virtual worlds

MDPI

Article

Cognitive Load During Student-Created Virtual Worlds: A NASA-TLX Assessment Using Spatial.io

Dimitrios Magetos [1,*], Sarandis Mitropoulos [2] and Christos Douligeris [1]

[1] Department of Informatics, University of Piraeus, 18534 Piraeus, Greece; cdoulig@unipi.gr
[2] School of Science and Technology, Hellenic Open University, 26331 Patras, Greece; smitropoulos@eap.gr
* Correspondence: dmagetos@unipi.gr

Abstract

Despite the increasing adoption of metaverse technologies, particularly in educational contexts, the cognitive impacts of students designing their own immersive experiences remain underexplored. This study addresses that gap by providing empirical baseline data on the cognitive load of students designing virtual worlds. All 100 participants were second-year undergraduate computer science students, aged 20 to 22, who completed a five-week virtual world development project using the Spatial.io platform. Cognitive load was assessed using the NASA Task Load Index (NASA-TLX), which is a measure of six dimensions: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. The total weighted NASA-TLX showed moderate cognitive load (M = 48.42, SD = 12.18, 95% CI [45.99, 50.85]), with a percentage of 9% of respondents indicating high cognitive load (scores > 60). Temporal Demand was the highest-rated dimension (M = 14.32, SD = 3.84), followed by Mental Demand (M = 13.68, SD = 3.52), while Physical Demand was the lowest (M = 5.23, SD = 2.94). The Spearman correlation analysis indicated that there were strong correlations between Temporal Demand and Effort ($\rho = 0.62$, $p < 0.001$) and Temporal Demand and Mental Demand ($\rho = 0.58$, $p < 0.001$) with Frustration demonstrating moderate correlations with most cognitive dimensions. In this sample, the NASA-TLX showed a good internal consistency (0.82). No statistically significant difference was found in the overall workload scores or in individual subscales ($p > 0.05$), indicating that the cognitive load experienced by male and female participants in the development of a virtual world was similar, but the small sub-sample of female participants ($n = 25$) reduces the statistical power of the comparison. These results provide a descriptive baseline for cognitive workload in virtual world creation led by students and provide evidence-based guidance into how educators can develop immersive development programs.

Keywords: virtual worlds; cognitive load; NASA-TLX; immersive learning environments; student-created content; Spatial.io; constructionism; metaverse

Check for updates

Academic Editors: Nefeli Georgakopoulou, Sotiris Diplaris, Stefanos Vrochidis, Giuseppe Amato, Maurice Benayoun and Beatrice De Gelder

Received: 12 December 2025
Revised: 26 April 2026
Accepted: 6 May 2026
Published: 11 May 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

1.1. Virtual Worlds in Higher Education

Over the past decade, virtual, immersive technologies have grown in use within the education industry, and they have proven to have significant benefits in experiential learning and multi-user interactivity [1,2]. Virtual worlds are lasting, three-dimensional, computer-generated worlds, where users socialize with each other as avatars [3]. Such environments open up educational opportunities beyond the physical limitations of a traditional classroom. Unlike e-learning, which is usually done in two dimensions, virtual

worlds engage the user in contextual three-dimensional worlds [4]. In a virtual world learners can interact with others in real time as they manipulate objects in a common space.

The platforms of virtual worlds have been embraced in educational environments all over the world, in fields such as architecture and medicine, as well as in language learning and cultural education [5,6]. Such environments have been linked to a number of learning benefits, like improvement of teamwork skills, improved student involvement, and abstract thinking compared to the conventional teaching techniques [7,8]. Moreover, the shift to remote and hybrid learning models, boosted by the COVID-19 pandemic, has also shown that virtual worlds can indeed retain the social presence and interactive learning opportunities in the distributed educational setting [9]. Ease of use of metaverse apps such as Spatial.io, Mozilla Hubs, and AltspaceVR have minimized the impediments to incorporating virtual worlds into higher education as well [10].

However, educational application of virtual worlds has primarily placed students in the role of consumers of pre-developed knowledge but not creators of virtual worlds [11]. Although there is rich research on the navigation, exploration, and interaction of learners in virtual worlds, much less focus has been directed towards student-constructed virtual worlds and the cognitive load involved in their creation [12]. The approach of constructionist pedagogy—the idea of learning through creating shareable artifacts [13]—made student-created virtual environments interesting as a subject of educational research. As students become consumers that become active creators of virtual content, they are engaged in profound cognitive activities that encompass spatial reasoning, technical problem-solving, aesthetic designing, and systems thinking [14].

