn how to play need more practical information and a proper explanation of how to do so. Game designers often use *tutorials* to support this learning process.

To date, modern video games employ a wide variety of tutorial styles, ranging from gently easing players into the experience to forcing them to learn via trial and error. Even in the context of the same video game, a tutorial can be developed differently depending on the platform on which the game is experienced (e.g., PlayStation, PC, etc.) and the peripheral hardware used (e.g., gamepads, keyboard, etc.).

While the effective tutorial design is considered a crucial aspect of teaching the game mechanisms and retaining new players, more is needed to know how different tutorial styles may affect game learnability and player engagement [7]. As a result, game designers must rely on their own experience and intuition to design tutorials, with the risk that new players will struggle to grasp the most rudimentary game concepts.

To mitigate the above issue, in this paper we present a four-step approach aimed at supporting game designers in incorporating a tutorial within their game. Specifically:

- The core of the approach is a *classification framework*, which enables us to categorize the tutorials of some selected video games belonging to a specific genre based on a set of gaming features (e.g., player's freedom degree, context-sensitivity of the tutorial, etc.).
- The second step consists of developing a *single-level video game* equipped with a tutorial realized in different variants in a way that each variant reproduces the mechanisms underlying any of the most common gaming features found in the previous step.
- The third step concerns the enactment of a *user experiment* involving different groups of players, instructing them to play the main level of the (novel) video game in the absence of any tutorial or after having completed (only) one of the tutorial variants.
- The fourth step consists of administering *dedicated questionnaires* to the players involved in the user experiment. The target is to understand which tutorial variant was the most effective for the comprehension of the game mechanisms, combining players' evaluation with empirical measures collected during the user experiment.

The hypothesis behind our approach is that a deeper understanding of the gaming features underlying the tutorials of video games of a certain genre could allow game designers to create more learnable and engaging video games. To support this hypothesis, we have applied our approach to look at how video game companies that produce AAA first person shooter (FPS) titles have built the tutorial for their games. Then, we identified two main categories of tutorial types and developed a simple FPS prototype in which those tutorials are implemented. Finally, we performed a between-subjects experiment with 46 participants who played our prototype in three different settings ([first OR second OR no] tutorial + main level). This has allowed us to test the effectiveness of our approach and, on the other hand, to wrap up some interesting insights concerning tutorials.

The rest of the paper is organised as follows. In Section 2, we analyze background and related work. In Section 3, we discuss our approach to categorize tutorials based on the gaming features, while Section 4 discusses the implementation and evaluation of our approach in the FPS genre. Finally, in Section 5, we discuss lessons learned, conclusions and recommendations for future work.

## 2. Background and Related Work

When a player starts interacting within a game world, tutorials are typically the first part of the game that new players encounter, representing the so-called "*first-time gaming experience*". In this stage, players become aware of how to interact with the game world and the consequences of these interactions. Thus, from a learning perspective, tutorials are the natural first point of instruction for a video game.

Tutorials can exist in many forms. They can be integrated into the gameplay or can be completely separate and optional. Some games may offer a basic mandatory tutorial and optional advanced training, combining the different modalities. While the instructional nature of tutorials can easily lead to moments of boredom, their absence can lead to frustration at later parts of the game because players are unsure how to play and must learn through trial and error. Both cases can cause an increase in player drop-out rates, which can impact developers, especially in some business models, such as "free to play", where the revenue comes from users who continue to play (and often buy micro-transactions).

It is worth noting that tutorial instructions can also be provided during ordinary gameplay, when certain game mechanics are introduced for the first time or when the player gains new items or abilities [8]. In recent games, visual cues are used to inform players what can be interacted with, where to look and where to move within the game world [5,9]. Another important consideration is how tutorials may impact the replay value of a video game [10]. When a player decides to replay a game, assuming she has already become accustomed to the controls, the effectiveness of tutorials becomes relevant when the game includes completely different play styles depending on the selected class mechanism or character type. Players are often motivated by other factors, such as completing missing achievements, exploring different parts of a branching narrative to reveal alternative

endings [11], or unlocking other content (e.g., items, etc.). In some cases, replay value may also be driven by a player's desire to improve their skills [12].

### 2.1. Previous Works on Assessing Tutorials

Previous studies have explored ways to measure the effectiveness/enjoyment of tutorials (e.g., [13–15]). For example, a study by Andersen et al. [7] measured how much game complexity affects the perceived outcome of tutorials. Moreover, they defined a binary heuristic to analyze tutorials based on the techniques used to teach players. In addition, Green et al. [16] used the classification framework delivered in [7] and a game jam framework to build an Artificial Intelligence (AI) tool capable of developing tutorials for video games. Other studies investigated ways to assess tutorials through questionnaires and/or heuristics [13–15,17]. Ballew et al. [13] attempted to validate design heuristics presented by Federoff [18] by getting participants (n = 133) to play either a low- or high-rated PC video games. Their analysis revealed that 75% of the design heuristics were more applicable to the high-rated games than the low-rated ones. On the other hand, Johnson et al. [14] presented a detailed exploratory and confirmatory factor analyses of both the player satisfaction (PENS) [19] and the game experience questionnaire (GEQ) [20] leading to the development of revised models. Klimmt et al. [15] investigated the impact of game difficulty and player performance on game enjoyment and found that they are not fully in line with predictions derived from flow and attribution theory, suggesting that players change their view based on their own performance, which impacts their enjoyment, and that players also strategically switch between different sources of fun for maintaining a positive gaming experience. Lastly, Sweetser et al. [17] presented the *GameFlow* model for evaluating player enjoyment in video games, which was similar to the work by Johnson et al. [14] who explored criteria based on eight elements (concentration, challenge, skills, control, clear goals, feedback, immersion and social interaction) to successfully distinguish high-rated and low-rated video games and understand why a game could fail or succeed.

