information

Article

The Impact of Story Structure, Meaningfulness, and Concentration in Serious Games

Sofia Pescarin [1,*] and Delfina S. Martinez Pandiani [2]

[1] Institute of Heritage Science (ISPC), the National Research Council of Italy (CNR), 50019 Florence, Italy
[2] Department of Computer Science and Engineering, University of Bologna, 40126 Bologna, Italy
* Correspondence: sofia.pescarin@cnr.it

Abstract: This contribution analyzes the impact of factors related to story structure, meaningfulness, and concentration in the design of Serious Games. To explore them, the authors carried out an experimental evaluation aiming to identify relevant aspects affecting the cognitive-emotional impact of immersive Virtual Reality (VR), specifically Educational Environmental Narrative (EEN) Games. The experiment was designed around three main research questions: if passive or active interaction is preferable for factual and spatial knowledge acquisition; whether meaningfulness is a relevant experience in a serious game (SG) context; and if concentration impacts knowledge acquisition and engagement also in VR educational games. The findings highlight that passive interaction should only be encouraged for factual knowledge acquisition, that meaningfulness is a relevant experience and should be included in serious game design, and, finally, that concentration is a factor that impacts the experience in immersive games. The authors discuss potential design paths to improve both factual and spatial knowledge acquisition, such as abstract concept-oriented design, concluding that SGs should contain game mechanics explicitly supporting players' moments of reflection, and story structures explicitly aligned to educational facts.

Keywords: VR; serious games; meaningfulness; concentration; abstract concepts; game design

Citation: Pescarin, S.; Martinez Pandiani, D.S. The Impact of Story Structure, Meaningfulness, and Concentration in Serious Games. *Information* **2022**, *13*, 567. https://doi.org/10.3390/info13120567

Academic Editors: Lucio Tommaso De Paolis and Valerio De Luca

Received: 7 October 2022
Accepted: 30 November 2022
Published: 3 December 2022

1. Introduction

In 2019, the serious game (SG) "A Night in the Forum" (NiF) was published in the Sony PlayStation store for PS VR, as a prototype of a European project (REVEAL: Realizing Education through Virtual Environments and Augmented Locations). The game design adopted the findings of a previous evaluation of another, similar game in 2018 ("The Chantry" [1]). NiF is an Educational Environmental Narrative (EEN) videogame [2] dedicated to the Forum of Augustus in Rome. Its overarching goal is to engage young adult players (age +16) and their families with Roman history and archaeology, thus promoting cultural tourism and in general Cultural Heritage (CH), through videogames on widespread game platforms (see Supplementary Material for access to the videogame and its trailer). The game was developed taking into consideration two scenarios: (a) a guided play experience to be carried out in schools or museums and (b) a single user or family experience for home entertainment. Its educational aim is to involve players in Roman history, specifically that of the emperor Augustus and his Forum. The game's key learning outcomes were identified as the comprehension and recall of main historical facts, concepts, and spatial information. These outcomes also defined the evaluation metrics for the game's success. As such, 35 educational elements were chosen by the experts at the museum and used in the development of a framework that connected the gameplay with the narrative, as described in [2]. Relevant studies about game design and narratives were considered and applied to the design of NiF, including on the positive effect on comprehension of a well-defined story structure [3–8], on declarative knowledge acquisition [9], and on motivation and engagement [10]. Additionally, in the context of Virtual Reality (VR), works on how spatial memory may work better with active navigation [11–13] were also considered.

As such, crucial design insights were gathered from the fundamental study by Thorndike (1977) [4], which is still recognized as a valuable tool for strengthening memory recall in contemporary studies [8]. Ref. [4] analyzed content summaries performed by different participants who received the same, simple narrative story, but presented in different variants. The experiments suggested that "story comprehensibility and recall were correlated and were found to be a function of the amount of identifiable plot structure in the passages" [4] (p. 104). This study found that recall was higher for high-level organizational story elements, rather than lower-level details, and that recall was facilitated by the repetition of story structure, as repeating story content produced proactive interference. Based on these insights, and aiming to maximize memory and comprehension, we designed the game so that it unfolds through a main story (the frame) and secondary stories (episodes), connected to in-game specific tasks and discoveries. In this sense, the narratives of NiF were organized hierarchically (Figure 1), with all story elements explicitly structured in an effective order, following Thorndike's theory [4] (p. 79).

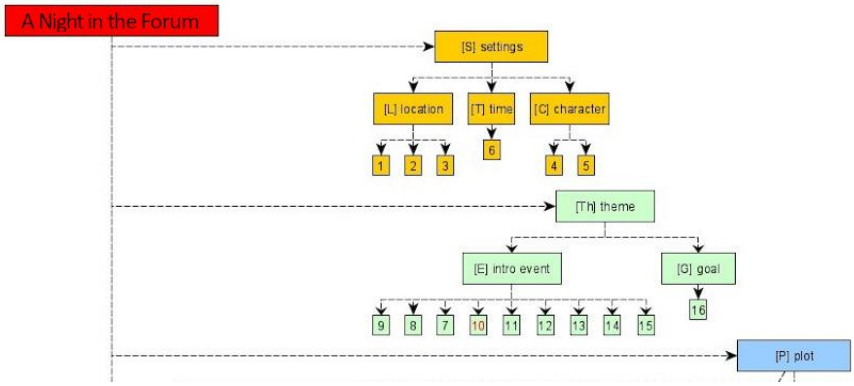

Figure 1. Subpart of the NiF game tree structure, demonstrating the methodology of the hierarchical story structure creation.

However, this theory was developed with regard to simple linear narratives, while the NiF narrative is nonlinear, as well as significantly more complex. Moreover, in addition to potential educational goals, SGs such as NiF are also meant to involve players in engaging experiences. For this reason, we wanted to better understand whether Thorndike's theory was also valid in more complex and interactive narratives, whether game mechanics could impact learning outcomes, and what might be the role of meaningfulness.

Following these questions, a first evaluation on "The Chantry" was initially conducted and described in [8], mainly focused on spatial memory and, specifically, on whether or not an explicit structure and implicit storyline are keys for recalling spatial information, and on how interaction impacts "learning and feeling". The results highlighted that an explicit story structure led to better spatial knowledge acquisition, while factual knowledge improves with passive interaction (e.g., guided 3D virtual environments). This evaluation highly suggested that SG mechanics could have an impact on specific learning outcomes, such as memorization and knowledge acquisition, but only if passive interaction was adopted, which led to an unfortunate reduction of the sense of presence and engagement.