Building virtual worlds demands the ability to master development platforms, make complex design choices, and combine different media elements—all of which can be cognitively challenging [15]. To effectively carry out virtual world development projects, educators need to comprehend and control these cognitive requirements [16]. Since the virtual world technologies are becoming more and more integrated into the higher education curriculum, it is timely and even necessary to explore the cognitive issues that students experience in the process of their development [17].

1.2. The Challenge of Cognitive Load in Virtual World Development

While digital environments present new educational opportunities, the cognitive demands associated with their development—particularly when students are the creators—remain insufficiently understood. Cognitive Load Theory (CLT), proposed by Sweller in 1988, provides a foundational framework for understanding how the limitations of human working memory affect learning [18]. Working memory has a limited capacity for processing new information; when instructional demands exceed this capacity, learning is impaired [19].

Developing a virtual environment demands considerably greater cognitive complexity than navigating or consuming a pre-built digital space [12,14]. Students who become creators, as opposed to consumers, should learn skills in working with design software; make choices about spatial and aesthetic style; combine multiple media; enable interactive functionality; and design the experience of a future user [12,14]. All these activities impose a burden on the working memory and, when combined, could overload the cognitive abilities in the absence of adequate instructional support [16]. Cognitive load studies on other creative online tasks—including game design, programming projects, and multimedia authoring—suggest that creation tasks consistently impose a higher workload than consumption tasks, particularly in the early stages of skill acquisition [20,21]. Unlike the passive consumption of virtual content, the act of creation requires students to translate conceptual design intentions into functional three-dimensional digital artifacts [22].

Further complications are the necessity to study special software interfaces, learn about the spatial relations in three-dimensional forms, and resolve technical problems that are inevitable in the design process [15]. New creators must learn three-dimensional design and generate sub-substantive material, which merely further increases the learning curve [11]. The overlap of these needs—learning to work with development tools and make something meaningful—introduces extraneous cognitive load that consumes resources needed in the cognitive processing to be able to effectively learn [19].

Although virtual world development is increasingly incorporated into higher education curricula, empirical evidence regarding the cognitive load students experience during this process remains scarce [12,15]. The majority of available studies assess cognitive load when consuming virtual worlds and not when creating them [9,23], which poses the great gap in the cognition of the requirements of content development. In the absence of such evidence, teachers do not have the empirical foundation to inform the design of the instruction, project timelines, and scaffolding supports required to effectively support students. This can be narrowed by the use of validated tools designed to measure the complex, multidimensional cognitive load that is inherent to creative online activities [20].

1.3. Measuring Cognitive Load: The NASA-TLX Approach

The measurement of cognitive workload in multidimensional, complicated tasks necessitates special tools based on theory and empirically validated. The NASA Task Load Index (NASA-TLX), which was created by Hart and Staveland in 1988 [24], is one of the most widely used subjective workload assessment tools across domains, including aviation, healthcare, education, and human-computer interaction [23,25]. Unlike physiological measures that require specialized equipment, or performance-based approaches that require secondary tasks and risk interfering with primary task execution, NASA-TLX provides multidimensional subjective workload ratings that can be collected immediately after task completion, requiring minimal disruption to the learning environment.

NASA-TLX evaluates workload on six scales: Mental Demand (cognitive effort needed), Physical Demand (physical activity needed), Temporal Demand (time pressure felt), Performance (perceived success of work completed), Effort (intensity of work needed), and Frustration (stress and irritation felt) [21]. Such a multidimensional approach provides a more detailed workload profile reflective of the relative impact of each type of demand on the overall workload experience.

The instrument employs a two-step process: the participants are asked to rate each dimension on a 21-point scale, followed by the pairwise comparisons that will allow identifying the relative significance of each dimension to the particular task at hand [21]. Such a weighting process acknowledges the fact that qualitatively different forms of workload might be present in each task [21]. In response to variations in workload and its correlation with performance results, NASA-TLX has been made one of the most popular and tested tools for subjective workload evaluation [22].