Tutorials are an essential element of an interactive experience. Players who are new to playing games or a particular genre will be overwhelmed by the key combinations or tactics used to play a game. The existing literature has explored the design and/or implementation of tutorials in video games (e.g., [7,16,21,22]). For example, Paras [22] conducted a study on his journey to making a game, which evolved into the development and testing of an effective in-game tutorial system that improves player performance without negatively affecting the experience of play. Two critical success factors are highlighted for the definition of a tutorial: the type of training and the timing of the training.

Many studies have explored how to create tutorials and embed them directly into the game design mechanics [8]. Other works (such as [23,24]) focused on providing guidelines for tutorials, considering the context in which they are delivered. Indeed, there is a vast difference between a tutorial that must be experienced within a virtual reality environment versus one that will be used in a first- or third-person player perspective. For example, both Frommel et al. [23] and Zaidi et al. [24] explored ways in which tutorials should be delivered within a virtual environment. Resulting from their studies, Frommel et al. concluded that the development and integration of context-sensitive tutorials elicited higher positive emotions, lower negative emotions and higher motivation. They further state that developers should not regard tutorials as separate introduction levels to a game but rather as part of the overall game experience. In this direction, Zaidi et al. concluded that the user-centered design approach for their tutorial greatly supported a player's ability to learn about the environment and subsequently understand the game. As a result, the player found the experience highly entertaining, engaging and usable.

### 2.2. Novelties

Compared with the existing works, the novelty of our approach lies in the systematic sequence of methodological steps required to assess the impact of tutorials for video games of a certain genre. Rather than providing a rigid set of guidelines consisting of pre-fixed

parameters, our intuition is that assessing tutorials requires a deep understanding of the essential gaming features underlying the mechanisms of a specific game genre. Only with this knowledge at hand is it possible to apply a suitable classification framework to identify the most effective tutorial types. In addition, by relying on a single-level video game realized ad hoc for the game genre of interest, we can concretely create tutorial variants that emulate the selected tutorial types and quantify their impact on the player's experience. This assessment performed "on the field" will support game developers in deciding whether to accept or reject the results they would have obtained by only applying the classification framework. In particular, using the dropouts index as a parameter to evaluate how effective a tutorial is makes this work very useful even from other perspectives of research (e.g., publishers who care about consistent incomes, developers that want users to experience their games fully and so on). This information also helps to inform game developers to improve the continued gameplay rates of players by reducing dropouts early on in the game experience due to bad tutorials or explanations.

### 3. Approach

In this section, we provide details on the main methodological step required by our general-purpose approach to assess the impact of different tutorial types.

After having selected a specific game genre of interest, we must identify a reasonable number of games that: (i) are representative of that genre; (ii) provide gameplay diversity; and (iii) cover a good time span. Each selected game can be evaluated against a classification framework that categorizes its in-game tutorials based on the features it provides, chosen from an extensive list. The anatomy of the proposed classification framework is presented in Section 3.1.

The approach also requires the realization of a game prototype (of the genre under analysis) equipped with a tutorial realized in different variants. Any variant will reproduce the most common combinations of gaming features found using the classification framework. The prototype must contain at least one game level after the tutorial to check if players can put into practice what they have learned. We notice that the proper development of the game prototype is the charge of the game designers. The details of a sample game prototype developed for the first-person shooter (FPS) genre are shown in Section 4.

Finally, conducting a comparative user experiment enables us to investigate players' performance, while interacting with the main level of the game prototype in the absence of any tutorial or after completing one of the implemented tutorial variants. The composition of the user groups should be defined in a between-subjects fashion, i.e., each user experiences only one tutorial variants, and their outcomes are then compared. On the other hand, given the nature of the experiment, the within-subject design would make it hard to understand the impact of the learning effect due to the testing of more tutorial variants, with the risk of biasing the results. The design of the questionnaires to collect the players' feedback and the analysis of the evaluation techniques for interpreting the results are discussed in Section 3.2.

#### 3.1. Classification Framework

We developed our classification framework relying on the literature frameworks of Andersen [7] and Paras [22], which provide a good mixture of technical and non-technical parameters to categorize tutorials. In addition, since video games have evolved substantially during the last decade, we added four parameters to modernize the classification, considering that the frameworks of Andersen and Paras were published in 2012 and 2006, respectively.

The following parameters were taken from the classification framework of Andersen [7]:

1. **Tutorial presence:** *true* if the game under analysis provides some in-game tutorial.
2. **Context sensitivity:** *true* if the instructions for a certain action or game mechanic are shown to the players only when they really need to use them during the gameplay.
3. **Freedom:** *true* if the players are provided with some freedom during the tutorial, e.g., they can make choices in the game.
4. **Availability of help:** *true* if, during the tutorial, the game understands when the player is in need of help and reacts to that through textual, visual, or graphical cues.

Then, from Paras [22] we selected the following parameters:

5. **Printed:** *true* if the game provides printed documentation with the game instructions.
6. **On-screen text:** *true* if the game delivers instructions to the players through text.
7. **Voice:** *true* if the game delivers instructions through voice.
8. **Video:** *true* if the game explains mechanics or commands employing dedicated videos, including cut scenes.
9. **Helping avatar:** *true* if an in-game assistant supports the player. In-game assistants range from a non-playable character (NPC) to fictitious characters speaking to the player through a phone call, etc. We notice that a helping avatar is considered part of the game world; i.e., it is not simply a voice or a text providing specific instructions.
10. **Controller diagram:** *true* if the game shows the players (on request) an image describing how commands are mapped to the input device.

Finally, we decided to include in the framework four additional parameters to capture trends that recently emerged in video games:

11. **Command scheme customization:** *true* if the game lets the player change the command scheme mapping.
12. **Skippable:** *true* if the tutorial can be completely skipped.
13. **Story integration:** *true* if the tutorial is integrated into the game's story.
14. **Practice tool presence:** *true* if the players are provided with a safe place to practice and get used to commands and game mechanics.