To understand how storytelling and game mechanics could reinforce factual and spatial knowledge acquisition in games, without losing the sense of presence and engagement, we therefore prepared an evaluation on NiF, with the goal of identifying factors potentially involved in the game experience and their effect on learning outcomes.

2. Research Questions

Although NiF was produced by the same project and with similar goals to "The Chantry", the two games have critical differences. Firstly, NiF has a higher complexity, as its 3D scenario is wider and includes two versions of the same site (the archaeological site today and its reconstruction in the 1st AD). Additionally, the game has a more complex interaction modality,

with objects that can be picked up, stored, and moved, and task-oriented game mechanics. Finally, NiF includes more complex storytelling, developed in two levels: an explicit story that frames the entire narrative (a tourist lost in time that becomes the guardian of the forum), and a secondary narrative (Augustus and the imperial life in Rome).

The purpose of this study was to verify whether the results obtained while evaluating "The Chantry" were also valid for more complex games such as NiF. We also wanted to include other game design topics in this assessment: conceptual knowledge acquisition; how meaningfulness may be connected to knowledge acquisition and memorization; and the impact of concentration. Consequently, we defined three research questions: Research Question 1 (RQ1) aimed at investigating how passive/active interaction may affect factual/spatial knowledge acquisition and memorization in complex serious games with non-linear narratives; Research Question 2 (RQ2) aimed at investigating the connection between meaningfulness and conceptual knowledge in SGs; and Research Question 3 (RQ3) focused on the impact of concentration on knowledge acquisition and recall in SGs.

2.1. RQ1: Passive/Active Interaction and Factual/Spatial Knowledge in SGs

In order to better understand if and how the results of the first evaluation on "The Chantry" were also valid for more complex games, we took as reference the narrative of the NiF game and created different versions of it to verify under which conditions factual knowledge and spatial knowledge were better recalled. We decided to analyze not only the role of interaction (active/passive), but also the role of media channels (multimodal/written/audio) and communication styles (descriptive/narrative). Our questions were: (RQ1a) Are memory and comprehension strengthened by passive interaction more than by active interaction?, (RQ1b) Are the results valid for both factual and spatial knowledge?, and (RQ1c) How are they related to specific media channels and to different communication styles? As an outcome of these analyses, we wanted to highlight channels and styles that might be preferred when designing an SG with specific goals.

2.2. RQ2: Conceptual Knowledge and Meaningfulness in SGs

A new field known as "videogame tourism" is demonstrating how games can be used by cultural institutions and the tourism industry to exploit cultural destinations [14,15]. Studies in the tourism sector show that people search for "authentic, memorable and meaningful experiences" [16,17]. Meaningfulness can be defined as a crucial aspect also in SGs which make use of such cultural destinations, such as NiF. Looking at recent SGs and Virtual Museum experiences, we have noticed that the educational approach is mostly limited to factual knowledge, while other approaches are less considered. On the other side, results of EU projects such as EMOTIVE [18] have shown that meaningfulness is not only connected to information provision (factual knowledge), but to emotional attachment to CH, developed through personal engagement, active reflection and interpretation, personal meaning making, and participatory dynamics [19]. As such, gameplay designed around universal abstract concepts (i.e., beauty, love, death, etc.) that are part of human life, triggered by provocative questions, could strengthen meaningfulness. Meaning-making is an active process through which people revise or reconsider an event or experience. It has also been studied in the context of psychotherapy [20,21]. Within this work, we aimed to study the connection between play and meaningfulness, searching for personal meaning-making elements in gameplay experiences. Our research questions were: (RQ2a) How is meaningfulness connected to player experience, and how are emerging concepts connected to meaning-making?, and (RQ2b) how and which concepts emerge during game play? As an outcome, we aimed to identify factors that could influence personal reflection and meaning-making.

2.3. (RQ3) Concentration and Knowledge Acquisition in SGs

Finally, we identified concentration as a potentially relevant element to be considered, not only when evaluating SGs, but also in their design. Several studies have now high-

lighted that we are increasingly distracted, mostly due to the continuous technological solicitations and to "the never ending stream of information and addiction to smartphones and social media", with "consequences on our creativity, productivity and concentration" [22]. For this study, we consider concentration as "the action of maintaining one's attention on a particular task for a certain period of time without becoming distracted" [22] (p. xxix). Concentration has a direct relation to memory [23] and therefore to how much we recall factual, spatial, and conceptual information. VR has been demonstrated by several studies to have potential benefits for attention deficits, while others confirmed its advantages in cognitive performance [24]. We aimed to investigate if this is also true when VR is not specifically designed for attentional benefits, to understand whether there is an attentional benefit for users playing an SG, and whether distraction limits knowledge acquisition and engagement in interactive media. Based on these investigations, we identify appropriate design strategies to handle concentration and to retain players' attention as much as possible, including persuasive technologies, based on neurological and cognitive results that are already used for this purpose [25] (p. 66–69). Therefore, our last research question regarded how concentration impacts user experience and possibly how specific priming "concentration tasks" might improve performance and engagement: (RQ3) Does carrying out a concentration task, before other cognitive tasks, increase players' performance?

3. Experimental Methodology

At the beginning of 2020, we set up the experiment to investigate the three mentioned aspects, intending to define new design requirements. We chose NiF and involved a homogeneous group of participants. The overall evaluation would have required three iterations (during February–May 2020 with 60 testers), but we have concluded only the first two, due to the unexpected worldwide pandemic. The first iteration was used to verify the evaluation protocol and limit the number of conditions. The second iteration aimed to verify the research questions and limit the number of media during the third iteration.

3.1. Subjects

A total of 23 individuals were recruited to participate in a series of tasks (reading, listening, playing), followed by questionnaires and drawing activities, for a total of between 60 and 180 min (Table 1). During the first iteration, 3 subjects participated. During the second iteration, 20 subjects participated. The group included university students between 22 and 29 years of age, of various genders, and coming from 7 different countries. They all had a high education level in a domain connected to the humanities. All had at least a little knowledge of Roman History and were interested in Archaeology, as a general topic. Most of them reported visits to archaeological sites (18/23) and museums (22/23) from time to time. Only a minority had experience in playing games, either on smartphones or on platforms, such as PS VR (7/23). Regarding the preferred medium of story acquisition, 65% of the testers reported that they preferred to read stories, and only 5% preferred listening to them. Testers were assigned randomly to a group (between A and H). Each group was required to perform a sequence of tasks. The table below describes the tasks carried out by the participants and the adopted evaluation methods (Table 1).

Table 1. The three main analyses and guiding research questions. Meanings of the abbreviations are found in Appendix A.

