

Development of a model for explaining the learning outcomes when using 3D virtual environments in informal learning settings

Emmanuel Fokides¹  · Penelope Atsikpasi¹

Received: 11 March 2018 / Accepted: 17 April 2018 / Published online: 24 April 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract The study presents the development and testing of a model for explaining the learning outcomes when individuals use 3D virtual environments (VEs) in informal learning settings. For that matter, a VE was developed, presenting the work of a sculptress, namely Nausica Pastra. The following subjective factors were considered so as to build a research model: perceived usefulness, perceived ease of use, motivation, presence, perceived application's realism, as well as the enjoyment when using VEs. Self-reported data together with the results of tests embedded in the VE were gathered from 612 individuals. Structural Equation Modelling was employed for model testing and parameter estimation. The analyses of the results revealed a good model fit and 53% of the variance in the learning outcomes was explained. Out of the twenty research hypotheses, fifteen were supported. It was found that the most significant factors affecting the learning outcomes were motivation, perceived ease of use, perceived usefulness, and enjoyment. The implications of the findings for experts involved in the development of virtual museums are also discussed.

Keywords 3D virtual environments · Informal learning · Structural equation modeling · Subjective constructs · Virtual museum

1 Introduction

Almost all educational systems, by incorporating various technological innovations in the curriculum, try to provide a high quality and a wide range of knowledge, to connect

✉ Emmanuel Fokides
fokides@aegean.gr

Penelope Atsikpasi
premnt16002@aegean.gr

¹ Department of Primary Education, University of the Aegean, Rhodes, Greece

the learning subjects with daily practices, to release the education of its spatial and temporal limitations and to foster a number of skills, abilities, and learning dispositions, called the “21st century skills” (Trilling and Fadel 2009). A technology that supports the realization of all of the above, both in formal, non-formal, and informal settings, are 3D virtual environments (VEs) (Duncan et al. 2012).

Researchers have tried to evaluate the learning outcomes when using VEs in all levels of education, focusing on the teaching methods or on certain technical aspects, while the inclusion of psychological or subjective factors that may have an impact on the users’ learning experience is rather limited (e.g., Merchant et al. 2012; Tüzün and Özdiç 2016). Moreover, there is no common consensus on which factors to include, how they interplay with each other, or even how they are defined. On the other hand, by knowing how specific factors positively or negatively affect users, we can, with appropriate interventions, maximize the effectiveness of VEs.

This study contributes towards this direction, by developing and testing a model for examining a number of factors that was theorized to shape the learning experience when using VEs in informal learning settings. What is more, how the included factors interact with each other was also examined. It has to be noted that informal learning was chosen because it is the least studied in relation to VEs. For that matter, a virtual museum was developed presenting the work of Nausica Pastra, a Greek sculptor/artist with an interesting international carrier, who, on the other hand, was not widely known in her mother country. Following a literature review, the factors that were included in the model were perceived usefulness, perceived ease of use, perceived application’s realism, presence, motivation, as well as the enjoyment when using VEs. Also, the learning outcomes, what the users were able to learn when using the VE, were included as one of the model’s constructs. In the sections to follow, a brief review of the literature regarding informal learning and virtual museums is presented, followed by the literature review regarding the factors affecting the learning outcomes when using VEs. Next, the research rationale, methodology, and the results analyses are presented. Subsequently, the results are discussed, and the conclusion completes the work.

2 Informal learning, museums, and (3D) virtual museums

Informal learning is spontaneous, meaning that it is a non-systematic and often a non-organized process that can be evaluated on the basis of one’s overall lifelong learning (Coombs et al. 1973). It can also be defined as the sum of the learning activities not taking place in a formal classroom (Gerber et al. 2001). Informal learning is also experience-based, non-routine and often tacit (Marsick and Watkins 1990). Taking into account the above, it can be argued that informal learning originates accidentally and sporadically, does not follow a specified curriculum, and the learner is highly motivated to learn. Also, since it is not planned pedagogically or systematically organized, it is difficult to quantify and test it. The most significant part of the educational research is related to formal education and planned teaching; informal learning is often overlooked and not regarded as a particularly valid form of learning. On the other hand, there are several studies highlighting the importance of informal learning in an individual’s life (e.g., Fonseca et al. 2016; Harrop and Turpin 2013). Examples of where informal learning can take place include after-school activities, community-based programs, libraries, at home, and museums.

Informal learning and museums are closely related; a substantial number of studies examined the former in the context of the later. Museums are institutions having as a purpose to collect, preserve, and display artifacts (permanently or temporarily) of artistic, cultural, or scientific importance with the objective to educate the public (Alexander et al. 2017). Archaeological sites, monuments, traditional settlements, urban and natural environments can also act as museums. From an educational perspective, museums try to engage visitors in learning experiences, by stimulating their curiosity and interest (Hooper-Greenhill 2013). Consequently, learning in museums is mostly experiential, depending on the individual's interests and preferences. As already mentioned, several studies have been conducted regarding learning in museums, in all age groups and in all kinds of settings. These demonstrated that the individuals experienced enjoyment and positive emotions, being free to browse through the museum and interact with its exhibits at their own pace (e.g., Dilli and Dümenci 2015; Miglietta et al. 2008). At the same time, the visitors acquired a varying degree of useful information and knowledge and developed cognitive, affective, as well as linguistic skills (Ampartzaki et al. 2013; Hackett 2014; Synodi 2014).

In recent years, a new form of museums has emerged, that of digital/virtual museums. These environments can be stand-alone applications, websites, mobile applications, or a combination of the above. They draw on the characteristics of physical museums, complementing or enhancing the experience one has, through personalization and interactivity. The virtual museums can act as the digital footprints of real museums, as exhibitions created from diverse resources, or they can be fictional environments with equally fictional exhibits. Regardless if these digital museums represent real or imaginary environments, their visitors have the ability to communicate, explore, or even modify the digital space and objects (Pujol and Lorente 2013). Virtual museums are quite popular and attract the research interest for the examination of the experiences (learning or otherwise) the users have when “visiting” them (e.g., Li and Zhou 2016; Naya and Ibáñez 2015; Pescarin et al. 2012).

There are several ways to develop such interactive environments, however, VEs have notable advantages compared to other types of applications (i.e., web pages and 2D multimedia applications) because they (a) can achieve the ambience of a physical museum, (b) allow free navigation inside the simulated space, (c) allow increased interactions with the 3D objects/exhibits as well as with other users present in the VE, thus, offering a more realistic experience (Pescarin et al. 2012), and (d) the 3D digital artworks can be viewed from multiple perspectives allowing the user to get an idea of their actual dimensions, placement, and use (Skamantzari and Georgopoulos 2016). Indeed, several studies examined the learning outcomes when using 3D virtual museums in a variety of settings such as (a) formal or non-formal education (e.g., Seok and Kim 2014), (b) by combining visits to a real museum and virtual visits to its digital counterpart (e.g., Jones and Alba 2016), and (c) in the context of serious games (e.g., Paliokas and Sylaiou 2016). On the other hand, research on the impact of 3D virtual museums, and of VEs in general, in strictly informal learning settings is rather sparse (e.g., Voss et al. 2013).

3 Factors affecting the learning experience when using 3D virtual environments

On the basis of what was presented in the preceding section, it can be argued that 3D virtual museums provide a good basis for examining VEs in relation to informal learning. VEs

offer unique and realistic experiences to users, opportunities for personalized learning (Rapanotti et al. 2012), and, in general, their educational value is widely accepted (Allison et al. 2012; Merchant et al. 2014). Although most researchers reported -more or less- satisfactory learning outcomes when using them, there are many different, and sometimes contradictory, explanations why these outcomes were achieved. In addition, the focus was on formal education (Merchant et al. 2014). Far fewer studies examined psychological and subjective factors as well as their interactions. The study of the above in informal learning settings proved to be even more sparse.