In educational settings NASA-TLX has been useful in assessing the cognitive load of technology-enhanced learning tasks such as virtual reality and immersive learning environments [9,23,26]. However, despite its established utility in educational technology research, NASA-TLX has not previously been applied to assess the cognitive workload of students engaged in virtual world creation—the specific gap this study addresses.

1.4. Research Gap and Study Objectives

Despite the increasing incorporation of virtual worlds into higher education and the availability of validated workload measurement tools, there remains a substantial gap in the literature. Cognitive load during consumption of virtual worlds has been studied

recently [8,11]; however, the cognitive load experienced by the students in the process of developing a virtual world has not been empirically investigated. There are a number of significant implications of this gap to research and practice. To start with, there is a lack of empirical data to guide project scope, timeframes, and teaching scaffolds when educators build courses integrating virtual world creation [12,16]. Second, the lack of workload data restricts the possibilities of educators to recognize and implement pedagogical strategies that are specifically applicable in the context of the development of virtual worlds. Third, with the growing use of platforms like Spatial.io and Mozilla Hubs in education [10], the lack of empirical cognitive load information does not allow educators to assess their appropriateness in a variety of student groups.

The proposed descriptive baseline study will offer an empirical measure of the cognitive workload of students engaged in the creation of a virtual world with students engaged in this task for the first time, using the NASA-TLX instrument. It was a cohort study involving 100 undergraduate computer science students in a five-week project on the Spatial.io platform that involved creating a virtual world. The current study is an addition to the literature since it presents empirical baseline information on the cognitive demands of the development of virtual worlds—a context that has received limited empirical attention to date. The study addresses three specific research questions:

RQ1: What levels of cognitive load, as measured by NASA-TLX, do computer science students experience when developing virtual worlds using the Spatial.io platform?

RQ2: Which dimensions of the NASA-TLX (Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration) contribute most significantly to overall cognitive workload during virtual world development?

RQ3: Are there statistically significant differences in NASA-TLX workload scores between male and female students during virtual world development?

These findings provide empirical baseline data to inform future research and curriculum design decisions related to virtual world development projects in higher education, supporting educators seeking to leverage immersive technologies while maintaining cognitive demands within a range conducive to effective learning.

2. Literature Review and Theoretical Framework

2.1. Virtual Reality and Extended Reality in Education

The application of virtual reality (VR) and extended reality (XR) technologies in education during the past 20 years has changed its initial use as an experiment to the integration into the teaching and learning processes [1,9]. XR (including virtual reality, augmented reality (AR), and mixed reality (MR)) offers interactive, immersive, and real educational experiences with affordances that are not easily found on other types of remote learning [4,23]. These technologies can mimic learning experiences that would be hard, expensive, or hazardous to recreate in the real world and create opportunities to engage in experiential learning, whether it be virtual laboratories, historical reenactment, or simulations of complex systems and collaborative design [5,7].

According to meta-analytic evidence, VR-based instruction may have moderate to large positive effects on the learning outcomes relative to traditional methods, especially in tasks that require spatial knowledge, procedural skills, and transfer of learning to real-world contexts [7]. A meta-analysis of 69 studies conducted by Merchant et al. in 2014 found that the VR environments that involved games and simulations led to better educational results ($d = 0.47$), and the effect was stronger when the learner had more control over the objects and was provided with immediate feedback [7]. More recently, Radianti et al.'s systematic review of 91 higher education immersive VR applications identified effective learning design features, including realism, interactive feedback, collaboration,

and effective scaffolding to support learners through cognitively demanding tasks [9]. These findings suggest that enhanced learning outcomes are not simply a byproduct of immersive technology itself but of effective pedagogy centered on its use [23].

Various theoretical frameworks complement the application of VR and virtual worlds in education. Constructivist viewpoints emphasize the usefulness of virtual environments in enabling students to develop knowledge by means of exploration and manipulation of the digital objects [13,27]. The social constructivist views give special attention to peer learning and knowledge co-construction by collaboration and discourse within virtual collaborative settings [9]. The Cognitive Affective Model of Immersive Learning (CAMIL) suggested by Makransky and Petersen [23] is a cognitive approach to immersive learning that explains immersive learning in terms of cognitive dimensions such as attention, spatial reasoning, and generative learning strategies and affective dimensions such as self-efficacy, motivation, and emotional engagement. The recognition of the potential presence of cognitive costs in immersion is also made in CAMIL [23].