We excluded any parameter from our framework to capture the presence of third-party material made available to explain or tweak the game mechanics, e.g., online content, game mods (usually realized by fans), etc. Unless this content is requested or approved by the game developers, it is not something players can always rely upon in place of a tutorial.

### 3.2. Questionnaires Development and Evaluation Techniques

We developed two questionnaires to collect the users' feedback related to their gaming experience with the game prototype. Specifically, we mixed elements from psychology and human-computer interaction (HCI) [25–28]. We focused on the following points:

- **Demographic information (DI):** Users' experience with video games and with the particular genre in which the analysis will focus, along with age and gender.
- **Simple debriefing (SD):** It is interesting to know if the user usually skips tutorials in video games when possible and if, just after having played the game, s/he felt like the tutorial was helpful.
- **Learnability of the game (LG):** To measure if the game is easier to learn when users are presented with one particular tutorial.
- **How much users felt loaded (UL):** Some questions to understand the amount of mental workload required to the users while playing the game.
- **Performance self evaluation (PE):** To measure how users felt successful while playing the game.
- **User experience (UX):** Some thoughts from the users about their experience while playing the game.

We split the questions into two sub-questionnaires: (i) a pre-questionnaire to be submitted to the users before they start playing the game; and (ii) a post-questionnaire to be administered right after users finish playing the game. In the post-questionnaire, we

decided to pair our questions with two well-known and established questionnaires: the user experience (UX) questionnaire [29] and the NASA task load index (TLX) [30]. The UX questionnaire focuses on measuring usability aspects (efficiency, perspicuity, dependability) and user experience aspects (originality, stimulation) during the interaction with a user interface. The TLX is a multi-item questionnaire developed to measure perceived workload, i.e., the amount of effort users have to exert to perform a task on a user interface. Lastly, we decided to use log file analysis to obtain some empirical values of the participants' performance. For the sake of space, we will not include the complete list of questions in this paper, but we will look at the focus of each questionnaire:

1. **Pre-Questionnaire:** It was built to obtain demographic information on the participants and to understand their expertise with video games. The pre-questionnaire provides the following items (cf. DI):
  - (a) Which is an email address to which we can send the game? (*open answer*)
  - (b) What is your age? (*open answer*)
  - (c) Which is your gender? (*female/male/prefer not to say/other*)
  - (d) How much are you experienced with video games? (Likert scale from 1—*I never play games* to 5—*I always play games*)
  - (e) How much are you experienced with the genre of video games under analysis? (Likert scale from 1—*I never play this genre of games* to 5—*I always play this genre of games*)
  - (f) Please indicate your favorite video game's genre (*open answer, optional*).
  - (g) Consent data collection (*yes/no*).
2. **Post-Questionnaire:** It was built to understand how each tutorial variant can influence the player experience in the context of a certain game genre. The post-questionnaire provides the following items:
  - (a) Questions about learnability (cf. LG):
    - I have perfectly understood how to perform actions (with examples) for playing the game. (Likert scale from 0—*Never* to 3—*Always*)
    - The information provided throughout the game (with examples) are clear. (Likert scale from 0—*Never* to 3—*Always*)
  - (b) Debriefing questions (cf. SD):
    - Before playing the game, would you have skipped the tutorial if it was possible? (*yes/no*)
    - Please explain in a few lines the reason of your choice (*open answer, optional*).
    - After you have finished playing the game, do you think that the tutorial was useful? (*yes/no*).
    - Please explain in a few lines the reason for your choice (*open answer, optional*).
  - (c) NASA Task Load Index [30], in which every item is evaluated with a Likert scale ranging from 1—*Low* to 10—*High* (cf. UL):
    - How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.) to play the game?
    - How much pressure did you feel, due to the pace at which the game is set, or due to the setting of the game?
    - How successful do you think you were in playing the game?
    - How hard did you have to mentally work to accomplish your level of performance?
    - How frustrated (e.g., insecure, discouraged, irritated, stressed and annoyed) did you feel while playing the game?
  - (d) User eXperience Questionnaire [29]: It consists of a set of pairs of contrasting attributes. The attributes are at the opposite values on a Likert scale ranging

from (1) to (7). Users can express opinions on the game by choosing the value that most closely reflects their impression (cf. UX).

- (1) Annoying–(7) Enjoyable
- (1) Not understandable–(7) Understandable
- (1) Creative–(7) Dull
- (1) Valuable–(7) Inferior
- (1) Boring–(7) Exciting
- (1) Not interesting–(7) Interesting
- (1) Easy to learn–(7) Difficult to learn
- (1) Unpredictable–(7) Predictable
- (1) Fast–(7) Slow
- (1) Inventive–(7) Conventional
- (1) Obstructive–(7) Supportive
- (1) Good–(7) Bad
- (1) Complicated–(7) Easy
- (1) Unlikable–(7) Pleasing
- (1) Usual–(7) Leading edge
- (1) Unpleasant–(7) Pleasant
- (1) Secure–(7) Not secure
- (1) Motivating–(7) Demotivating
- (1) Meets expectations–(7) Does not meet expectations
- (1) Inefficient–(7) Efficient
- (1) Clear–(7) Confusing
- (1) Impractical–(7) Practical
- (1) Organized–(7) Cluttered
- (1) Attractive–(7) Unattractive
- (1) Friendly–(7) Unfriendly
- (1) Conservative–(7) Innovative

3. **File Logging:** to obtain empirical measures of users' performance directly from the game. This step could vary greatly depending on the genre of the video games on which the study focuses, but it should contain every parameter related to users' in-game actions that can be captured. In particular, we chose to log the following parameters:

- Time spent in the tutorial.
- Time spent in each level.
- Number of times the user tried each level.
- Number of deaths in each level (relevant only for some game genres).
- Number of deaths in the tutorial (relevant only for some game genres).
- Time stamps for important events.
- Numbers of enemies killed (relevant only for some game genres).
- Final result (defeat/victory/drop-out)
- Empirical measure of performance (dependent from the game prototype).