In order to have a better understanding on which factors affect the user's learning experience when using VEs, a quite extensive literature review was carried out spanning from 2007 to 2017, revealing a substantial number of well-organized and interesting studies that provided useful information and ideas. For instance, deNoyelles et al. (2014) examined the role of self-efficacy. In another case, the effects of presence on the learning outcomes were examined (Hassell et al. 2012). Once again presence together with enjoyment (Sylaiou et al. 2010) or with perceived usefulness (Tüzün and Özdiñç 2016) were the factors of interest in other studies. Jia et al. (2014) found that perceived system efficacy and self-efficacy had an impact on the effectiveness of VEs in informal education. Also, Naya and Ibáñez (2015) found that factors such as enjoyment and perceived ease of use affect the learning outcomes, while Yaman et al. (2008) considered the instructional support and the learner's personal interests. Presence and usability were used by Merchant et al. (2012) for explaining the observed learning outcomes in VEs. Finally, Lee et al. (2010) developed a model that tried to explain the learning outcomes in VEs by using a number of VE's features together with cognitive beliefs, presence, control, motivation, and reflective thinking.

The vast majority of the studies that were found had little in common; different factors were used as key determinants of the learning experience, different types of VEs were examined, and the learning subjects/learning settings were also diverse. Also, much of the research did not deal with the learning experience per se but with factors that, beyond any doubt, can shape it (e.g., usability). What is more, there is no common consensus on the definition of some factors. For example, the terms presence, immersion, and flow were used interchangeably and were even examined using the same questions. Nevertheless, by summarizing the literature review, six factors were identified, commonly used for explaining the learning outcomes, that can be grouped into the following categories:

3.1 VEs technical features

- Perceived realism. Though a VE's realism depends on certain technical features (e.g., how detailed the 3D models are and to what extent a user can interact with them), individuals perceive it differently. Studies indicated that realism acts as a facilitator of learning (Dalgarno and Lee 2010; Harrington 2012; Lee et al. 2010). Also, in this study, realism was viewed as the sum of other factors such as perceived audio and visual adequacy and perceived interactions' adequacy.
- Perceived ease of use. Perceived ease of use (PEU) has been found to play a key role in one's experience when using VEs (e.g., Lee et al. 2010). As with perceived realism, PEU was also viewed as the sum of other factors such as perceived usability and perceived feedback's adequacy.

3.2 State of mind

- **Enjoyment.** Studies have shown that enjoyment and, in general, the positive feelings towards a VE contribute to knowledge acquisition (Faiola et al. 2013).
- **Presence.** Presence is one of the most troubling factors, suffering from definitional issues because there is a tendency to use this term as a synonym for engagement, flow, or immersion. It refers to the phenomenon of one acting and feeling as being in the digital world (Sanchez-Vives and Slater 2005). Though presence has an impact on the learning outcomes (e.g., Bulu 2012; Lee et al. 2010), there is a lot of controversy regarding its exact role. There are those who supported that it depends on the use of specialized equipment (e.g., North and North 2016) and those who supported that it depends on the user's personality (e.g., Nunez 2004). In order to examine the above, it was decided to include presence as a construct, view it as a synonym of immersion, and examine its role.

3.3 Learning facilitating features

- **Perceived usefulness.** Perceived usefulness is the degree to which one believes that his/her performance would be enhanced by using an ICT tool, (Davis et al. 1989). It is commonly used as a construct in studies regarding VEs' impact on learning (e.g., Lee et al. 2010; Sharda et al. 2004). In terms of its impact on knowledge acquisition, it can be viewed as perceived learning facilitator, how one thinks that the use of a specific ICT tool can help him/her acquire some knowledge.
- **Motivation.** Motivation is not a static attribute, but it is sensitive to how the content is presented and inherently volatile (Linnenbrink and Pintrich 2002). Research has indicated that, in VEs, motivation and the learning outcomes are closely connected (McLellan 2004). It has to be noted that in informal learning intrinsic motivation (one's need to learn about the application's subject) is already present and quite strong. Thus, it was interesting to examine whether an application can have an impact on one's already strong motivation to learn.

4 Research model and hypotheses

Figure 1 presents the proposed research model, which includes the factors presented in the previous section. Since the literature suggests diverse relationships among the factors, it was decided to add and test all possible relationships. The study's hypotheses were based on these relationships:

- H1a-e: Perceived realism influences (a) presence, (b) enjoyment, (c) motivation, (d) perceived usefulness, and (e) the learning outcomes.
- H2a-e: Perceived ease of use influences (a) presence, (b) enjoyment, (c) motivation, (d) perceived usefulness, and (e) the learning outcomes.
- H3a-d: Presence influences (a) enjoyment, (b) motivation, (c) perceived usefulness, and (d) the learning outcomes.

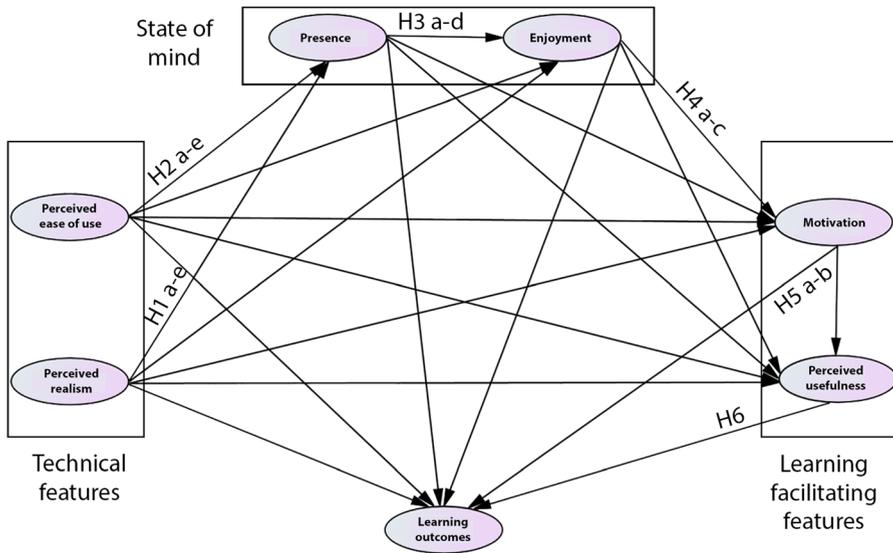


Fig. 1 The research model

- H4a-c: Enjoyment influences (a) motivation, (b) perceived usefulness, and (c) the learning outcomes.
- H5a-b: Motivation influences (a) perceived usefulness and (b) the learning outcomes.
- H6: Perceived usefulness influences the learning outcomes.

In this study, the dependent variable was the learning outcomes, while all the other acted as independent variables. Because structural equation modeling was going to be used for model testing, perceived ease of use and realism were the exogenous variables, while presence, motivation, enjoyment, and perceived usefulness were the endogenous variables.

5 Method

To test the research model, a study was designed and implemented utilizing a 3D virtual museum developed specifically for its purposes. The target group was adults, data were collected using a questionnaire and tests, and the project lasted for six months. The above are further elaborated in the coming sections.

5.1 Participants and duration

An invitation was issued addressed to anyone interested to participate in the project. This invitation was posted on social networks as well as in large communities of virtual worlds (i.e., Second Life and OS Grid). Participants were informed that they were going to “visit” an art museum and complete a short questionnaire, that the study was

conducted on a voluntary basis, and that by entering the virtual museum consent to participate was deemed to have been given. Detailed instructions were provided on how to login to the virtual world and on other technical issues (how to connect, what software to use, and the necessary settings). Participants were also informed that the time-period they remained connected and their IP addresses were recorded. The virtual museum remained active for a period of six months (June to December 2017) in order to collect a sufficient sample size.

5.2 Materials

As already mentioned, a 3D virtual museum was developed for the project's needs, using OpenSimulator (<http://opensimulator.org/>), an open source platform which allows the development, management, and hosting of VEs. The museum presented the work of the sculptress Nausica Pastra who pursued an international career but she remained relatively unknown in Greece. This was exactly the reason for selecting this artist. Since the participants were Greeks, there was little chance the museum's visitors to be familiar with her life and work. Consequently, it would be easier to (a) check whether the participants acquired any knowledge and (b) to attribute this to their "visit" to the museum rather than to any other factor.