	1-Active Interaction and Factual/Spatial Knowledge	**2-Conceptual Knowledge and Meaningfulness**	**3-Concentration—Factual Knowledge**
Research Questions	(RQ1a), (RQ1b), (RQ1c)	(RQ2a), (RQ2b)	(RQ3)
Tasks	pG/rS/lSN	pG/rS/lSN/d	f
Evaluation Methods	Q2a/b/c/d	Q2a/b/s/d	Q1, Q2

3.2. Methodology

In order to approach RQ1, we decided to include a comparative analysis of three main communication channels: active visual multimodal (play task: pG) in which participants were asked to play the NiF game; passive written (read task: rS) in which participants were asked to read the story of the game; and passive audio (listen task: lSN) in which they just had to listen to the story. In order to get more fine-grained insights regarding knowledge acquisition with respect to different communication styles, we decided to further develop two versions of the read story: a narrative version (SN) and a descriptive/expository version (SD). To do so, we took as reference work carried out in 2014 by the V-MUST.net EU project on terminology [26], where "descriptive communication" is defined as a style where "events, monuments, artifacts, artworks, customs or beliefs are defined, described and interpreted by a Sender who aims at informing and making aware a Receiver" [27], while "narrative communication", as a style in which a Sender provides information about events, monuments, etc. by arranging them to create an account of a subject by the Receiver's side [26–28]. These different communication channels and styles were assigned to the participants, in accordance to their group (Table 2): playing the game (pG) for groups A, B (gA,B); reading the narrative version of the story (rSN) for groups C, E (gC,E); reading the descriptive version of the story (rSD) for groups D, F (gD,F); and listening to the narrative story (lSN) for groups G, H (gG,H). No time limit was set for the tasks to avoid subjects feeling time pressure. At the end of the tasks, participants filled out two questionnaires specifically designed to get insight into factual knowledge acquisition (Q2a) and on spatial knowledge acquisition (Q2b,c,d).

Table 2. Groups, tasks and participants, and evaluation tools used.

Group	Tasks	Participants	Tools
A	f, pG	4	(s),(Q2a,b,c,d),(d)
B	pG	4	(s),(Q2a,b,c,d),(d)
C	f, rSN	2	(s),(Q2a,b),(d)
D	f, rSD	2	(s),(Q2a,b),(d)
E	rSN	2	(s),(Q2a,b),(d)
F	rSD	2	(s),(Q2a,b),(d)
G	f, lSN	2	(s),(Q2a,b)
H	lSN	2	(s),(Q2a,b)

In order to explore RQ2, we aimed to capture meaningfulness in interactive media as a structured evaluation tool for a SG prototype, in order to further support the development of applications that include specific meaning-making practices [29]. Unfortunately, there are not many focused studies on how to structurally analyze meaningfulness in interactive media. Nevertheless, we identified a number of approaches to analyze user interest during the game, meaningfulness becoming a sort of filter in the experience [30]: by searching for repeated words [31], by identifying causal links among terms [32], by finding connections between two elements [33], by analyzing ordered lists to understand perceived priority, identifying similitudes or metaphors used by the player (i.e., "this is like ... "), and searching for user motivation [34], or by searching for triggers [33,35] (p. 611). Another approach is to search or solicit personal interpretation and understanding through open-ended oral or written interviews (i.e., users are asked to summarize the experience), as writing and describing have been proven to "force structure onto thoughts and feelings that previously had not been clearly organized" [36]. Furthermore, in VR games such as NiF, which immerse players in 3D environments with geographic aspects, meaningfulness can be captured by exploring how this geography is experienced and identifying links with players' personal lives. An experimental approach inspired by psychogeography studies and applied to the exploration of emotional or unconscious aspects of an experience, as described by Debord [37] and Coverley [38], has already been applied to the VR/AR domain in [1,39]. As such, for the purpose of the current study, participants were asked,

after they performed their task, to write a short summary (write summary: s) of the experience, with what they recalled. Users who played the VR game were also asked to sit with a facilitator and draw (draw map: d) what they had experienced and captured their attention, in the form of a map or a drawing with details or comments.

Finally, in order to explore RQ3, i.e., to investigate the impact of concentration on knowledge acquisition and recall, we designed a "Focus exercise" (f), to be performed by half of the participants before the other tasks, with the goal of improving participants' concentration. We assumed that carrying out (f) before other cognitive tasks would increase performance in spatial and factual knowledge acquisition. This (f) task was designed by adapting a performance of the artist M. Abramovic, known as "Counting the Rice" [40]. This exercise is part of Abramovic's method developed over decades of research on performance and immaterial art, and originally debuted in Abramovic's *Cleaning the House* workshop, to "develop the participants' endurance, concentration" and to "allow the public to stretch their physical and mental limits". As such, the action of counting the rice represented a 'warm-up' exercise to gain concentration and prepare the body and mind for performance art [2].

3.3. Experimental Conditions

A $3 \times 2 \times 2$ design was used, simplified to obtain 8 conditions. We set 3 conditions for the media used (the way the information was presented): PLAY-GAME (pG), READ-STORY (rS), and LISTEN-STORY (lS), and 2 conditions for the concentration: FOCUS-EXERCISE (f), NON-FOCUS-EXERCISE. As mentioned, two conditions were set only for the written story—NARRATIVE and DESCRIPTIVE. For the LISTEN-STORY task, we decided to use only the narrative version of the story, to have a better match with the game experience, whose audio corresponded to it (Table 2). After participants had completed the required tasks, they were tested using different evaluation methods (Tables 1 and 2):

- (s): all participants were asked to summarize the experience;
- (Q2a): all participants filled Questionnaire 2a, for factual and conceptual knowledge;
- (Q2b): all participants filled Questionnaire 2b for spatial knowledge;
- (Q2c): only players (7 participants), filled Questionnaire 2c on spatial knowledge;
- (Q2d): only players (same 7) filled Questionnaire 2d on spatial knowledge;
- (d): 13 participants (all players, with a control group represented by the readers of the narrative story) were assigned the "Draw" task.

3.4. Experimental Procedure

The following protocol defined the procedure and modalities of the evaluation. Subjects were invited on a voluntary basis. On the date, they were introduced to the experiment, given a short explanation, and randomly assigned a letter (used also to anonymize the subjects). They were asked to read and sign an Informed Consent form, and complete a first anonymous questionnaire (Q1) meant to acquire demographic data through open questions and Likert scales, including the level of expertise on topics and past behavior regarding museums/archaeological sites, on the use of gaming technologies and preferred communication channels. After completing Q1, one of the testers, assigned to the group of players, reported potential problems with motion sickness and decided to leave the experiment, so only 19 testers participated in the subsequent tasks. These remaining testers were asked to proceed with the tasks, in the exact sequence indicated by the letter, in different dedicated spaces. When all tasks were completed, participants were asked to fill in a second questionnaire (Q2) and to participate in the drawing activity (d).