The systematic reviews in recent years show an increasing usage of virtual worlds in the various divert disciplines of education. Applications of VR simulations are applied in health sciences and medicine to practice clinical procedures, to build empathy, and to view anatomy [5]. Architecture and engineering classes make use of virtual spaces to provide design reviews and collaborative design planning [15]. Some of the roles of language teaching programs that use digital environments include authentic communication practice, cultural immersion, and simulations of learning [2,9]. Cultural heritage agencies and museums have developed educational virtual exhibitions, and via them, students are able to access collections and heritage sites irrespective of geographic location [6]. Table 1 provides a summary of the important studies that have contributed to the knowledge of virtual worlds in education, showing the extent of its application.

Table 1. Key Virtual Reality and Virtual World Education Studies.

Study	Year	Focus	Sample
Mikropoulos & Natsis [1]	2011	Systematic review of educational virtual environments	53 empirical studies
Dalgarno & Lee [4]	2010	Learning affordances of 3D virtual environments	Conceptual analysis
Hew & Cheung [11]	2010	Review of 3D immersive virtual worlds in education	15 empirical studies
Duncan et al. [12]	2012	Taxonomy of virtual world usage in education	107 publications
Merchant et al. [7]	2014	Meta-analysis of VR-based instruction effectiveness	69 studies
Radianti et al. [8]	2020	Systematic review of immersive VR in higher education	91 applications
Mystakidis et al. [9]	2021	Social VR environments in higher education	37 studies
Makransky & Petersen [23]	2021	Cognitive Affective Model of Immersive Learning (CAMIL)	Theoretical framework with empirical validation
[28]	2022–2025	Recent study on virtual worlds or metaverse in education	TBD

Nevertheless, the literature does not adequately address the cognitive demands placed on students who create, rather than merely consume, virtual content [11,12]. Although there is extensive research on the approach and interaction of students in designed virtual environments, much less is known empirically about the cognitive demands of student-created virtual environments [12]. Pedagogically, this difference is important: constructionist theories of learning suggest that learning takes place more deeply when people create shareable artifacts [13], and in the virtual world, the process of creating the virtual world can be more valuable to learning than a form of passive consumption. But lack of systematic measurement of the cognitive requirements of creation activities leaves teachers with scanty

evidence to make instructional design choices about project complexities, time constraints, and needed scaffolds. The present study addresses this gap by providing empirical baseline data on the cognitive workload of students engaged in virtual world creation, extending the existing literature beyond consumption-focused research toward the underexplored domain of student-led content development.

2.2. Cognitive Load Theory

Cognitive Load Theory (CLT) was initially introduced by Sweller in 1988, which can be used as a basic building block to consider the implication of the human cognitive architecture on learning and problem-solving [18]. Active working memory that processes and manipulates information has a small capacity and is able to hold a few new elements at a time [18,19]. The long-term memory, on the other hand, is capable of holding an essentially infinite number of knowledge structures or schemas that can be referred to and used in a variety of situations [18,19].

One of the key concepts of CLT is cognitive load classification into three types, which, combined, form the overall cognitive effort needed at any stage in the course of learning. **Intrinsic load** is calculated by the complexity of the material in itself—that is, the number of interacting elements in the material that have to be processed at the same time [18,19]. **Extraneous load** uses up cognitive resources but does not make a difference to learning and should be avoided by designing effective instruction [19]. **Germane load** is a constructive cognitive work aimed at creating and automating new schemas. Good instructional design aims to control intrinsic load, reduce extraneous load, and help maintain germane load to enable a meaningful acquisition of schema [18,19].

The implications of CLT for the design of technology-enhanced learning environments, such as virtual world development, are important. The various cognitive demands that are imposed on students involved in creating virtual worlds are multiple and simultaneous [15,16]. The complexity of handling three-dimensional spatial relationships, coupled with the integration of various media elements at the same time, gives rise to intrinsic load [15,16]. Extraneous load can be created by unintuitive software interfaces, absence of instructional support, and technical difficulties that do not pertain to the overall learning goals [16]. The positive cognitive processes that result in germane load are designing virtual environments and considering design options [22].