The administration of the post-questionnaire to each user involved in testing the game prototype equipped with no tutorial or with one of its implemented variants enables us to collect quantitative data to be assessed through a suitable evaluation technique. Given the comparative nature of the experiment and the presence of (at least) two independent testing groups (we are assuming a between-subjects design of the experiment), we believe that employing the ANOVA (analysis of variance) test is a suitable solution to determine if there is a statistically significant difference between the questionnaire results from the different groups [31]. Then, for each item of the questionnaire with such a difference, a two-sample *t*-test with a 95% confidence level can be applied to determine if the means of participants

using different tutorial variants differ. Finally, the level of statistical significance can be obtained by analyzing the resulting  $p$ -value, choosing 0.05 as the threshold value [28].

#### 4. Use Case

To evaluate our approach, we applied it to a selection of popular and highest-selling AAA FPSs. FPS is a video game genre centered on weapon-based combat from a first-person perspective [32], with the player experiencing the action through the eyes of the protagonist and controlling the player character in a three-dimensional space [33].

Specifically, we incorporated our approach in the FPS genre as follows: (i) we applied the classification framework defined in Section 3.1 against a selection of different in-game tutorials developed for 32 selected AAA FPSs. This allowed us to extract a parameter-based description of the two most common types of tutorial in this category of video games; (ii) we developed a prototype of a single-level FPS in three different versions: one with no tutorial at all and the other two equipped with the tutorial variants extracted from the previous step; (iii) we performed a controlled comparative experiment through online user testing. The experiment was realized by relying on a between-subjects design. We involved 46 users divided into three groups, so each group tested only one of the three versions of the FPS; and (iv) we collected feedback through pre- and post-questionnaires and kept track of the players' activities in the game using a dedicated logging mechanism.

In the following sections, we will look at each step in more detail.

##### 4.1. Applying the Classification Framework

The first step was to apply the classification framework to analyze already existing tutorials. We selected 32 tutorials from respective AAA FPS games based on their popularity and gameplay style. The list of selected games is shown in Table 1, together with their assessment against the classification framework.

We handpicked those titles from top-selling FPSs throughout the last two decades, mixing traditional FPSs with narrative-driven ones, such as the "Elder Scrolls", and niche titles such as "Mirror's Edge", a Parkour game. This allowed us to build both a vertical analysis of how tutorials evolved during these years, and a horizontal one, on how tutorials can vary when we move towards different, but very similar, genres of video games.

By looking at the results of the classification framework, we identified the two most frequent combinations of parameters/gaming features available in the analyzed tutorials:

1. The most popular one, with 10 entries in the table consisted of: presence of tutorial, context sensitivity, NOT freedom, help availability, text usage, voice usage, NOT video usage, helping avatar, NOT skippable, story integration, NOT practice tool, NOT printed, command scheme customization and controller diagram. This combination of gaming features is commonly found in narrative-driven FPSs, such as "Destiny".
2. Excluding the games belonging to the previous selection from the table, we are left with great diversity. Some video games, though, differ only for the evaluation of one or two parameters. Thus, we decided to select "Overwatch" (and its combination of gaming features) as representative of the games not belonging to the first selection. Therefore, the second combination consisted of: presence of tutorial, context sensitivity, freedom, help availability, text usage, voice usage, NOT video usage, helping avatar, skippable, NOT story integration, practice tool, NOT printed, command scheme customization and controller diagram. Games of this second group can be found by searching in the table for rows where the parameter "Practice Tool Presence" is true.

**Table 1.** Data gathered from 32 tutorials of FPSs available in the market.

Game	Tutorial Presence	Context Sensitivity	Freedom	Availability of Help	Text	Helping Avatar	Video	Voice	Skippable	Story Integration	Practice Tool	Printed	Command Scheme Customization	Controller Diagram
Overwatch	Y	Y	N	Y	Y	Y	N	Y	Y	N	Y	N	Y	Y
Apex Legends	Y	Y	Y	Y	Y	Y	N	Y	N	N	Y	N	Y	Y
Destiny	Y	Y	N	Y	Y	Y	N	Y	N	Y	N	N	Y	Y
Destiny 2	Y	Y	N	Y	Y	Y	N	Y	N	Y	N	N	Y	Y
COD MW2	Y	Y	N	N	Y	Y	N	Y	N	N	N	N	Y	Y
COD Ghosts	Y	Y	N	N	Y	Y	N	N	N	Y	N	N	Y	Y
COD BO2	Y	Y	N	N	Y	N	N	N	N	Y	N	N	Y	Y
Doom 2016(Normal)	Y	Y	N	N	Y	N	N	N	N	Y	N	N	Y	Y
Doom 2016(Hard)	N	Y	N	N	N	N	N	N	N	N	N	N	Y	Y
Mirror's Edge	Y	Y	N	Y	Y	Y	Y	Y	N	N	Y	N	Y	Y
Mirror's Edge Catalyst	Y	Y	N	Y	Y	Y	Y	Y	N	Y	Y	N	Y	Y
Bioshock	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N	N	Y	Y
Bioshock 2	Y	Y	Y	Y	Y	Y	Y	N	N	Y	N	N	Y	Y
Bioshock Infinite	Y	Y	Y	Y	Y	N	Y	N	N	Y	N	N	Y	Y
Portal	Y	Y	N	Y	Y	Y	N	Y	N	Y	N	N	Y	Y
Portal 2	Y	Y	N	Y	Y	Y	N	Y	N	Y	N	N	Y	Y
Spec Ops: the Line	Y	Y	N	N	Y	Y	N	Y	N	Y	N	N	Y	Y
Half-life	N	N	N	N	N	N	N	N	N	N	N	Y	Y	Y
Half-life 2	Y	Y	N	Y	Y	Y	N	Y	N	Y	N	N	Y	Y
Singularity	Y	Y	N	Y	Y	Y	N	Y	N	Y	N	N	Y	Y
Fallout 3	Y	Y	Y	Y	Y	Y	Y	N	N	Y	N	N	Y	Y
Fallout 4	Y	Y	Y	N	Y	N	Y	N	N	Y	N	N	Y	Y
TES3: Morrowind	Y	Y	Y	N	Y	N	N	N	N	Y	N	N	Y	Y
TES4: Oblivion	Y	Y	Y	Y	Y	Y	N	N	N	Y	N	N	Y	Y
TES5: Skyrim	Y	Y	Y	Y	Y	Y	N	Y	N	Y	N	N	Y	Y
Metro Exodus	Y	Y	N	Y	Y	N	N	N	N	Y	N	N	Y	Y