3.5. Tasks

We completed two iterations. The experiment was preliminarily tested on 3 recruited individuals as a first iteration, to evaluate the integrity and tiresomeness of the tasks at hand. After their positive evaluations, the experiment was extended to the other 20 testers.

The second iteration followed the procedure described above. The following sections include a description of the tasks and evaluation tools.

Focus Exercise (f). This task was included as a "primer" before other cognitive tasks for half of the testers, envisioned as a concentration exercise to help participants be present in both time and space. As such, the A-C-D-G groups became the experimental group (10 testers) carrying out the concentration exercise, while the other groups (B-E-F-H) became the control group (10 testers). The groups (gA-C-D-G) were asked to leave their phones and watches, and take earplugs, pencils, and sheets of paper (Figure 2). For each of the testers, stations were prepared, composed of a mound of black and white rice mixed together, and an instruction sheet. This sheet contained a copy of the text used in the original "Counting the Rice", asking participants to enact the "very simple, ironic, and useless action" of counting the rice, which repeatedly becomes "automatic but not immediate" so as to achieve concentration, and a cease of the "attack of anxiety and worries of everyday life" (full text is available in [40]). The testers were informed that they could use the blank sheet of paper and pencil as they liked, including that there were no final objectives to the exercise, that they would be informed when the time was done, and that they could stop before if they so pleased. No watch was available in the room. The average time spent on this exercise was 60 min.

Figure 2. Focus Exercise (**left**); playing the game exercise (**center**); phones and watch left by participants (**right**).

Read-Story task (rSN, rSD). As previously described, we created a narrative story, developed with first-person dialogues as an adventure of a tourist in the archaeological site (this version reports the exact dialogues of the game) (rSN); and a descriptive/expository version, in which the story is developed in third person with an external point of view, describing the possible actions and interactions of the tourist in the game (rSD). Below is an example of two extracts:

> *(rSN) There are ruins around [...]. Suddenly a deep and solemn voice somewhere: "In my 19th year, on my own initiative and expense, I raised an army and set the state free from the oppressive domination of a faction". Where did that come from?!*
>
> *(rSD) "While looking around with curiosity, the tourist finds an object on the stairs of a building with few still standing columns. The object is a shiny helmet. While the tourist moves closer to it, a solemn and deep voice starts talking, telling him that when he (the voice) was 19 years old, he raised an army to set the state free from the domination of a faction."*

Listen-Story task (lSN). The dialogues of the game were used to create an audio version of the story. An MP3 was edited and provided to the participants, who were asked to use their phones and headphones.

Play-Game task (pG). This task involved subjects in the gameplay of the entire video game. A description of the game is included in [2].

3.6. Data Collection

Data were collected through summaries, questionnaires, and drawings.

Summary (s). All participants were asked for a succinct summary of the story, with a focus on the sequence of events and important details. Our goals, in this case, were to identify the parts of the story recalled and their positions in the hierarchy, but also to explore concepts and keywords connected to meaningfulness.

Questionnaire 2 (Q2). It was designed to test factual knowledge (recall) and conceptual knowledge. It was aimed at understanding how much EEN Games can be a benefit for educational and touristic sectors, by analyzing what subjects understood and retained about the story. Q2 was divided into 4 sections.

Questionnaire 2a (Q2a). It was specifically designed to verify memory recall, through a series of 28 open-ended questions (about names, locations, characters, roles, functions, etc.), connected to the key educational information covered by the story, with the goal of verifying how much of this information was understood and recalled. It was completed by all participants.

Questionnaire 2b (Q2b). It was aimed at analyzing spatial memory, through 11 open-ended questions (including the request to describe what the forum looked like, listing as many places or areas of the forum as they could remember) and 11 multiple-choice questions. It was completed by all participants.

Questionnaire 2c and Questionnaire 2d (Q2c and Q2d). These were requested only of game players, aimed at an in-depth understanding of spatial knowledge acquisition. Q2c presented two photos taken at the archaeological site and two 3D reconstructions captured in the game and testers were asked to pair them (Figure 3). Q2d showed an aerial photo of the site and a list of places visited during the game, to be correctly located in the photo. Ten participants completed the two questionnaires.

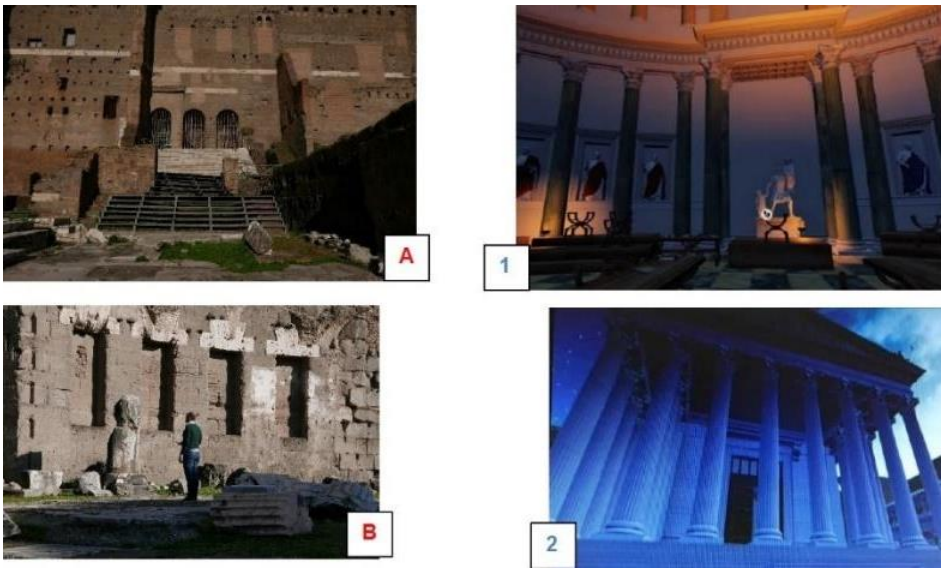

Figure 3. Exercise for pairing photos of archaeological sites and 3D reconstructions. Participants where asked to pair the images archeological sites (eg., (**A**) and (**B**)) with 3D reconstructions of the same monuments (eg., (**1**) and (**2**)).

Draw Map (d). Before leaving the experiment, 13 participants were invited to a separate area where they could sit and talk with a facilitator. They were provided with a creativity kit (paper sheets, colored markers, glue, scissors, etc.) and were asked to draw a map representing what they remembered about the experience, eventually also leaving comments. Thirteen drawings were collected (Figure 4).