Most of Pastra's creations were geometric forms based on mathematical functions. Also, the team had at its disposal high-resolution photographs of her sculptures. Therefore, creating their exact digital replicas was a relatively easy task. All were placed in a 512X512 meters imaginary natural environment (Fig. 2). While there was no specific order in their placement, the exhibits were grouped in small and big ones. The former were placed inside two small buildings, while the latter were placed in open spaces. Screens placed next



Fig. 2 Screenshots from the 3D virtual museum

to each exhibit provided information for them. Also, other screens provided information (in the form of websites, texts, and other audio-visual material) for Pastra's life and milestones in her carrier. The users could only walk, flying was disabled. This was done because the museum was divided into three areas (two main areas, and an exit area) separated by walls and gates. For gaining access to the next area, users had to answer some questions in order to lift the gates (see section "Instruments" for further details). The development of the application took about two months.

5.3 Instruments

For measuring the model's factors, a questionnaire was used. For developing it, a number of questionnaires (with high reliability and validity) used in other studies related to VEs were considered. Specifically, presence was measured using four questions from the presence questionnaire used by Novak et al. (2000). For measuring perceived usefulness, four questions from the Computer Attitude Scale (Selwyn 1997) were adapted. This scale is predominantly used in studies examining users' intentions to use various technologies in an educational context (e.g., Fokides 2017; Teo and Noyes 2011). Four questions coming from the same scale were used for measuring perceived ease of use. From Witmer's and Singer's questionnaire (1998), four items were used for measuring the perceived application's realism. The Intrinsic Motivation Inventory (Tamborini et al. 2010) provided four items for measuring enjoyment. Finally, from Martens's et al.'s questionnaire (2007) four questions were adapted for measuring motivation.

The questions were translated into Greek by two pairs of experts (a computer science professional and a psychologist with expertise in VEs), all being proficient in the English language. The two versions were then back-translated into English and viewed by another pair of experts. The final version was obtained through a consensus meeting of all experts, aiming to assess the semantic adaptation. Thus, the final version of the questionnaire was formed, having a total of twenty-four five-point Likert-type questions (worded "Strongly Agree", "Agree", "Neutral", "Disagree" and "Strongly Disagree"). The questionnaire is presented in the Appendix.

To effectively communicate substantive knowledge in informal learning conditions is a formidable task, and, probably, even more difficult is to quantify the acquired knowledge. Nevertheless, this task was accomplished by using three short tests, with fourteen questions each (right-wrong, multiple choice, and fill-in-the-blanks), all relevant to what was presented in the museum (i.e., the life and important milestones in Pastra's life, her sculptures, and artworks). The first two tests were embedded in the gates blocking access to the different areas of the museum, while the third was presented to the users together with the questionnaire. It has to be noted that taking the tests was not mandatory; the gates could be lifted either by answering the questions or by simply skipping the tests altogether. Also, it was not required to answer the questions correctly for allowing access to another area; even if all answers were wrong the gates were lifted. In the exit area, there was a large screen presenting the final set of questions and the questionnaire. As with the other two sets of questions, participants were given the choice to take the test and the questionnaire or simply skip this step and log-out. The reason for doing so, was because the study examined informal learning; under no circumstances, users should be left with the impression that they were tested for what they have learned.

5.4 Procedure and data processing

After logging-in, participants were free to explore the virtual environment in any way they wanted, examine the exhibits in any order they liked, and study the information screens for as long as they considered necessary. The only task they were given was to reach the exit area of the museum. Also, they were asked to remain logged-in for at least three-quarters of an hour, which was considered the minimum amount of time for a user to visit all the areas.

Scores were obtained by allocating numerical values to the responses in the questionnaire: “Strongly Agree” scored 5, “Agree” scored 4; “Neutral” scored 3; “Disagree” scored 2 and “Strongly Disagree” scored 1. Also, three scores were computed on the basis of the correct answers in the tests and were added as variables, forming the learning outcomes factor. From the data analyses the following cases were excluded (a) questionnaires and tests from users who remained logged-in for less than forty-five minutes, (b) multiple visits to the museum from the same IP address and, subsequently, multiple completion of the questionnaire and tests (only the first visit was considered valid), (c) questionnaires without three complete sets of tests and vice versa, and (d) questionnaires without variation in their responses. As a result, the final sample size was 612 participants (out of the initial 875).

6 Results

Coming to the results, the distribution of males and females was fairly even (49% and 51% respectively). Up to the age of thirty was 61% of the participants, while above the age of forty was only 10% of them. Scores from each of the questionnaire’s items were summed to provide total scores on each construct (including the tests’ scores) (Table 1). As a normative guide to interpretation, a score below the 25th percentile (19) can be interpreted as a relatively low score in the tests, whereas a score above the 75th percentile (32) can be interpreted as a relatively high one.

6.1 Exploratory and confirmatory factor analysis

Although all the questionnaire’s items came from already validated scales, an Exploratory Factor Analysis (EFA) and a Confirmatory Factor Analysis (CFA), were conducted in order to (re)establish its overall validity and reliability. The study’s 612 cases

Table 1 Factors’ mean scores

	M	SD	Percentiles		
			25	50	75
Presence (max = 20)	12.45	3.03	10.00	12.00	15.00
Perceived usefulness (max = 20)	15.82	2.82	14.00	16.00	18.00
Perceived ease of use (max = 20)	15.03	2.76	13.00	15.00	17.00
Motivation (max = 20)	13.68	3.23	12.00	14.00	16.00
Realism (max = 20)	12.56	2.54	11.00	13.00	14.00
Enjoyment (max = 20)	13.66	3.00	12.00	13.00	16.00
Tests’ scores/Learning outcomes (max = 42)	25.59	7.72	19.00	26.00	32.00

were considered very good (Comrey and Lee 2013) and even satisfied Everitt's (1975) strict rule for at least 20:1 sample to variable ratio. The data were suited for factorial analysis because (a) the kurtosis and skewness indices were small and below the recommended level of |10| and |3| respectively, as suggested by Kline (2005), (b) the Kaiser-Meyer-Olkin Measure of Sampling Adequacy index was .912, (c) the Bartlett's Test of Sphericity was significant ($p < .001$), and (d) the extraction communalities were above .5, as shown in Table 2 (Hair et al. 2006).

Table 2 Exploratory factor analysis

Item	Factor loadings							Extraction communalities
	Enj	PEU	Real	Pre	PU	Score	Mot	
Enj4	.89							.74
Enj2	.88							.84
Enj1	.78							.62
Enj3	.75							.59
PEU4		.87						.65
PEU1		.76						.64
PEU2		.74						.66
PEU3		.74						.65
Real2			.87					.70
Real1			.74					.54
Real4			.71					.58
Real3			.70					.54
Pre2				.86				.75
Pre3				.78				.61
Pre1				.78				.62
Pre4				.76				.60
PU1					.84			.71
PU2					.81			.72
PU3					.78			.63
PU4					.69			.61
Score2						.96		.93
Score1						.96		.98
Score3						.91		.88
Mot1							.87	.67
Mot2							.82	.72
Mot3							.78	.75
Mot4							.70	.68
Eigenvalues	1.41	2.88	2.07	1.80	1.35	1.14	1.02	
% variance explained (Total 68.76)	37.50	9.41	6.50	5.40	4.11	3.25	2.59	
Cronbach's α Total = .916	.87	.89	.87	.90	.84	.90	.97	

Extraction Method: PAF. Rotation Method: Oblique. Values < .30 are omitted for clarity of presentation. Enj: Enjoyment, PEU: perceived ease of use, PU: perceived usefulness, Pre: presence, Real: realism, Mot: motivation, Score = tests' scores/learning outcomes

For assessing the underlying structure of the twenty-seven items (the twenty-four questionnaire's items, plus the three tests), principal axis factor analysis (PAF) with oblique rotation was selected, because (a) PAF accounts for the covariation among variables (Kline 2005), and (b) oblique rotation produces more accurate results when the research involves human behavior (Costello and Osborne 2005). As hypothesized, seven factors were extracted (the questionnaire's six factors, plus the learning outcomes) using (a) the scree test (Costello and Osborne 2005) (Fig. 3) and (b) Kaiser's (1960) criterion (eigenvalue >1). No variables were dropped, all items loaded high on their respective factors ($> .6$), each factor averaged above the .7 level, there were no significant cross-loadings between items, and there were no correlations between the factors greater than .7 (Hair et al. 2006) (Table 2). A 68.76% of the total variance was explained by the seven components, which was considered very good. The reliability of all constructs was assessed using Cronbach's alpha and it was found to be between .84 and .97, well above DeVellis's (2003) recommendations ($> .70$). Due to the above, the internal consistency of the constructs and of the overall scale was considered more than satisfactory.