Table 1. Cont.

Game	Tutorial Presence	Context Sensitivity	Freedom	Availability of Help	Text	Helping Avatar	Video	Voice	Skippable	Story Integration	Practice Tool	Printed	Command Scheme Customization	Controller Diagram
Battlefield 1	Y	Y	N	Y	Y	Y	N	Y	N	N	N	N	Y	Y
SW Battlefront 2	Y	Y	N	Y	Y	Y	Y	Y	N	Y	N	N	Y	Y
STALKER: Shadow of Chernobyl	Y	Y	N	Y	Y	Y	Y	Y	N	Y	N	N	Y	Y
Valorant	Y	Y	N	Y	Y	N	N	Y	Y	N	Y	N	Y	Y
Team Fortress 2	Y	Y	M	Y	Y	N	Y	N	Y	N	Y	N	Y	Y
Paladins	Y	Y	N	Y	Y	Y	N	Y	Y	N	Y	N	Y	Y

#### 4.2. Game Prototype Development

We decided to develop a simple FPS equipped with two tutorial variants reflecting the combinations of parameters found by applying the classification framework and a main level. Technically, we relied on Unreal Engine 4 (<https://www.unrealengine.com/>, accessed on 20 October 2022) to implement the prototype. Table 2 provides an overview of the gaming features available in the two implemented tutorial variants.

**Table 2.** Summary of the gaming features included in first and second tutorial variants.

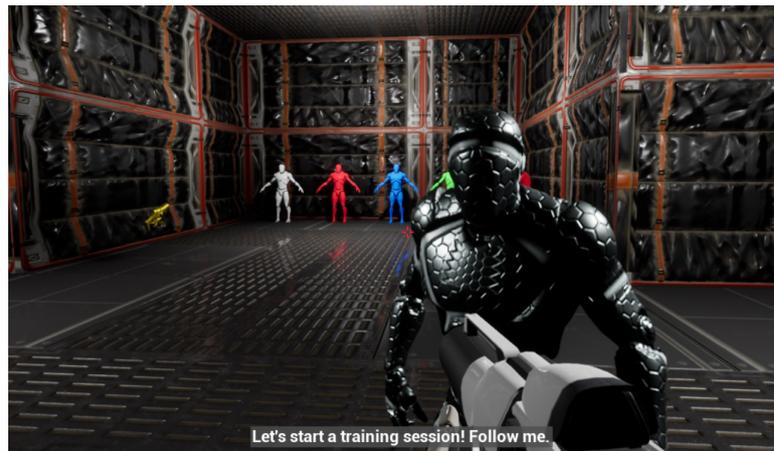
Element	First Tutorial Variant	Second Tutorial Variant
Tutorial Presence	yes	yes
Context Sensitivity	yes	yes
Freedom	no	no
Availability of Help	yes	yes
Text	yes	yes
Helping Avatar	yes	yes
Voice	yes	yes
Video	no	no
Skippable	no	yes
Story Integration	yes	no
Practice Tool	no	yes
Printed	no	no
Command Scheme Customization	yes	yes
Controller Diagram	yes	yes

In the first tutorial variant, the player wakes up after a big explosion, in a spaceship that has been invaded by (four different colored) aliens. With the help of a friendly non-playable character (NPC), the player will learn the basic elements of the game while traversing a linear space where events are triggered. The player can only move to the next part of the environment once s/he demonstrates an understanding of what s/he needs to do (explained by the NPC). If players perform a particular action before the corresponding explanation, the NPC will skip it.

The second tutorial variant implements the second combination of parameters as discussed in Section 4.1. The player is placed inside an open training room where s/he is greeted by a friendly NPC that provides her/him with a fixed linear sequence of short instructions about controls and game mechanics (see Figure 1).

Players are required to apply everything that was explained in tutorials, but in a context that aims at increasing the difficulty, using lights and small spaces to make users feel more anxious or under pressure. The gameplay of the prototype focused on:

- Unlimited ammunition with no reload time;
- One bullet, of the right color, is enough to kill an enemy;
- Players can pick up golden guns. This will give them the ability to shoot golden bullets that can kill every enemy for 10 s;
- Players die if an enemy stays within a specific radius from them for some time;
- The last room of the level contains the final boss that requires 10 shots of a special gun to be killed;
- Doors can be opened by shooting colored buttons;
- In the main level, enemies are spawned in infinite waves;
- Adaptive difficulty that gently invites users to play by combining exploration and shooting.



**Figure 1.** In-game screenshot of the second tutorial variant.

#### 4.3. User Testing

The game prototype was preliminarily evaluated through a heuristic analysis performed by team members, not by neutral external experts, to find potential usability issues. This enabled us to refine the interaction mechanisms of the game before releasing it to the final users.