Figure 4. Some of the resulting drawings (d), from group A (**left**) and E (**right**).

3.7. Data Analysis

The performance of participants on the 28 + 22 questions (Q2a, Q2b) was analyzed. First, each tester's responses were scored in terms of correctness and completeness on a scale from 0 to 3 (0 = no or incorrect answer, 1 = partial correct answer, 2 = correct answer, 3 = extra detailed correct answers). Subsequently, the results of testers were aggregated by group (A-H), and each group was given two scores:

- a BSP (Binary Score Percentage): the percentage number of questions that were answered at least partially correct averaged over participants in the group (minimum possible: 0, maximum possible: 100%);
- a CSA (Correctness Score Average): the average correctness of answers, calculated by summing the score value (0–3) of each of the answers and averaging over participants in the group (minimum possible: 0, maximum: 3)

The scores were further decomposed to obtain more detailed information about the performance of each group only in the first part of the questionnaire (Q2a) and only in the second (Q2b). As such, each group was given additional BSP and CSA scores for each of the two questionnaires. For each of the research questions, group scores were compared based on the different tasks that each group had undertaken. For example, to investigate the potential impact of (f) exercise, BSP and CSA scores were compared between groups that were required to participate in it and those that were not. Given the extremely reduced size of the groups, no independent statistical tests were performed, but rather score averages were calculated and are herein presented with no further independent analyses. Moreover, the summaries were analyzed in terms of length, completeness, sequences of reported facts, and wording. The drawings (d) were also considered in terms of the emergence of details and peculiarities.

4. Results and Discussion

4.1. RQ1 Results: Passive/Active Interaction and Factual/Spatial Knowledge

For each group, BSP and CSA scores were calculated over both Q2a and Q2b, and over each part of the questionnaire individually (see Table 3).

Table 3. BSP and CSA for each group on Q2a, Q2b, and on both.

	All	A (p)	B (p)	C (rSN)	E (rSN)	D (rSD)	F (rSD)	G (lN)	H (lN)
Overall BSP	51%	54%	43%	57%	53%	45%	31%	39%	38%
Overall CSA	0.87	1.01	0.71	1.09	0.84	0.72	0.61	0.64	0.56
Q2a BSP	46%	45%	39%	72%	57%	43%	38%	43%	36%
Q2a CSA	0.87	0.89	0.71	1.54	0.99	0.79	0.79	0.79	0.66
Q2b BSP	57%	63%	47%	43%	48%	46%	25%	36%	39%
Q2b CSA	0.87	1.12	0.71	0.64	0.70	0.66	0.43	0.50	0.46

RQ1a: Factual Knowledge Recall and Types of Interaction. To verify if factual knowledge was facilitated by passive or active interaction, we grouped the results in those produced by players (active interaction: gA-B, 7 participants) and by non-players (passive interaction: gC-D-E-F-G-H, 12 participants). We first considered results obtained in Q2a, because of its specificity for factual knowledge recall (Table 4). From these results we observed the tendencies of those involved in linear tasks, i.e., reading and listening, to better perform on factual knowledge (recalling facts), with pG testers obtaining a CSA of 0.8 and a BSP of only 42%, while non-players obtained a CSA of 0.9 and reaching 48.2% of BSP.

Table 4. Performance of players and non-players in Q2a questionnaire.

	All	**A-B (pG)**	**C-D-E-F-G-H**
Q2a BSP	46%	42%	48.2%
Q2a CSA	0.87	0.8	0.9

RQ1b Results: Spatial Knowledge Recall and Types of Interaction. We further analyzed the results considering if they were also applicable to spatial knowledge acquisition. Therefore, we compared the results of players with non-players in Q2a and Q2b (Table 5). The results, in this case, showed a tendency of players to better recall spatial knowledge, with a BSP of 55% and a CSA of 0.91, while non-players stop at 39.5% and 0.56. Therefore, overall, players performed 13% better when they had to recall spatial information, but 6.2% worse when they had to recall information such as facts, names, etc. This tendency also confirms that the results of the previous evaluation could be applied also to more complex narratives and games. Considering the overall results of both questionnaires, we observed that players performed better when educational games dealt with both factual and spatial information, considered equally important. To have more insights on spatial memory, we also considered the performance on Q2c and Q2d, required only of pG participants. We compared the results and verified that players performed better on Q2c, with an average of 67.5% correct answers, than on Q2d (Table 5), with only an average of 30% correct answers.

Table 5. Performance of players and non-players in Q2b and overall Q2a+Q2b questionnaires.

	All	**A-B (pG)**	**C-D-E-F-G-H**
Q2b BSP	57%	55%	39.5%
Q2b CSA	0.87	0.91	0.56
Overall BSP	51%	48.5%	43.83%
Overall CSA	0.87	1.36	0.74
Q2c	—-	67.5%	—-
Q2d	—-	30%	—-

*RQ1c Results: Factual Knowledge Recall and Types of Media and Communication Styles***.** To better understand how factual and spatial knowledge understanding and recall varied in accordance with the media channel types and communication styles, we analyzed the results of Q2a and Q2b, based on the tasks assigned to different groups, such as playing, reading the narrative/descriptive story and listening to the narrative story (pG (A, B); rSN (C,E); rSD (D, F); and lSN (G, H)). Regarding correctness of Q2a, rSN obtained the highest score, with a BSP of 64,5% correct answers. The other groups obtained BSP scores of 42% (pG), 40,5% (rSD) and 39,5% (lSN), not reaching even half the number of correct answers. CSA performance showed similar results, rSN obtained the highest score, with a CSA of 1.26; while the other groups (pG, rSD, and lSN) obtained CSA scores of 0.8, 0.79, and 0.72, respectively (Table 6). When we considered Q2b, the spatial questionnaire, we observed a partially different tendency, with pG reaching, in this case, a better result, with a BSP of 55%, while rSN obtained 45,5%, lSN 37.5%, and rSD only 35.5%. With CSA we

noticed a similar pattern to Q2a, with players able to recall much more and detailed spatial information (0.91), than the others rSN (0.67), rSD (0.54), and lSN (0.48).

Table 6. Performance according to communication media and style.