CFA was performed by imputing the factor structure that emerged during the EFA into AMOS 25. The convergent validity was considered satisfactory because (a) the standardized estimates were acceptable (ranging from .71 to .99), (b) the items explained at least half of the variance of the latent variable they belonged (ranging from .52 to .98), (c) the average variance extracted (AVE) for each factor was above the .50 level, and (d) the measurement items loaded with significant t -values on their constructs (Table 3) (Hair et al. 2010). For evaluating the questionnaire's discriminant validity, the square root of the AVE for any given factor, as well as the correlations between this factor and all other factors were calculated. The discriminant validity was considered satisfactory because, in all cases, the variance a construct shared with its measures was more than the variance this factor shared with any other factor (Fornell et al. 1982) (Table 4).

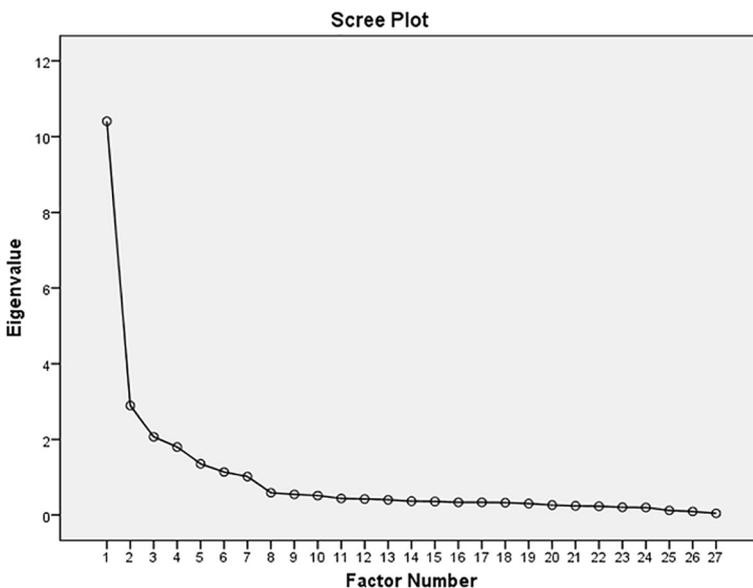


Fig. 3 Scree plot of the eigenvalues

Table 3 Results for the measurement model

Item	SE	<i>t</i> -value	<i>R</i> ²	AVE
Enj4	.87	31.25	.69	.68
Enj2	.95	21.59	.73	
Enjη1	.73	21.28	.71	
Enj3	.73	–	.64	
PEU4	.72	19.09	.88	.62
PEU2	.84	18.76	.98	
PEU1	.82	2.63	.93	
PEU3	.75	–	.66	
Real2	.81	18.10	.60	.58
Real1	.72	19.32	.74	
Real4	.77	18.64	.71	
Real3	.74	–	.60	
Pre2	.87	2.88	.61	.64
Pre4	.76	21.88	.57	
Pre3	.78	21.60	.75	
Pre1	.78		.55	
PU2	.84	24.98	.59	.68
PU1	.86	22.62	.52	
PU3	.78	22.75	.66	
PU4	.81	–	.57	
Score2	.96	73.01	.68	.93
Score1	.99	52.64	.71	
Score3	.94	–	.52	
Mot1	.78	23.44	.53	.69
Mot2	.84	23.87	.54	
Mot3	.86	22.98	.90	
Mot4	.83	–	.75	

This value was fixed at 1.00 for model identification purposes

SE, standardized estimate; *AVE*, average variance extracted

In conclusion, the results of the EFA and CFA confirmed the questionnaire's factorial structure and indicated that it had more than satisfactory validity and reliability.

6.2 Initial model testing

Since the study was based on perceptual processes that were measured at one point in time, Common Method Variance (CMV), the variance which is caused by the measurement method, was a concern (Podsakoff et al. 2003). For ruling out its presence, two tests were conducted (a) Harman's single-factor analysis (Podsakoff and Organ 1986), and (b) the comparison of the standardized regression weights before and after the addition of a common latent factor (CMV) (Gaskin 2013). It was found that CMV

Table 4 Convergent and discriminant validity

Factor	CR	AVE	Enj	PEU	Real	Pre	PU	Score	Mot
Enj	.89	.68	.82						
PEU	.87	.62	.41	.79					
Realism	.85	.58	.39	.49	.76				
Pre	.87	.64	.07	.13	.14	.80			
PU	.89	.68	.51	.58	.54	.22	.82		
Scores	.98	.93	.56	.58	.40	.17	.64	.96	
Mot	.90	.69	.58	.72	.53	.10	.67	.63	.83

CR, Critical ratio; AVE, Average Variance Extracted

Diagonal in bold: square root of the AVE extracted from the observed variables. Off-diagonal: correlations between constructs

was not an issue because (a) Harman's single-factor analysis was <50% (36.51%), and (b) the standardized regression weights were not that different (difference < .1).

All the fit indices of the initial model appeared to be within the recommended range of values, with the exception of χ^2 (Table 5). On the other hand, χ^2 usually indicates significant differences when the sample size exceeds 200 cases (Hair et al. 2006), as in this study ($N = 612$). Therefore, this outcome was considered acceptable.

6.3 Structural equation modeling

For testing the initial research model (Fig. 1), SEM was performed using AMOS 25. Before conducting the analysis, it was checked whether the multivariate assumptions were met. It was found that (a) there were no influential cases and outliers, (b) abnormal Cook's distance was not an issue (all cases < .025), (c) a curve estimation for all the relationships in the model revealed that, in some cases, linearity was slightly lower than the strongest relationship between variables but it was still significantly linear, and (d) multicollinearity was not an issue because the highest VIF value that was observed was 2.28, well below the recommended maximum of 3, while Tolerance was far greater than the recommended minimum of .1 (O'Brien 2007).

Table 5 The initial fit indices of the research model

	Result	Recommendation	Reference
χ^2	$\chi^2 (300, N = 612) = 732.08, p < .001$	ns	Schumacker and Lomax 2010
χ^2/df	2.44	1–3	Kline 2005
SRMR	.03	< .05	McDonald and Ho 2002
TLI	.96	$\geq .95$	Hu and Bentler 1999
NFI	.94	> .90	Bentler and Bonett 1980
RMSEA	.049	< .05	McDonald and Ho 2002
CFI	.97	$\geq .95$	Hu and Bentler 1999

ns, not significant

When conducting SEM, it is suggested to use control variables. Gender and age were used as such, making the assumption that they had an impact on all the other factors. The results of the SEM analysis are presented in Table 6.

Since there were some non-statistically significant effects, and others were rather weak, the removal of these paths may result in a more robust yet simpler model. For examining this, a hierarchy of $2^{24} = 16,777,216$ models was formed by making all the direct effects optional. These models were analyzed using the Specification Search Facility available in AMOS 25. The final model was the one with the smallest BCC_0 value ($BCC_0 = 6.65$) (Burnham and Anderson 1998). The fit statistics for the final model remained satisfactory [$\chi^2(353, N = 612) = 810.27, p < .001, \chi^2/df = 2.30, SRMR = .035, TLI = .96, NFI = .94, RMSEA = .046, CFI = .96$]. Indeed, in this model, some paths were removed, leading to the rejection of five out of the initial twenty hypotheses, as shown in Table 7 and Fig. 4. The final model also included two more paths, that of age to enjoyment and perceived usefulness, because age proved to have a significant impact on these factors.