Using CLT with technology-enhanced learning has generated a number of evidence-based teaching mechanisms, such as chunking up complicated information, worked examples to lessen unneeded problem-solving needs, deleting duplicative material, and adaptive scaffolding that diminishes with developing mastery [29]. Nevertheless, the use of such methods of immersive three-dimensional authoring environments is little explored. Three-dimensional environments designed with a specific purpose of instruction pose particular cognitive demands not found in other types of instruction, such as real-time spatial navigation, control of three-dimensional objects, and control of a variety of design features simultaneously [30].

This study measures overall subjective workload using NASA-TLX instead of trying to separate the intrinsic, extraneous, and germane load factors. Whereas CLT offers a good theoretical background to comprehend the cognitive demands, NASA-TLX presents a validated multidimensional scale of student workload in six interdependent dimensions: Mental Demand, Temporal Demand, Effort, Frustration, Physical Demand, and Performance [24]. The Discussion (Section 5.2) discusses the mapping of NASA-TLX results to the tripartite framework of CLT.

2.3. Cognitive Theory of Multimedia Learning in Immersive Environments

The Cognitive Theory of Multimedia Learning (CTML) was created by Mayer and his associates to study the way people learn using a combination of words and images in consideration of the principles of cognitive load [22]. CTML is based on three main assumptions: (1) the dual-channel assumption, according to which humans process information in two channels (separate visual and auditory); (2) the limited capacity assumption, according to which each channel is capable of processing only a limited amount of information at a time; and (3) the active processing assumption, according to which meaningful learning requires the learner to select, organize, and integrate incoming information with existing knowledge [22].

Though CTML was initially designed to be used in two-dimensional multimedia presentation, later studies have explored its use in three-dimensional and immersive space [23]. The virtual worlds are not confined to the conventional multimedia since they enable their users to control movement through the space, interact with objects that move, and have the perception of depth and spatial feedback [4,8]. The CAMIL framework directly addresses these considerations, implying that immersive virtual reality has cognitive benefits, including better spatial processing, embodied cognition, and more attention, and cognitive costs, including disorientation, attentional depletion, and overload because of complex interfaces, compared to traditional desktop multimedia [23]. The quality of instructional design and the degree to which the aspects of immersion are aligned to the learning objectives determine the ratio of these advantages and expenses.

In CTML terms, virtual world creation is particularly a cognitively difficult task. The construction of a virtual world entails the concomitant combination of different design elements—spatial structure, graphic design, interactivity, navigation flows, and a multiplicity of media—into a platform having its own cognitively challenging interface [12,15]. The concept of design coherence is especially hard to manage when dealing with inexperienced designers that do not have enough experience to determine what is a vital design element or a redundant design element. This can be attributed to the rule of temporal contiguity, according to which similar narration and animation should be displayed simultaneously to reduce the cognitive load required to combine the two in the mind [22]. Learners must be coherent in two or more of these dimensions of design simultaneously, making the task even more challenging when it comes to creating virtual worlds [31].

It is crucial to note that CTML is not the only theoretical perspective that can be used to instruct learning in an immersive setting; others include the constructivist views on learning, situated cognition theories, and embodied learning. The current research is not a direct test of the CTML principles; instead, it relies on the concepts of CTML as a subset of a larger theoretical construct to learn the importance of cognitive workload measurement in virtual world development situations. Through empirical recording of the cognitive load that students bear when undertaking creation tasks, this research gives information to be used in the future to apply CTML and the theories that are inspired by it to pedagogy through immersive technologies. Section 5.2 further explains how the principles of CTML can be applied to the observed patterns in the workloads in this study.

2.4. NASA-TLX in Educational and Virtual Reality Contexts

The NASA Task Load Index (NASA-TLX) is an index that was created in the aviation sector but has since been confirmed in various fields, such as health, human-computer interaction, and educational technology [24,25]. Its multidimensional design lies in the understanding that the work cannot possibly be effectively summarized in a single measure in a global fashion due to the different task demands that cannot be effectively summarized, such as mental demands, physical demands, temporal demands, performance demands,

effort demands, and emotional demands on work, among others, contributing to the total workload measurement in a different fashion [24]. The psychometric validity and sensitivity of NASA-TLX in a broad spectrum of task settings have been established by extensive research [25]. In particular, NASA-TLX will be applicable in new technology environments where physiological measurement would prove to be obtrusive or impractical [25].