The prototype was finally released in three different versions: one with no tutorial and the other two equipped with the tutorial variants described in the previous section. The user experiment consisted of a (i) recruitment phase, (ii) pre-questionnaire, (iii) testing phase, and (iv) post-questionnaire and file logging.

##### 4.3.1. Recruitment Phase

A call for participation was distributed over the Internet via Telegram groups, Discord servers, Reddit channels and other social media networks. These calls were distributed at different times of the day to obtain a globally diverse audience.

##### 4.3.2. Pre-Questionnaire

Firstly, participants were asked to complete the pre-questionnaire presented in Section 3.2. Once completed, a link to download one of the three versions of the game was provided to them. We divided participants into three different groups based on the version of the game they downloaded. Users were randomly allocated to one of the three testing scenarios, i.e., first/second tutorial variant or no tutorial, followed by the main game level.

##### 4.3.3. Testing Phase

A user could play the game prototype until completing the main level. The estimated time for playing the tutorial and the main level and answering the pre- and post-questionnaires was around 30 min.

Each of the three groups played a different version of the game:

- G1: *"Narrative Tutorial Group"* (15 participants): played a version of the game equipped with the first tutorial variant followed by the main game level.
- G2: *"Simple Tutorial Group"* (16 participants): played a version of the game equipped with the second tutorial variant followed by the main game level.
- G3: *"No Tutorial Group"* (15 participants): played a version of the game equipped only with the main game level (no tutorial available).

##### 4.3.4. Post-Questionnaire and File Logging

When completed or closed, the game automatically opens the post-questionnaire presented in Section 3.2. This was useful to ensure that we captured participants' responses as soon as they had finished playing.

The game provides an action logger to record all the interactions between the players and the game. We decided to log time stamps of important events (e.g., door opened, tutorial completed, buttons being shot) and aspects related to time and performance (e.g., time spent in the tutorial or in the main level, amount of deaths, enemies killed). At the end of the gaming activity, a log file keeping track of these interactions and events was automatically compressed in a password-protected .zip file by and released to the user. We decided to collect log files by asking participants to attach them to the post-questionnaire (through a dedicated feature to upload them). Uploading log files to the questionnaire was not mandatory.

#### 4.4. Results

Throughout the phase of recruitment, we had 114 potential participants who answered the initial form. Among them, 46 users played the game (organized in three independent groups as explained in Section 4.3.3) and answered the post-questionnaire. Participants were between 17 and 26 years old (65.8%) and primarily male (70%). A total of 40.4% of users declared themselves to be hardcore gamers, i.e., they play FPSs on a daily basis, at least 2 h a day, generally every day of the week. Then 22.8% of users declared they play FPSs of various kinds 4–5 h a week. Finally, the remaining 36.8% reported being casual gamers, seldomly playing FPSs.

Following the methodological steps of our approach (see Section 3), we applied ANOVA on each item of the post-questionnaire to compare the results obtained from the three independent user groups, namely G1 (Narrative Tutorial Group), G2 (Simple Tutorial Group) and G3 (No Tutorial Group), cf. Section 4.3.3. For each item, when group differences were evaluated as significant, we employed a two-sample *t*-test with a 95% confidence level to verify if the means of participants using different tutorial variants differ. Before running the two-sample *t*-test, we first used the Kolmogorov Smirnov statistic (KS Test) to establish the normality of the distribution of the collected data [34]. Then, we checked that the variances and standard deviations in both groups were approximately equal [31]. Finally, we measured the level of statistical significance, analyzing the resulting *p*-value. The results of the analysis, which include only the 5 items (out of 37) where a statistically significant difference was measured, are shown in Table 3.

Among the most interesting results, it appears evident that:

- Users in G3 (no tutorial) felt themselves more successful during the gaming experience but perceived the game pace as slower and, in general, the game mechanics more difficult to learn with respect to the other groups;
- Users in G1 (first tutorial variant) found the game more friendly than users in G3 and more understandable if compared with users in G2 and G3;

**Table 3.** Results of the ANOVA and two-sample *t*-test analysis. G1 is the “Narrative Tutorial Group”, G2 is the “Simple Tutorial Group” and G3 is the “No Tutorial Group”.

Element	ANOVA	T-Test Analysis
<i>Perceived level of success (1–10)</i>	<i>p</i> -value: 0.012 F-value: 4.87	<i>p</i> -value: 0.002 for G3 > G1 <i>p</i> -value: 0.018 for G3 > G2
<i>Not-Understandable/Understandable (1–7)</i>	<i>p</i> -value: 0.011 F-value: 5.01	<i>p</i> -value: 0.003 for G1 > G3 <i>p</i> -value: 0.003 for G1 > G2
<i>Easy/Difficult to learn (1–7)</i>	<i>p</i> -value: 0.038 F-value: 3.54	<i>p</i> -value: 0.012 for G1 < G3 <i>p</i> -value: 0.026 for G2 < G3
<i>Fast/Slow Pace (1–7)</i>	<i>p</i> -value: 0.016 F-value: 4.53	<i>p</i> -value: 0.004 for G1 < G3 <i>p</i> -value: 0.015 for G2 < G3
<i>Friendly/Unfriendly (1–7)</i>	<i>p</i> -value: 0.034 F-value: 3.66	<i>p</i> -value: 0.004 for G1 < G3

In addition, we analyzed the log files, keeping track of the users’ in-game actions. Because we made attaching log files to the questionnaire optional, we only collected files from 39 users, out of the 46 who played the game. The results obtained highlight that

the Simple Tutorial Group (G2) had a higher drop-out index than the Narrative Tutorial Group (G1). This is important to acknowledge because drop-outs can heavily influence a video game's success and financial outcome. Moreover, from a timing perspective, users of G1 needed on average less time—around 3.6 min—to complete the game's main level compared with the users in G2, who needed 5.1 min.