6.4 Mediation

Mediation models provide a more accurate picture of the effects the independent variables have on the dependent. Hayes's (2013) bootstrapping technique was used and it was found that (a) although realism did not directly affect the learning outcomes, it had indirect effects through perceived usefulness, motivation, and pleasure, (b) perceived ease of use had both a direct and an indirect impact on the learning outcomes (through enjoyment and motivation), (c) presence influenced the learning outcomes only indirectly through perceived usefulness, (d) enjoyment besides its direct impact on

Table 6 Direct effects in the proposed model

Path	Path coefficient (β)	<i>t</i> -value	<i>p</i>
Enj \leftarrow PEU	.304	5.984	< .001
Enj \leftarrow Real	.269	5.350	< .001
Enj \leftarrow Age	-.155	-3.990	< .001
Mot \leftarrow PEU	.514	1.923	< .001
Mot \leftarrow Enj	.305	8.237	< .001
Mot \leftarrow Real	.163	4.053	< .001
PU \leftarrow PEU	.106	1.964	.050
PU \leftarrow Real	.182	4.269	< .001
PU \leftarrow Pre	.119	3.619	< .001
PU \leftarrow Enj	.182	4.419	< .001
PU \leftarrow Mot	.369	6.018	< .001
PU \leftarrow Age	.138	4.339	< .001
Scores \leftarrow PEU	.197	3.828	< .001
Scores \leftarrow Enj	.266	6.674	< .001
Scores \leftarrow PU	.283	5.780	< .001
Scores \leftarrow Mot	.166	2.768	.006

Non-significant effects are omitted

Table 7 Hypothesis testing results

Hypothesis	Path	Path Coefficient (β)	<i>t</i> -value	<i>p</i>
H1a	Real → Pre	.15	3.14	.002
H1b	Real → Enj	.27	5.37	< .001
H1c	Real → Mot	.16	4.05	< .001
H1d	Real → $\Upsilon\Delta M$.18	4.21	< .001
H1e	Real → Scores	Not supported		
H2a	PEU → Pre	Not supported		
H2b	PEU → Enj	.31	6.03	< .001
H2c	PEU → Mot	.51	1.94	< .001
H2d	PEU → PU	.11	1.97	.049
H2e	PEU → Scores	.21	4.12	< .001
H3a	Pre → Enj	not supported		
H3b	Pre → Mot	not supported		
H3c	Pre → PU	.12	3.68	< .001
H3d	Pre → Scores	not supported		
H4a	Enj → Mot	.31	8.24	< .001
H4b	Enj → PU	.18	4.44	< .001
H4c	Enj → Scores	.24	6.21	< .001
H5a	Mot → PU	.37	6.04	< .001
H5b	Mot → Scores	.13	2.21	.027
H6	PU → Scores	.30	6.67	< .001
*	Age → Enj	−.15	−3.95	< .001
*	Age → PU	.14	4.34	< .001

* Not included in the initial model/hypotheses

the learning outcomes, had indirect ones through motivation and perceived usefulness, and (e) motivation had both direct and indirect (through perceived usefulness) impact on the learning outcomes (Table 8).

6.5 Post-hoc statistical power

The method described by Soper (2016) was used for conducting the post-hoc power analysis for unsupported direct effects. For the six predictors of the dependent variable, an observed R^2 of .53 for this factor, a probability level of .05, and a sample size of 612, the observed statistical power was 1.00. Thus, the proposed model demonstrates an excellent power to detect the significant effects, while the non-significant effects that were observed were indeed not significant.

7 Discussion

The first thing that it is worth noting is the validity and reliability of the questionnaire used in this study. Indeed, both the EFA and CFA confirmed the above. At the same

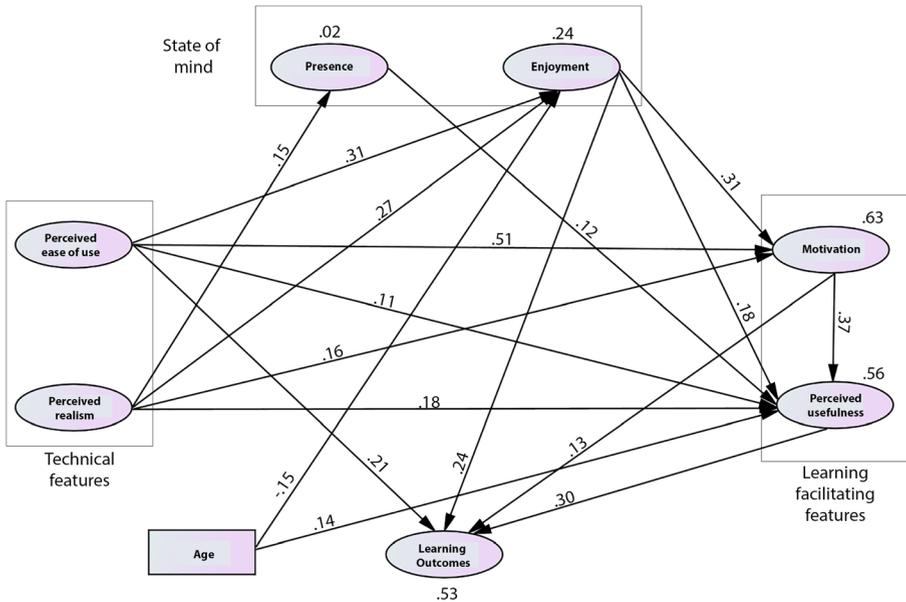


Fig. 4 The final model (non-significant paths were omitted for clarity of presentation)

time, the number of participants in which it was administered satisfied the strictest rule of thumb regarding the sample size (Everitt 1975). Consequently, it can be argued that, besides being a short one in terms of how many items it has, the questionnaire is a quite robust tool for measuring the factors included in the model.

Presence and perceived application’s realism had the lowest mean scores ($M = 12.45$ and $M = 12.56$ respectively). This finding was expected as it is almost certain that participants did not use any specialized equipment that could enhance their sense of presence when “visiting” the virtual museum. On the other hand, the mean scores of perceived usefulness and perceived ease of use were well above the mid-point ($M = 15.82$ and $M = 15.03$ respectively). It seems that the users regarded the application as easy to use, as well as a useful one. The role of these factors will be further elaborated

Table 8 Mediation results

Hypothesis	Estimate	<i>p</i>	Mediation effect	Direct effect
Real → PU → Scores	.22	.001	yes	no
Real → Mot → Scores	.10	.013	yes	no
Real → Enj. → Scores	.30	.001	yes	no
PEU → Mot → Scores	.35	0.018	yes	yes
PEU → Enj → Scores	.36	< .001	yes	yes
Pre → PU → Scores	.11	< .001	yes	no
Enj → Mot → Scores	.16	.016	yes	yes
Enj → PU → Scores	.19	.001	yes	yes
Mot → PU → Scores	.35	< .001	yes	yes

in the paragraphs to follow. As for the tests' scores, which acted as an indicator of the learning outcomes, these can be considered as satisfactory since (a) they were slightly above the mid-point ($M = 25.59$, $SD = 7.72$, $\max = 42$) and (b) they were achieved in informal settings and not in an organized educational intervention. These results confirm previous research which recognized the educational value of 3D virtual museums and indicated that indeed users were able to acquire some knowledge (e.g., Li and Zhou 2016; Naya and Ibáñez 2015; Pescarin et al. 2012).