In educational settings, NASA-TLX has proven valuable for evaluating the cognitive demands imposed by technology-enhanced learning activities. The NASA-TLX has been highly flexible and adaptable, allowing researchers to study workload in various educational technologies, such as computer-based simulations and online collaborative assignments [23,25].

In studies involving virtual reality and immersive environments, NASA-TLX has become a preferred instrument for workload assessment, as it can be administered electronically after task completion, allowing uninterrupted immersion during the task itself [9,23,26]. Makransky et al. have demonstrated that various levels of cognitive workload can be generated with the help of VR environments depending on the level of difficulty of the tasks, the level of expertise of the learner, and the quality of instructional design [23]. Based on their empirical findings, an adequately designed immersive VR system can lead to an equivalent workload to a task performed on a desktop computer, and poorly designed VR-based environments can cause excessive workload through disorientation or disastrous interfaces or sensory oversaturation [23]. These findings emphasize the applicability of empirically measured workloads in the de-designing of educational VR and in its evaluation.

Although NASA-TLX has been extensively used to investigate workload when VR technology is used to teach, the use of this tool to investigate the student-led creation of a virtual world has not been reported in the literature. This stands out as a significant distinction, since the act of creating and designing interactive three-dimensional spaces involves a significantly different and certainly more cognitive process than is the act of navigating or consuming an existing three-dimensional space. Students who used NASA-TLX to create a virtual world are a novel contribution to the educational technology literature and will be used to form a baseline of workload data to inform future research in the area.

2.5. Synthesis and Research Gap

The reviewed literature provides some of the fundamental basis of the current study. To begin with, meta-analytic and systematic review data supports the educational value of virtual reality (VR) and extended reality (XR), especially in spatial learning, the acquisition of procedural skills, and an increased degree of learner engagement [7,8]. Second, cognitive load and multimedia learning theories describe the impact of the constraints of the working memory in complex technological systems on learning and the necessity to control intrinsic, extraneous, and germane processing to design effective instructional design [16,18,19,22]. Third, NASA-TLX is a validated, multidimensional, and reliable tool to measure subjective workload, and its sensitivity has been reported in educational and VR settings [23–25].

Although research exists on the use of virtual worlds for navigation and interaction within pre-built environments, there is limited empirical investigation of the cognitive load experienced by students when creating their own virtual worlds—a gap that has meaningful consequences for both research and educational practice [11,12]. Pedagogically, the constructionist theory promotes learning by creating a virtual world as a more cognitively stimulating experience than passive consumption, but there is limited empirical evidence on the cognitive requirements of creating tasks [13]. Educators are increasingly utilizing the offered virtual world platforms such as Spatial.io and Mozilla Hubs [10],

which further underscores the relevance of empirical data of workloads to guide their application in pedagogy. In theory, the cognitive load of the long-term creative process of the three-dimensional design spaces is supposed to differ greatly from the short-term, task-oriented VR experiences that have dominated earlier studies [23]. Figure 1 shows the theoretical relationship and conceptual framework of the study.