	All	A-B (pG)	CE (rSN)	DF (rSD)	GH (lSN)
Q2a BSP	46%	42%	64.5%	40.5%	39.5%
Q2a CSA	0.87	0.8	1.26	0.79	0.72
Q2b BSP	57%	55%	45.5%	35.5%	37.5%
Q2b CSA	0.87	0.91	0.67	0.54	0.48
Overall BSP	51%	48.5%	55%	38%	38.5%
Overall CSA	0.87	1.36	0.96	0.66	0.6

4.2. RQ2 Results: Conceptual Knowledge and Meaningfulness

We analyzed summaries (s) and drawings (d) to identify factors potentially related to meaningfulness. Within (s), we searched for significant keywords revealing values and purposes, repeated terms, connections among terms and causal links, ordering of events or other elements, mistakes, and unusual wording. We then also analyzed output from Q2a and Q2b considering the richness of details and emerging concepts as a guide toward meaningfulness, comparing results of pG with non-pG, and among media and communication style types. We finally studied (d) results from a psychogeography perspective, and considered and integrated this study with the results obtained in Q2c and Q2d specifically on spatial conceptualization.

We found indications of embodiment in the summaries through the use of the first person, with 60% of the players using the first person in (s). Regarding keywords and details provided in (s) and in Q2, we noticed that both players and non-players used similar concepts (such as citizenship, protection, power, justice/judge, law, government, divinity, family, descendance, inequality, cult, orientation/directions, protest, religion, deification), with more and various concepts used by readers (i.e., trust, leadership, poverty/richness, memory, immigration, social groups/sociality/social status, role, victory). Listeners used fewer concepts, in line with a lack of details in the answers.

Players and non-players used adjectives in (s), that could give indications of special meaning (i.e., "the tribunal had one main chair covered with red satin"); in some cases connected to personal interpretation (i.e., "trustworthy"), readers used words that give information on their imagination, with details not originally included in the stories (i.e., "the basement [was] all dark", "the stairs were steep [temple]", "[shelf] had lockers with numbers"). Only non-players recalled the types of Roman disputes that were mentioned in the game ("financials, migrations, family", "immigration, inheritance, city's law", "about slaves", "citizenship causes, fights for heredity, personal attacks and violence, murder", "family issues, problems with money and credit"), while 90% of the players could not recall them.

We also noticed that elements connected to meaningfulness were preceded by terms such as "like", "similar to", or "a kind of", which reveal the personal cultural, cognitive, and emotional context of the subject. Players have used the highest number of these expressions (i.e., Roman wax tablets were "like books", "like iPad", or the forum square was "like a football field", or the inner part of the temple was "like a small prison" and the gigantic statue of Augustus was "like a protector", "like a judge", "a kind of demon"), while we found these similarities reported by non-players (and only by readers) in only three cases.

Considering Q2a and Q2b, we noticed that in Q2a non-players, specifically rSN, provided the highest number of details in their answers (CSA: 1.54), while in Q2b players who performed (f) added the most detailed answers (CSA: 31.33), thus suggesting that concentration might have an impact also on meaningfulness. On average, considering the

two main groups of players and non-players, we noted a similar tendency, with non-players being more detailed (CSA: 0.9) in Q2a, and players being more accurate in Q2b (CSA: 1.36).

Looking at the map drawings (d), we noticed that all drawings were different from each other and that there were no significant differences in the groups (A,B,E) (Figure 4). We considered details, 2D and 3D elements, and comments involving aspects not related to the game. In 80% of the drawings, details were added, differing with regard to the typology (statues, tablets, columns); 60% contained written comments, including the name of the statues they recalled or objects found in the game. Only 20% reported personal comments, external to the game, While 80% tried to simplify and synthesize the experience, using 2D maps and including in them sketches of objects. There were no specific differences also considering players who had performed the focus exercise.

4.3. RQ3 Results: Concentration vs. Factual and Spatial Knowledge Recall

To investigate the potential impact of the focusing exercise (f), the BSP and CSA scores between groups that were required to participate in the exercise (A, C, D, G) and those that did not (B, E, F, H) were compared. Additionally, a more detailed comparison of BSP and CSA scores was completed between groups that did the same exact sequence of tasks, except for (f): A vs. B; C vs. E; D vs. F; G vs. H.

Consistently, on average, the groups that participated in (f) had higher performance scores than those that did not. Specifically, those participating in (f) exercise attained higher BSP (M = 49%) and CSA (M = 0.87) than those who did not (BSP M = 41%, CSA M = 0.68). This tendency holds for each of the questionnaires (Q2a and Q2b) as well. Furthermore, we observed that for groups involved in the same tasks (players, readers, and listeners), compared solely on the basis of participation on (f), the overall scores were always higher for those that did (f) exercise. The performance difference was especially noticeable in the groups who played the game (A and B), both in the BSP (Group A M = 54%, Group B M = 43%), and in the CSA (Group A M = 1.01, Group B M = 0.71). The biggest difference was found for these groups' BSP specifically in Q2b (Group A M = 63%, Group B M = 47%).

5. Conclusions

In conclusion, this article discusses the results of an experiment carried out in 2020 aimed at exploring factors with an impact on immersive VR Serious Games and specifically on Educational Environmental Narrative (EEN) Games. In particular, the analysis we presented regards an experiment designed around the serious game (SG) "A Night in the Forum" (NiF), published in the Sony PlayStation store for Virtual Reality headsets (PS VR). Although partial and with a limited number of testers due to the COVID-19 global pandemic, the evaluation we performed highlighted some factors affecting the cognitive-emotional impact of immersive SGs and specifically EEN games such as NiF. These factors need to be introduced and considered in game design.

Impact of Story Structure. Regarding the tendencies of linear narrative and passive interaction to improve factual knowledge acquisition (RQ1a), the results confirm previous evaluation conclusions [8]. Those involved in linear tasks in fact perform better in recalling facts than did players. Extending the analysis to spatial knowledge (RQ1b), however, we observed that while players tend to recall worse factual knowledge, they perform better when recalling spatial information. We have also understood, from the low performance obtained in Q2d, that there is an overall cognitive complexity in recognizing 3D reconstructed environments and locating them in real places, especially if viewed from a different perspective (in this case an aerial view (Figure 5) instead of a 1st person perspective (Figure 3)). In order to obtain better performance in these tasks, VR environments and mechanics should be specifically designed with this caveat in mind. This evaluation showed that the factual information that players best recalled corresponded to explicit knowledge in the main narrative of the game (higher hierarchy in Figure 1). On the other hand, players performed poorly on implicit knowledge and educational information that required more reasoning, interpretation, and reflection, while the players might have been

absorbed in the immersiveness of the spatial environment, with little possibility for more complex reasoning. We can conclude, therefore, that SGs should include appropriate game mechanics that explicitly support players' moments of reflection and be designed with a story structure more aligned to educational facts.

Figure 5. Testers were asked to locate game places in an aerial image of the archaeological site.