Coming to the final model, in order its predictive power to be meaningful, it has to demonstrate high R^2 s as well as significant and substantial structural paths that are close to .20 (ideally above .30) (Chin 1988). Then again, statistically significant interaction terms below this threshold are also important (Chin et al. 2003). In the light of the above, out of the initial twenty hypotheses, fifteen were supported by the data. The learning outcomes (the dependent variable) were adequately explained ($R^2 = .53$), while perceived usefulness, enjoyment, perceived ease of use, and motivation were its most significant determinants, because the paths linking them to the dependent variable were significant ($\beta = .30$, $\beta = .24$, $\beta = .21$, and $\beta = .13$ respectively). Taking into account that the model relates informal learning with VEs, it is quite plausible that many factors, not included in the model, might have influenced the results (e.g., the participants' personality and temperament). Therefore, the fact that just four factors managed to interpret more than half of the variance of the learning outcomes, is noteworthy. Thus, it can be concluded that the model possesses the power to adequately explain the learning outcomes when individuals use VEs in informal settings. Moreover, perceived usefulness was also adequately explained ($R^2 = .56$) by motivation ($\beta = .37$), perceived realism ($\beta = .18$), enjoyment ($\beta = .18$), age ($\beta = .14$), presence ($\beta = .12$), and perceived ease of use ($\beta = .11$). The most sufficiently explained factor was that of motivation ($R^2 = .63$), which was significantly influenced by perceived ease of use ($\beta = .51$), enjoyment ($\beta = .31$), and realism ($\beta = .16$). In contrast, presence ($R^2 = .02$) and enjoyment ($R^2 = .24$) were not adequately explained. Perceived realism was the only factor -weakly- affecting presence ($\beta = .15$), while perceived ease of use ($\beta = .31$), perceived realism ($\beta = .27$), and age ($\beta = -.15$) had an impact on enjoyment.

Associating the findings of previous research with the findings of the present study is a quite difficult task, given that the proposed model examined multiple factors, in contrast to other studies where only one or just a few factors were examined (e.g., Tüzün and Özdiñç 2016; Merchant et al. 2012). Even so, the most influential factor in shaping the learning outcomes was perceived usefulness, followed by enjoyment and perceived ease of use. Motivation was also influential but its impact was less important. In addition, these factors acted as mediators of the effects of all the other constructs to the learning outcomes (see Table 8). The significant impact of perceived ease of use on the learning outcomes has been noted by other researchers (e.g., Lee et al. 2010) and this also holds true for enjoyment (Faiola et al. 2013; Wrzesien and Raya 2010). On the other hand, the direct effect of perceived usefulness to the learning outcomes is in contrast to the findings of other studies where no such effect was noted (e.g., Lee et al. 2010). Additionally, motivation was expected to have a strong direct impact on the learning outcomes as suggested by other researchers (e.g., McLellan 2004). Alas, out of all the factors affecting the learning outcomes, it had the lowest impact. Then again, it had the strongest effect on perceived usefulness ($\beta = .37$). Plausible explanations for these

inconsistencies in the findings of the present study, in relation to previous research, are the use of different models, target group differences, and the learning settings.

As already mentioned in a previous section, it was considered interesting to check whether a VE can effect motivation, since intrinsic motivation is already quite strong in informal learning (Marsick and Watkins 1990). Indeed, it was found that perceived ease of use had an impressive impact on motivation ($\beta = .51$), confirming the findings of previous research (e.g., Lee et al. 2010). Thus, it can be concluded that it is possible to enhance one's motivation to learn when the VE is not considered difficult in its use.

Enjoyment, although not adequately explained ($R^2 = .24$), seems to play a substantial role (see Fig. 3) as noted by others (Faiola et al. 2013; Harrington 2012; Wrzesien and Raya 2010). Then again, other factors that shape enjoyment need to be identified. Besides being the least explained construct ($R^2 = .02$), presence's role was minimal, having a rather weak effect only on perceived usefulness. While research indicated that it has significant effects on the learning outcomes (e.g., Bulu 2012; Hassell et al. 2012; Lee et al. 2010), this is not supported by the findings of the present study. It is worth noting that, even if presence was removed, the final model was not going to be affected that much. Thus, it is not an exaggeration to state that presence is not a factor in VEs when specialized equipment is not used, as North and North (2016) suggested. As for the perceived application's realism, it proved to be an influential factor, having direct effects on all constructs except on the learning outcomes. On the other hand, it had indirect ones through perceived usefulness, enjoyment, and, to a lesser degree, through motivation (see Table 8). While these findings confirm previous research noting the importance of realism (e.g., Dalgarno and Lee 2010; Harrington 2012; Lee et al. 2010), the absence of a direct impact on the learning outcomes is puzzling.

The model was found to be invariant across genders; males and females are expected to have similar learning outcomes when using VEs, as reported by others (Wrzesien and Raya 2010). Finally, age seems to play an interesting, yet not that strong, role. It negatively affects enjoyment and positively affects perceived usefulness. This means that the older the participants were, the more they recognized the usefulness of the VE, but the less they enjoyed using it. Given that the sample had an unbalanced age distribution, this finding has to be viewed with some caution.

The study's findings have implications for experts involved in the development of virtual museums. On the basis of the results, the impact of enjoyment needs to be enhanced even further. Thus, more attention is required on features that can increase the sense of enjoyment when one is using a VE. This can be achieved by adding game-like features, as these facilitate both enjoyment and motivation (Kozlov and Johansen 2010). Also, the sense of the VE's realism can be further developed, since it was found to have a moderate effect on almost all of the model's constructs. Then again, caution is advised, because increased realism means very detailed 3D models and equally detailed VEs. Such complex applications require advanced computers in order to run smoothly and not all users have such computers. Finally, the design of the virtual museum per se (e.g., how the exhibits are placed, how the associated information is presented, lighting, and the inclusion of activities) requires attention as these details can have an impact on the learning outcomes. That is because the design influences whether the users' attention in a museum (virtual or otherwise) is focused and active or not (Screven 1993).

8 Conclusion

Within the theoretical framework laid by previous research, a model was developed for explaining the learning outcomes when adults use VEs in informal learning settings. While the results were thought-provoking, there are limitations to the study that require some attention. First, the sample, though its size was more than adequate for statistical analysis, consisted only of Greek adults. Therefore, the results cannot easily be generalized to other samples. Moreover, the participants' socio-economic profile was not recorded in detail since there was no focus on specific groups and their age range was rather narrow. As a result, the impact of factors such as the educational level is unknown, while a more balanced age distribution might have resulted in different outcomes. Finally, the VE presented the work of an artist; totally different results might have been obtained if another theme/subject was chosen (e.g., science concepts). Future validations will establish (or reject) the model's applicability. The role of presence certainly requires an in-depth examination, since it was the least explained factor. Additional factors can be considered for inclusion that might explain the learning outcomes in VEs more thoroughly. Future studies can also examine whether the model is invariant in different participants' profiles or if it can be used in any type of VE and, thus, increase its usefulness to the scientific community.

Nevertheless, experts involved in informal learning can use this model in order to devise suitable strategies and policies for content delivery. The model can also act as a tool for measuring the learning outcomes and/or the effectiveness of VEs in various learning settings. In conclusion, the study contributes to the growing body of research on the impact of VEs in learning and it is hoped that the model will be of use to researchers and educational experts.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Appendix

Table 9 The final questionnaire

Factor	Item
Perceived usefulness/ Learning facilitator	PEL1 I feel that the VE eased the way I learn
	PEL2 The VE provided a much easier way to learn compared to other methods
	PEL3 Why use the VE? There are easier ways to learn*
	PEL4 The VE made learning more interesting
Perceived ease of use	PEU1 Learning how to use the VE was easy
	PEU2 I found the VE easy to use
	PEU3 I constantly needed help because it was not easy for me to use the VE*
	PEU4 Becoming skillful at using the VE was easy
Presence	Pr1 My immediate surroundings “disappeared” when using the VE
	Pr2 When using the VE, I often forgot where I am

Table 9 (continued)