## 5. Discussion, Future Work and Concluding Remarks

In this paper, we presented a general-purpose approach aimed at supporting game designers in the identification of the most suitable tutorial style for a specific genre of video games. The approach consists of four methodological steps targeted at: (i) categorizing the tutorials of video games of a specific genre against a classification framework based on a set of predefined gaming features; (ii) developing a game prototype equipped with different variants of a tutorial that reproduces the most common combinations of gaming features found using the classification framework; (iii) performing a user experiment involving independent user groups to investigate the players' performance while interacting with the main level of the game prototype in the absence of any tutorial or after completing one of the implemented tutorial variants; and (iv) administering dedicated pre- and post-questionnaires to the players involved in the experiment for collecting quantitative data to be assessed through a suitable evaluation technique.

We performed a controlled comparative experiment administering a pre- and a post-questionnaire to 46 users who played the game (divided into three independent groups). The questionnaire results were analyzed using ANOVA and two-sample *t*-test.

The results revealed that the first tutorial variant, played by the Narrative Tutorial Group (G1), was the most understandable and user-friendly overall. Therefore, even if no general claim can be made, these results may support any developer aiming to realize an effective tutorial for FPS games. Indeed, by looking at the analysis made through the classification framework, a developer can implement a tutorial that includes the gaming features in the first tutorial variant.

While our approach is thought to be general-purpose and is customizable based on the specific game genre (i.e., further/existing parameters of the classification framework can be added/revised to capture tutorial aspects that are particularly relevant for the specific game genre, new questions can be added to the pre- and post-questionnaire, etc.), in the future it should be tested against non-AAA games and other game types (e.g., serious games, etc.) to broaden its validity. While we performed our user experiment online, we plan to test the approach in the future employing field studies.

In this paper, we chose to evaluate our approach with AAA games two reasons:

- AAA game modalities are currently being pushed as a standard for the future [?]. We believe that delivering an approach for realizing effective tutorials for AAA games can also positively affect the tutorial design for other kinds of games, where carefully crafting a tutorial requires a significant effort.
- AAA game developers are less likely to invest many resources in an aspect of the game (i.e., the tutorial design) that is time-consuming, provides less entertainment to the final user, or can potentially be skipped [8]. In this respect, our approach aims to mitigate the burden of designing effective tutorials by supporting game designers in selecting the most suitable subset of components for tutorial creation.

Finally, while we acknowledge the importance of studying the impact of tutorials to support replaying specific games, the approach proposed in this paper has been tested only in the case of first-game experiences. For this reason, we cannot claim anything about the approach's effectiveness when a game is replayed.

**Author Contributions:** Conceptualization, D.B., L.S.F., T.C. and A.M.; methodology, D.B.; software, D.B. and L.S.F.; validation, D.B. and A.M.; writing—original draft preparation, D.B. and L.S.F.; writing—review and editing, L.S.F. and A.M.; supervision, T.C. and A.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study. Written informed consent has been obtained from the patient(s) to publish this paper.