Impact of Meaningfulness. RQ2 investigated the connection between conceptual knowledge and meaningfulness. The experiment did not lead to clear results, although we did observe engagement and embodiment in the written reports of the multimodal game experience, as identified by specific keywords and recognizing the use of the first person. Other elements connected to meaning-making found in the summaries were located by terms such as "similar to", "like", or "kind of" and revealed subjectivity of the cultural, cognitive, and emotional context. Game players, as opposed to readers or listeners, used the highest number of these expressions, indicating how playing can potentially solicit more personal meaningfulness. More strategies to solicit reasoning and reflective moments can be included in game design, as they contribute to self-awareness and other positive effects [41,42]. Additionally, we can also conclude that concentration had an impact on meaningfulness. On the other hand, we observed only a limited number of elicited concepts, probably due to the fact that the design of the game was not aimed at making players focus on concepts. Thus, it is yet to be verified whether concept-oriented design could be a factor influencing personal reflection and meaning-making. With this study, we also explored potential new ways to evaluate concepts' emergence. The best method identified for this task has been the analysis of summaries, yet we likely need more structured and studied protocols on how to use psychogeography as an evaluation method.

Impact of Concentration. We observed that concentration has an impact on knowledge acquisition and also on engagement. RQ3 experimental results show that concentration had a positive impact on knowledge acquisition, not only when reading or listening, but also when engaged in VR immersive experiences. The results show a tendency towards higher performance for those who performed the attention exercise (f), in both Q2a and Q2b, especially for the game players. We believe that SGs need to seriously consider concentration as a factor, and thus distraction as a potential issue. What kind of game design and mechanics can be adopted to strengthen concentration? In this experiment, we adopted a priming, "concentration" exercise, which, however, is not very suitable for most game sessions. Other proposals could come from persuasive technologies [25] (p. 66–69), or the inclusion of analog exercises while playing.

All in all, the presented evaluation is a first step towards studying difficult matters regarding Serious Games (SG) and their design. One important caveat is that the chosen evaluation methods are mainly based on subjects' observations in the form of responses to questionnaires or written descriptions and might limit objectivity. In further studies, we aim at expanding the evaluation to a more heterogeneous audience. We also aim to include the identification and measurement of bio signals, including automatic recording and tracking of actions and oral expressions, which previous studies correlate to concentration and other relevant factors.

Supplementary Materials: The trailer of the videogame is available online at https://www.youtube.com/watch?v=9AvCQexNbrU; the game is available at https://store.playstation.com/en-gb/product/EP2996-CUSA12481_00-THEIMPERIALFORUM/.

Author Contributions: Experiment conceptualization, S.P.; experiment methodology, S.P. and D.S.M.P.; software, VR TRON; game assets, VR TRON, CNR ISPC (D. Ferdani, B. Fanini); general REVEAL project coordination, J. Habgood—Sheffield Hallam University. All authors have read and agreed to the published version of the manuscript.

Funding: The research leading to this publication has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements "REVEAl-Realising Education through Virtual Environments and Augmented Locations" (Grant Agreement N. 732599), [https://cordis.europa.eu/project/id/732599] (accessed on 30 November 2022); and "SPICE-Social Cohesion, Participation, and Inclusion through Cultural Engagement" (Grant Agreement N. 870811), [https://spice-h2020.eu] (accessed on 30 November 2022).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data, previously anonymized, supporting reported results are stored in CNR ISPC data cloud [https://cloud.ispc.cnr.it/apps/files/?dir=/Data_Evaluation/2019_NiF/] (accessed on 30 November 2022).

Acknowledgments: We acknowledge the Museum of Imperial Forum in Rome (Lucrezia Ungaro and Paolo Vigliarolo) for their scientific contribution, and Carlo Teo Pedretti and Mattia Spadoni (University of Bologna) for their contribution in the organization of the experiment.

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

Appendix A. Abbreviations in alphabetical order

CH:	Cultural Heritage
d:	drawing a map with details or comments
EEN:	Educational Environmental Narrative
ENViG:	Environmental Narrative Videogames
f:	focusing or concentration exercise
l:	listening task
lSN:	listening a narrative version of the story task
NiF:	Night in the Forum videogame
pG:	playing game task
PS VR:	PlayStation VR
Q1:	Questionnaire for personal data and background information
Q2a:	Questionnaire for factual and conceptual knowledge
Q2b:	Questionnaire for spatial knowledge evaluation
Q2c & Q2d:	Questionnaire on spatial knowledge
r:	reading task
RQ1:	Research Question 1
RQ2:	Research Question 2
RQ3:	Research Question 3
rSD:	reading a descriptive version of the story task
rSN:	reading a narrative version of the story task
s:	written short summary of the experience
SG:	Serious Game