Factor	Item	
Enjoyment	Pr3	When I used the VE, the digital world was more real than the real one
	Pr4	When I used the VE, I felt that my mind was inside the digital world
	En1	My experience using the VE was enjoyable
	En2	Everything was so boring, I did not have any fun at all*
	En3	The whole experience was entertaining
Motivation	En4	I enjoyed the VE very much
	Mo1	I had the desire to learn more about Nausica Pastra**
	Mo2	Everything was so interesting that I explored all the VE
	Mo3	I wasn't interested at all in learning using the VE*
Realism	Mo4	This type of application did not hold my attention*
	Real1	The visual display quality of the application was quite impressive
	Real2	Interaction with the virtual objects seemed like real
	Real3	Sometimes the virtual objects became more real than the real ones
	Real4	The virtual objects seemed like real

* Item for which scoring was reversed

** Replace with the learning subject of the VE

References

- Alexander, E. P., Alexander, M., & Decker, J. (2017). Museums in motion: An introduction to the history and functions of museums. Rowman & Littlefield.
- Allison, C., Campbell, A., Davies, C. J., Dow, L., Kennedy, S., & McCaffery, J., (2012). Growing the use of virtual worlds in education: An OpenSim perspective. *Proceedings of the 2nd European Immersive Education Summit*, 1-13. Paris, France.
- Ampartzaki, M., Kypriotaki, M., Voreadou, C., Dardioti, A., & Stathi, I. (2013). Communities of practice and participatory action research: The formation of a synergy for the development of museum programmes for early childhood. *Educational Action Research*, 21(1), 4–27. <https://doi.org/10.1080/09650792.2013.761920>.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3), 588–606. <https://doi.org/10.1037/0033-2909.88.3.588>.
- Bulu, S. T. (2012). Place presence, social presence, co-presence, and satisfaction in virtual worlds. *Computers & Education*, 58(1), 154–161. <https://doi.org/10.1016/j.compedu.2011.08.024>.
- Burnham, K. P., & Anderson, D. R. (1998). *Model selection and inference: A practical information-theoretic approach*. New York: Springer-Verlag. <https://doi.org/10.1007/978-1-4757-2917-7>.
- Chin, W. W. (1988). Issues and opinion on structural equation modeling. *MIS Quarterly*, 22, vii–xvi.
- Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research*, 14(2), 189–217. <https://doi.org/10.1287/isre.14.2.189.16018>.
- Comrey, A. L., & Lee, H. B. (2013). *A first course in factor analysis*. Psychology Press.
- Coombs, P. H., Prosser, C., & Ahmed, M. (1973). *New paths to learning for rural children and youth*. New York: International Council for Educational Development.
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10(7), 1–9.
- Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology*, 41(1), 10–32. <https://doi.org/10.1111/j.1467-8535.2009.01038.x>.
- Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35(8), 982–1003. <https://doi.org/10.1287/mnsc.35.8.982>.

- DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Newbury Park, CA: SAGE Publications.
- Dilli, R., & Dümenci, S. (2015). Effect of museum education on teaching extinct animals lived in anatolia to pre-school children. *Eğitim ve Bilim*, 40(181), 217–230. <https://doi.org/10.15390/EB.2015.4653>.
- Duncan, I., Miller, A., & Jiang, S. (2012). A taxonomy of virtual worlds usage in education. *British Journal of Educational Technology*, 43(6), 949–964. <https://doi.org/10.1111/j.1467-8535.2011.01263.x>.
- Everitt, B. S. (1975). Multivariate analysis: The need for data, and other problems. *The British Journal of Psychiatry*, 126(3), 237–240. <https://doi.org/10.1192/bjp.126.3.237>.
- Faiola, A., Newlon, C., Pfaff, M., & Smyslova, O. (2013). Correlating the effects of flow and telepresence in virtual worlds: Enhancing our understanding of user behavior in game-based learning. *Computers in Human Behavior*, 29(3), 1113–1121. <https://doi.org/10.1016/j.chb.2012.10.003>.
- Fokides, E. (2017). A model for explaining primary school students' learning outcomes when they use multi-user virtual environments. *Journal of Computers in Education*, 4(3), 225–250. <https://doi.org/10.1007/s40692-017-0080-y>.
- Fonseca, D., Valls, F., Redondo, E., & Villagrasa, S. (2016). Informal interactions in 3D education: Citizenship participation and assessment of virtual urban proposals. *Computers in Human Behavior*, 55, 504–518. <https://doi.org/10.1016/j.chb.2015.05.032>.
- Fornell, C., Tellis, G. J., & Zinkhan, G. M. (1982). Validity assessment: A structural equations approach using partial least squares. In B. J. Walker et al. (Eds.), *An assessment of marketing thought & practice* (pp. 405–409). Chicago: American Marketing Association.
- Gaskin, J. (2013). *SEM series part 5a: Confirmatory factor analysis*. Retrieved from <https://www.youtube.com/watch?v=MCYmyzRzNIY>
- Gerber, B. L., Marek, E. A., & Cavallo, A. M. L. (2001). Development of an informal learning opportunities assay. *International Journal of Science Education*, 23(6), 569–583. <https://doi.org/10.1080/09500690116959>.
- Hackett, A. (2014). Ziggling and zooming all over the place: Young children's meaning making and movement in the museum. *Journal of Early Childhood Literacy*, 14(1), 5–27. <https://doi.org/10.1177/1468798412453730>.
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (Vol. 6). Upper Saddle River, NJ: Pearson Prentice Hall.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis: International version* (7th ed.). New Jersey: Pearson.
- Harrington, M. C. (2012). The virtual trillium trail and the empirical effects of freedom and fidelity on discovery-based learning. *Virtual Reality*, 16(2), 105–120. <https://doi.org/10.1007/s10055-011-0189-7>.
- Harrop, D., & Turpin, B. (2013). A study exploring learners' informal learning space behaviors, attitudes, and preferences. *New Review of Academic Librarianship*, 19(1), 58–77. <https://doi.org/10.1080/13614533.2013.740961>.
- Hassell, M. D., Goyal, S., Limayem, M., & Boughzala, I. (2012). Effects of presence, copresence, and flow on learning outcomes in 3D learning spaces. *Administrative Issues Journal: Education, Practice, and Research*, 2(1), 62–73. <https://doi.org/10.5929/2011.2.1.4>.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.
- Hooper-Greenhill, E. (2013). *Museums and their visitors*. Routledge.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55. <https://doi.org/10.1080/10705519909540118>.
- Jia, D., Bhatti, A., & Nahavandi, S. (2014). The impact of self-efficacy and perceived system efficacy on effectiveness of virtual training systems. *Behaviour & Information Technology*, 33(1), 16–35. <https://doi.org/10.1080/0144929X.2012.681067>.
- Jones, G., & Alba, A. D. (2016). Reviewing the effectiveness and learning outcomes of a 3D virtual museum: A pilot study. In F. M. Mendes Neto, & R. de Souza (Eds.), *Handbook of Research on 3-D Virtual Environments and Hypermedia for Ubiquitous Learning* (pp. 168–191). IGI global. <https://doi.org/10.4018/978-1-5225-0125-1.ch007>.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20, 141–151. <https://doi.org/10.1177/001316446002000116>.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.