**Data Availability Statement:** Data supporting reported results is not uploaded anywhere online due to privacy constraints of the participants, but can be anonymized and analyzed if needed.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Bergonse, R. Fifty Years on, What exactly is a video game? An essentialistic definitional approach. *Comput. Games J.* **2017**, *6*, 239–255. [[CrossRef](#)]
2. Veneruso, S.V.; Ferro, L.S.; Marrella, A.; Mecella, M.; Catarci, T. CyberVR: An Interactive Learning Experience in Virtual Reality for Cybersecurity Related Issues. In Proceedings of the International Conference on Advanced Visual Interfaces, Salerno, Italy, 28 September–2 October 2020.
3. Ferro, L.S.; Sapio, F. Another Week at the Office (AWATO)—An Interactive Serious Game for Threat Modeling Human Factors. In Proceedings of the International Conference on Human-Computer Interaction, Copenhagen, Denmark, 5–8 October 2020; Springer: Berlin/Heidelberg, Germany, 2020.
4. van der Stappen, A.; Liu, Y.; Xu, J.; Yu, X.; Li, J.; Van Der Spek, E.D. MathBuilder: A collaborative AR math game for elementary school students. In Proceedings of the Annual Symposium on Computer-Human Interaction in Play Companion—Extended Abstracts, Barcelona, Spain, 22–25 October 2019.
5. Squire, K.D. Video game-based learning: An emerging paradigm for instruction. *Perform. Improv. Q.* **2008**, *21*, 7–36. [[CrossRef](#)]
6. Tannahill, N.; Tissington, P.; Senior, C. Video games and higher education: What can “Call of Duty” teach our students? *Front. Psychol.* **2012**, *3*, 210. [[CrossRef](#)] [[PubMed](#)]
7. Andersen, E.; O’Rourke, E.; Liu, Y.E.; Snider, R.; Lowdermilk, J.; Truong, D.; Cooper, S.; Popovic, Z. The impact of tutorials on games of varying complexity. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, TX, USA, 5–10 May 2012.
8. White, M.M. *Learn to Play: Designing Tutorials for Video Games*; CRC Press: Boca Raton, FL, USA, 2014.
9. De Freitas, S. Are games effective learning tools? A review of educational games. *J. Educ. Technol. Soc.* **2018**, *21*, 74–84.
10. Gamito, S.; Martinho, C. Highlight the Path Not Taken to Add Replay Value to Digital Storytelling Games. In Proceedings of the International Conference on Interactive Digital Storytelling, Tallinn, Estonia, 7–10 December 2021; Springer: Berlin/Heidelberg, Germany, 2021; pp. 61–70.
11. Roth, C.; Vermeulen, I.; Vorderer, P.; Klimmt, C. Exploring replay value: shifts and continuities in user experiences between first and second exposure to an interactive story. *Cyberpsychology Behav. Soc. Netw.* **2012**, *15*, 378–381. [[CrossRef](#)] [[PubMed](#)]
12. Phoenix, D.A. How to Add Replay Value to Your Educational Game. *J. Appl. Learn. Technol.* **2014**, *4*, 20–23.
13. Ballew, T.V.; Jones, K.S. Designing Enjoyable Video games: Do Heuristics Differentiate Bad from Good? In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*; SAGE Publications Sage CA: Los Angeles, CA, USA, 2006.
14. Johnson, D.; Gardner, M.J.; Perry, R. Validation of two game experience scales: The player experience of need satisfaction (PENS) and game experience questionnaire (GEQ). *Int. J. Hum. Comput. Stud.* **2018**, *118*, 38–46. [[CrossRef](#)]
15. Klimmt, C.; Blake, C.; Hefner, D.; Vorderer, P.; Roth, C. Player performance, satisfaction, and video game enjoyment. In *Proceedings of the International Conference on Entertainment Computing*; Springer: Berlin/Heidelberg, Germany, 2009.
16. Green, M.C.; Khalifa, A.; Barros, G.A.; Togelius, J. “Press Space to Fire”: Automatic Video Game Tutorial Generation. *arXiv* **2018**, arXiv:1805.11768.
17. Sweetser, P.; Wyeth, P. GameFlow: A model for evaluating player enjoyment in games. *Comput. Entertain.* **2005**, *3*, 3. [[CrossRef](#)]
18. Federoff, M.A. Heuristics and Usability Guidelines for the Creation and Evaluation of Fun in Video Games. Ph.D. Thesis, The Australian National University, Acton, Australia, December 2004.
19. Ryan, R.M.; Rigby, C.S.; Przybylski, A. The motivational pull of video games: A self-determination theory approach. *Motiv. Emot.* **2006**, *30*, 344–360. [[CrossRef](#)]
20. Poels, K.; de Kort, Y.A.; IJsselstein, W.A. D3.3: *Game Experience Questionnaire: Development of a Self-Report Measure to Assess the Psychological Impact of Digital Games*; Technische Universiteit Eindhoven: Eindhoven, The Netherlands, 2007.
21. Aytemiz, B.; Karth, I.; Harder, J.; Smith, A.M.; Whitehead, J. Talin: A Framework for Dynamic Tutorials Based on the Skill Atoms Theory. In Proceedings of the AIIDE, Edmonton, AB, Canada, 13–17 November 2018.
22. Paras, B. Learning to Play: The Design of In-Game Training to Enhance Video game Experience. Ph.D. Thesis, School of Interactive Arts and Technology, Simon Fraser University, Burnaby, BC, Canada, 2006.
23. Frommel, J.; Fahlbusch, K.; Brich, J.; Weber, M. The effects of context-sensitive tutorials in virtual reality games. In Proceedings of the Annual Symposium on Computer-Human Interaction in Play, Amsterdam, The Netherlands, 15–18 October 2017.
24. Zaidi, S.F.M.; Moore, C.; Khanna, H. Towards integration of user-centered designed tutorials for better virtual reality immersion. In Proceedings of the 2nd International Conference on Image and Graphics Processing, Beijing, China, 23–25 August 2019.

25. Humayoun, S.R.; Catarci, T.; de Leoni, M.; Marrella, A.; Mecella, M.; Bortenschlager, M.; Steinmann, R. Designing mobile systems in highly dynamic scenarios: The WORKPAD methodology. *Knowl. Technol. Policy* **2009**, *22*, 25–43. [[CrossRef](#)]
26. Humayoun, S.R.; Catarci, T.; de Leoni, M.; Marrella, A.; Mecella, M.; Bortenschlager, M.; Steinmann, R. The WORKPAD user interface and methodology: Developing smart and effective mobile applications for emergency operators. In Proceedings of the International Conference on Universal Access in Human-Computer Interaction, Las Vegas, NV, USA, 15–20 July 2009; Springer: Berlin/Heidelberg, Germany, 2009; pp. 343–352.
27. Marrella, A.; Mecella, M.; Russo, A. Collaboration on-the-field: Suggestions and beyond. In Proceedings of the 8th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2011), Lisbon, Portugal, 8–11 May 2011 .
28. Dix, A. Statistics for HCI: Making Sense of Quantitative Data. *Synth. Lect. Hum. Centered Inform.* **2020**, *13*, 1–181.
29. Hinderks, A.; Schrepp, M.; Mayo, F.J.D.; Escalona, M.J.; Thomaschewski, J. Developing a UX KPI based on the user experience questionnaire. *Comput. Stand. Interfaces* **2019**, *65*, 38–44. [[CrossRef](#)]
30. Hart, S.G. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*; Sage Publications: Thousand Oaks, CA, USA, 2006; Volume 50, pp. 904–908.
31. Sauro, J.; Lewis, J.R. *Quantifying the User Experience: Practical Statistics for User Research*; Morgan Kaufmann: Burlington, MA, USA, 2016.
32. Rogers, S. *Level Up! The Guide to Great Video Game Design*; John Wiley & Sons: Hoboken, NJ, USA, 2014.
33. Fullerton, T. *Game Design Workshop: A Playcentric Approach to Creating Innovative Games*; CRC Press: Boca Raton, FL, USA, 2014.
34. Chakravarti, I.M.; Laha, R.G.; Roy, J. Handbook of methods of applied statistics. In *Wiley Series in Probability and Mathematical Statistics (USA)*; Wiley: Hoboken, NJ, USA, 1967 .
35. Bernevega, A.; Gekker, A. The Industry of Landlords: Exploring the Assetization of the Triple-A Game. *Games Cult.* **2022**, *17*, 47–69. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.