References

1. Habgood, J.; Moore, D.; Alapont, S.; Ferguson, C.; van Oostendorp, H. The REVEAL educational environmental narrative framework for PlayStation VR. In Proceedings of the 12th European Conference on Game-Based Learning, Sophia Antipolis, France, 5 October 2018; pp. 175–183.
2. Pescarin, S.; Fanini, B.; Ferdani, D.; Mifsud, K.; Hamilton, A. Optimising Environmental Educational Narrative Videogames: The Case of 'A Night in the Forum'. *J. Comput. Cult. Herit.* **2020**, *13*, 1–23. [CrossRef]
3. Mandler, J.M.; Johnson, N.S. Remembrance of things parsed: Story structure and recall. *Cogn. Psychol.* **1977**, *9*, 111–151. [CrossRef]
4. Thorndyke, P.W. Cognitive structures in comprehension and memory of narrative discourse. *Cogn. Psychol.* **1977**, *9*, 77–110. [CrossRef]
5. Lorch, R.F.; Lorch, E.P.; Inman, W.E. Effects of signaling topic structure on text recall. *J. Educ. Psychol.* **1993**, *85*, 281–290. [CrossRef]
6. Zabrucky, K.M.; Moore, D. Influence of text genre on adults' monitoring of understanding and recall. *Educ. Gerontol.* **1999**, *25*, 691–710.
7. Norris, S.P.; Guilbert, S.M.; Smith, M.L.; Hakimelahi, S.; Phillips, L.M. A theoretical framework for narrative explanation in science. *Sci. Educ.* **2005**, *89*, 535–563. [CrossRef]
8. Ferguson, C.; van Oostendorp, H.; van den Broek, E.L. The development and evaluation of the storyline scaffolding tool. In Proceedings of the 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games), Vienna, Austria, 4–6 September 2019; pp. 1–8.
9. Gustafsson, A.; Katzeff, C.; Bang, M. Evaluation of a pervasive game for domestic energy engagement among teenagers. *Comput. Entertain.* **2010**, *7*, 1–19. [CrossRef]
10. Wouters, P.; Van Oostendorp, H. Overview of instructional techniques to facilitate learning and motivation of serious games. In *Instructional Techniques to Facilitate Learning and Motivation of Serious Games*; Springer: Cham, Switzerland, 2017; pp. 1–16.
11. Brooks, B.M. The specificity of memory enhancement during interaction with a virtual environment. *Memory* **1999**, *7*, 65–78. [CrossRef]
12. Christou, C.G.; Bülthoff, H.H. View dependence in scene recognition after active learning. *Mem. Cogn.* **1999**, *27*, 996–1007. [CrossRef]
13. Conniff, A.; Craig, T.; Laing, R.; Galán-Díaz, C. A comparison of active navigation and passive observation of desktop models of future built environments. *Des. Stud.* **2010**, *31*, 419–438. [CrossRef]
14. Berardone, F. Videogame-Induced Tourism. In *Esperienze Oltre Lo Schermo*; Youcanprint: Puglia, Italy, 2017.
15. Dubois, L.-E.; Gibbs, C. Video game–induced tourism: A new frontier for destination marketers. *Tour. Rev.* **2018**, *73*, 186–198. [CrossRef]
16. Lord, G.D. Keynote Presentation: The Power of Cultural Tourism. In Proceedings of the Wisconsin Heritage Tourism Conference, Lac du Flambeau, WI, USA, 17 September 1999; LORD Cultural Resources Planning and Management: Toronto, ON, Canada, 1999; Volume 17.
17. Chen, H.; Imran, R. Cultural tourism: An analysis of engagement, cultural contact, memorable tourism experience and destination loyalty. *Tour. Manag. Perspect.* **2018**, *26*, 153–163. [CrossRef]
18. Katifori, A.; Roussou, M.; Perry, S.; Cignoni, P.; Malomo, L.; Palma, G.; Dretakis, G.; Vizcay, S. The EMOTIVE Project-Emotive Virtual Cultural Experiences through Personalized Storytelling. *Cira@ Euromed* **2018**, 11–20.
19. Perry, S.E. The enchantment of the archaeological record. *Eur. J. Archaeol.* **2019**, *22*, 354–371. [CrossRef]
20. Taylor, S.E. Adjustment to threatening events: A theory of cognitive adaptation. *Am. Psychol.* **1983**, *38*, 1161–1171. [CrossRef]
21. Park, C.L. Making sense of the meaning literature: An integrative review of meaning making and its effects on adjustment to stressful life events. *Psychol. Bull.* **2010**, *136*, 257. [CrossRef]
22. Van Der Stigchel, S. *Concentration: Staying Focused in Times of Distraction*; MIT Press: Cambridge, MA, USA, 2020.
23. Clark, A.; Chalmers, D.J. The extended mind. *Analysis* **1998**, *58*, 7–19. [CrossRef]
24. Bashiri, A.; Ghazisaeedi, M.; Shahmoradi, L. The opportunities of virtual reality in the rehabilitation of children with attention deficit hyperactivity disorder: A literature review. *Korean J. Pediatr.* **2017**, *60*, 337. [CrossRef]
25. Griffiths, M.D. Adolescent social networking: How do social media operators facilitate habitual use? *Educ. Health* **2018**, *36*, 66–69.
26. Pescarin, S. Museums and virtual museums in Europe: Reaching expectations. *SCIRES-IT-Sci. Res. Inf. Technol.* **2014**, *4*, 131–140. [CrossRef]
27. V-MUST. Terminology, Definitions and Types of Virtual Museums. (V-Must.net—Project Report: D2.1b). 2014. Available online: https://www.academia.edu/6090456/Terminology_definitions_and_types_of_Virtual_Museums (accessed on 30 November 2022).
28. Palombini, A. Storytelling and telling history. Towards a grammar of narratives for Cultural Heritage dissemination in the Digital Era. *J. Cult. Herit.* **2017**, *24*, 134–139. [CrossRef]
29. Roussou, M.; Perry, S.; Katifori, A.; Vassos, S.; Tzouganatou, A.; McKinney, S. Transformation through Provocation? In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*; Association for Computing Machinery: New York, NY, USA, 2019; Paper 627; pp. 1–13. [CrossRef]
30. Park, C.L.; Folkman, S. Meaning in the context of stress and coping. *Rev. Gen. Psychol.* **1997**, *1*, 115–144. [CrossRef]
31. Underwood, B.J.; Schulz, R.W. *Meaningfulness and Verbal Learning*; J.B. Lippincott: Philadelphia, PA, USA, 1960.

32. Pennebaker, J.W.; Francis, M.E. Cognitive, emotional, and language processes in disclosure. *Cogn. Emot.* **1996**, *10*, 601–626. [CrossRef]
33. Baumeister, R.F.; Vohs, K.D. The pursuit of meaningfulness in life. In *Handbook of Positive Psychology*; Oxford University Press: Oxford, UK, 2002; Volume 1, pp. 608–618.
34. Baumeister, R.F. *Meanings of Life*; Guilford Press: New York, NY, USA, 1991.
35. Emmons, R.A. Motives and life goals. In *Handbook of Personality Psychology*; Academic Press: Cambridge, MA, USA, 1997; pp. 485–512.
36. King, L.A.; Pennebaker, J.W. Thinking about goals, glue, and the meaning of life. *Ruminative Thoughts* **1996**, *10*, 97–106.
37. Debord, G. *Introduction to a Critique of Urban Geography*; Praxis Press: Glasgow, UK, 2008.
38. Coverley, M. *Psychogeography*; Pocket Essentials: Harpenden, UK, 2006.
39. Echavarria, K.R.; Dibble, L.; Bracco, A.; Silverton, E.; Dixon, S. Augmented Reality (AR) Maps for Experiencing Creative Narratives of Cultural Heritage. In Proceedings of the EUROGRAPHICS Workshop on Graphics and Cultural Heritage, Sarajevo, Bosnia and Herzegovina, 6–9 November 2019.
40. Abramovic, M. *Walk through Walls: A Memoir*; Crown Archetype: New York, NY, USA, 2017.
41. Baumeister, R.F.; Wilson, B. Life stories and the four need for meaning. *Psychol. Inq.* **1996**, *7*, 322–325. [CrossRef]
42. McAdams, D.P. Personality, modernity, and the storied self: A contemporary framework for studying persons. *Psychol. Inq.* **1996**, *7*, 295–321. [CrossRef]