- Kozlov, M. D., & Johansen, M. K. (2010). Real behavior in virtual environments: Psychology experiments in a simple virtual-reality paradigm using video games. *Cyberpsychology, Behavior, and Social Networking*, 13(6), 711–714. <https://doi.org/10.1089/cyber.2009.0310>.
- Lee, E. A. L., Wong, K. W., & Fung, C. C. (2010). How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Computers & Education*, 55(4), 1424–1442. <https://doi.org/10.1016/j.compedu.2010.06.006>.
- Li, L., & Zhou, J. (2016, December). Virtual reality technology based developmental designs of multiplayer-interaction-supporting exhibits of science museums: taking the exhibit of virtual experience on an aircraft carrier in China science and technology museum as an example. *Proceedings of the 15th ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Applications in Industry-Volume 1*, 409–412. ACM. <https://doi.org/10.1145/3013971.3014018>.
- Linnenbrink, E. A., & Pintrich, P. R. (2002). Motivation as an enabler for academic success. *School Psychology Review*, 31(3), 313–327.
- Marsick, V. J., & Watkins, K. (1990). *Informal and incidental learning in the workplace*. New York: Routledge.
- Martens, R., Bastiaens, T., & Kirschner, P. A. (2007). New learning design in distance education: The impact on student perception and motivation. *Distance Education*, 28(1), 81–93. <https://doi.org/10.1080/01587910701305327>.
- McDonald, R. P., & Ho, M. R. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods*, 7(1), 64–82. <https://doi.org/10.1037/1082-989X.7.1.64>.
- McLellan, H. (2004). Virtual realities. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology*. Erlbaum Associates: Mahwah, NJ.
- Merchant, Z., Goetz, E. T., Keeney-Kennicutt, W., Kwok, O. M., Cifuentes, L., & Davis, T. J. (2012). The learner characteristics, features of desktop 3D virtual reality environments, and college chemistry instruction: A structural equation modeling analysis. *Computers & Education*, 59(2), 551–568. <https://doi.org/10.1016/j.compedu.2012.02.004>.
- Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & Education*, 70, 29–40. <https://doi.org/10.1016/j.compedu.2013.07.033>.
- Miglietta, A. M., Belmonte, G., & Boero, F. (2008). A summative evaluation of science learning: A case study of the marine biology museum "Pietro Parenzan" (south East Italy). *Visitor Studies*, 11(2), 213–219. <https://doi.org/10.1080/10645570802355984>.
- Naya, V. B., & Ibáñez, L. A. H. (2015). Evaluating user experience in joint activities between schools and museums in virtual worlds. *Universal Access in the Information Society*, 14(3), 389–398. <https://doi.org/10.1007/s10209-014-0367-y>.
- North, M. M., & North, S. M. (2016). A comparative study of sense of presence of traditional virtual reality and immersive environments. *Australasian Journal of Information Systems*, 20. <https://doi.org/10.3127/ajis.v20i0.1168>.
- Novak, T. P., Hoffman, D. L., & Yung, Y. F. (2000). Measuring the customer experience in online environments: A structural modeling approach. *Marketing Science*, 19(1), 22–42. <https://doi.org/10.1287/mksc.19.1.22.15184>.
- deNoyelles, A., Hornik, S., & Johnson, R. D. (2014). Exploring the dimensions of self-efficacy in virtual world learning: Environment, task, and content. *Journal of Online Learning and Teaching*, 10(2), 255.
- Nunez, D. (2004, November). How is presence in non-immersive, non-realistic virtual environments possible? *Proceedings of the 3rd International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa*, 83–86. ACM. <https://doi.org/10.1145/1029949.1029964>.
- O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41(5), 673–690. <https://doi.org/10.1007/s11135-006-9018-6>.
- Paliokas, I., & Sylaiou, S. (2016, September). The use of serious games in museum visits and exhibitions: A systematic mapping study. *Proceedings of Games and Virtual Worlds for Serious Applications (VS-GAMES), 2016 8th International Conference, 1-8*. IEEE. <https://doi.org/10.1109/VS-GAMES.2016.7590371>.
- Pescarin, S., Pagano, A., Wallergard, M., Hupperetz, W., & Ray, C. (2012). Evaluating virtual museums: Archeovirtual case study. *Archaeology in the Digital Era*, 74(12).
- Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management*, 12(4), 531–544. <https://doi.org/10.1177/014920638601200408>.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903. <https://doi.org/10.1037/0021-9010.88.5.879>.

- Pujol, L. & Lorente, A., (2013). The virtual museum: A quest for the standard definition. Archaeology in the digital era. *Proceedings of the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA)*, 40–48.
- Rapanotti, L., Minocha, S., Barroca, L., Boulos, M. N., & Morse, D. R. (2012). 3D virtual worlds in higher education. *Informed Design of Educational Technologies in Higher Education: Enhanced Learning and Teaching*, 212–240. <https://doi.org/10.4018/978-1-61350-080-4.ch012>.
- Sanchez-Vives, M. V., & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience*, 6(4), 332–339. <https://doi.org/10.1038/nrn1651>.
- Schumacker, R. E., & Lomax, R. G. (2010). *A beginner's guide to structural equation modeling* (3rd ed.). New York: Routledge.
- Screven, C. (1993). Museums and informal education. *CMS Bulletin*, 1(1), 36–66.
- Selwyn, N. (1997). Students' attitudes toward computers: Validation of a computer attitude scale for 16-19 education. *Computers & Education*, 28, 35–41. [https://doi.org/10.1016/S0360-1315\(96\)00035-8](https://doi.org/10.1016/S0360-1315(96)00035-8).
- Seok, K. H., & Kim, Y. S. (2014). Implementation of interactive web 3D based virtual educational system using O3D-WebGL. *International Journal of Multimedia and Ubiquitous Engineering*, 9(4), 273–284. <https://doi.org/10.14257/ijmue.2014.9.4.29>.
- Sharda, R., Romano Jr, N. C., Lucca, J. A., Weiser, M., Scheets, G., Chung, J. M., ... Sleezer, C. M., (2004). Foundation for the study of computer-supported collaborative learning requiring immersive presence. *Journal of Management Information Systems*, 20(4), 31–63. <https://doi.org/10.1080/07421222.2004.11045780>.
- Skamantzari, M., & Georgopoulos, A. (2016). 3D Visualization for virtual museum development. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 41, 961–968. <https://doi.org/10.5194/isprsarchives-XLI-B5-961-2016>.
- Soper, D. S. (2016). *Post-hoc statistical power calculator for multiple regression* [Software]. Retrieved from <http://www.danielsoper.com/statcalc>
- Sylaiou, S., Mania, K., Karoulis, A., & White, M. (2010). Exploring the relationship between presence and enjoyment in a virtual museum. *International Journal of Human-Computer Studies*, 68(5), 243–253. <https://doi.org/10.1016/j.ijhcs.2009.11.002>.
- Synodi, E. (2014). Verbal communication in museum programs for young children: Perspectives from Greece and the UK. *Childhood Education*, 90(2), 116–126. <https://doi.org/10.1080/00094056.2014.894814>.
- Tamborini, R., Bowman, N. D., Eden, A., Grizzard, M., & Organ, A. (2010). Defining media enjoyment as the satisfaction of intrinsic needs. *Journal of Communication*, 60(4), 758e777. <https://doi.org/10.1111/j.1460-2466.2010.01513.x>.
- Teo, T., & Noyes, J. (2011). An assessment of the influence of attitude and perceived enjoyment on the intention to use technology among pre-service teachers: A structural equation modelling approach. *Computers & Education*, 57(2), 1645–1653. <https://doi.org/10.1016/j.compedu.2011.03.002>.
- Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times*. Wiley.
- Tüzün, H., & Özdiñç, F. (2016). The effects of 3D multi-user virtual environments on freshmen university students' conceptual and spatial learning and presence in departmental orientation. *Computers & Education*, 94, 228–240. <https://doi.org/10.1016/j.compedu.2015.12.005>.
- Voss, G. B., Nunes, F. B., Muhlbeier, A. R. K., & Medina, R. D. (2013). *Context-aware virtual laboratory for teaching computer networks: A proposal in the 3D OpenSim environment. Proceedings of the 15th Symposium on Virtual and Augmented Reality (SVR 2013)* (pp. 252–255). Brazil: Cuiaba, Mato Grosso. <https://doi.org/10.1109/SVR.2013.46>.
- Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3), 225–240. <https://doi.org/10.1162/105474698565686>.
- Wrzesien, M., & Raya, M. A. (2010). Learning in serious virtual worlds: Evaluation of learning effectiveness and appeal to students in the E-junior project. *Computers & Education*, 55(1), 178–187. <https://doi.org/10.1016/j.compedu.2010.01.003>.
- Yaman, M., Nerdel, C., & Bayrhuber, H. (2008). The effects of instructional support and learner interests when learning using computer simulations. *Computers & Education*, 51, 1784–1794. <https://doi.org/10.1016/j.compedu.2008.05.009>.