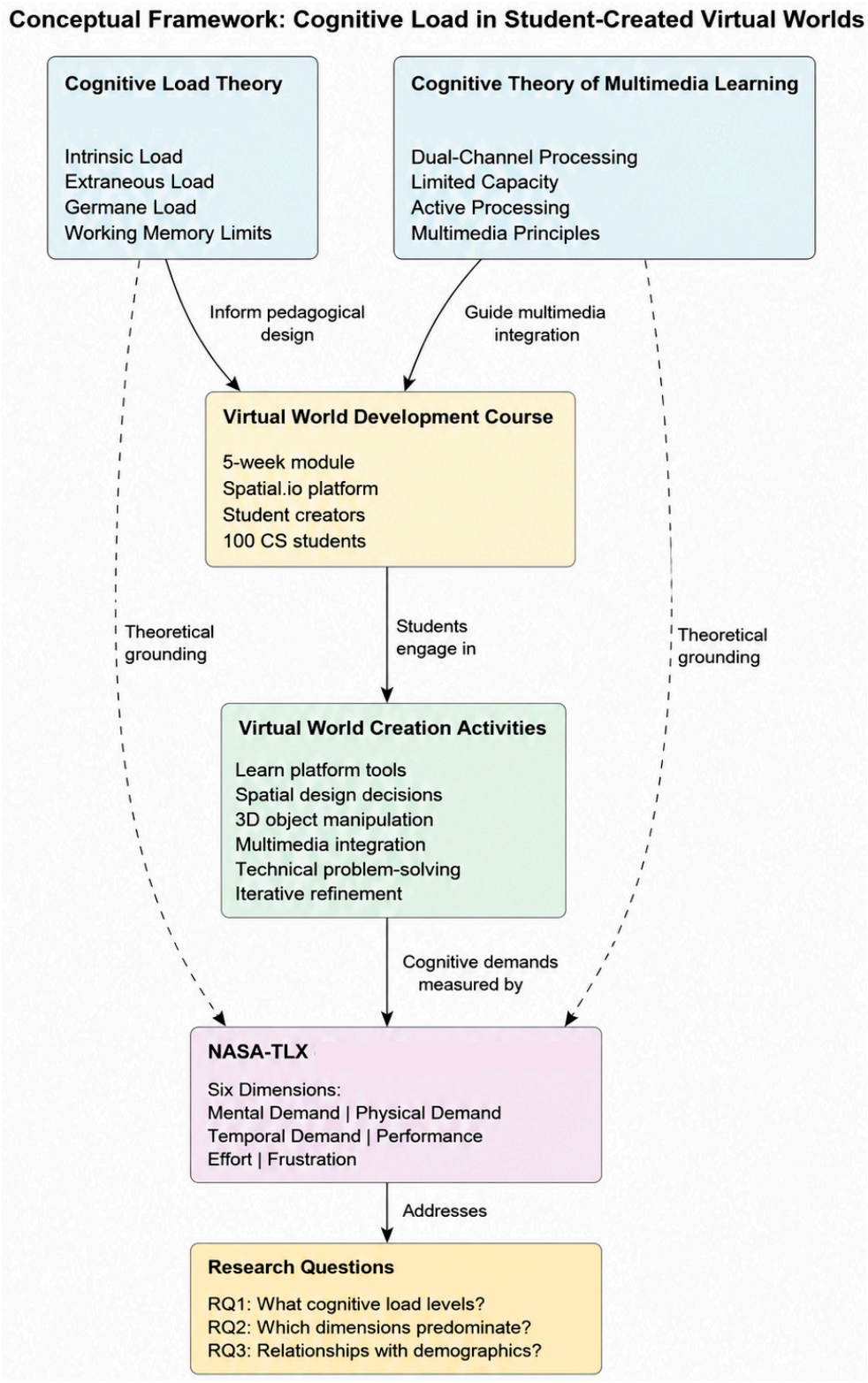


Figure 1. Conceptual framework that shows the relationships between the key theoretical and methodological constructs of the study. Solid arrows (→) are used for direct instructional or measurement

relationships, and dashed arrows (- - →) are used for theoretical grounding. The pedagogical design of the virtual world development course is based on the Cognitive Load Theory [18] and the Cognitive Theory of Multimedia Learning [22] and the cognitive demands of the course are measured with the NASA-TLX [24].

The current research will fill this gap by offering one of the first empirical evaluations of the cognitive load that students face when developing a virtual world.

Based on NASA-TLX data of 100 undergraduate computer science students after a five-week virtual world development project on Spatial.io, this research paper provides baseline cognitive workload data, reporting which NASA-TLX dimensions—Mental Demand, Temporal Demand, Effort, Frustration, Physical Demand, and Performance—were the most notable contributors to workload during student-led virtual world development. The Discussion (Section 5.2) reexamines the theoretical frameworks presented in this review in the context of the empirical results, projecting the patterns of workload observed on the tripartite load classification of CLT.

3. Materials and Methods

3.1. Study Design

The cross-sectional design was descriptive with a one-group observational design to assess subjective cognitive workload using the NASA Task Load Index after a five-week virtual world development project. This design was chosen to gather empirical baseline information that relates to an emerging educational practice where there is limited empirical evidence on cognitive workload [8,24]. The research is clearly defined as a descriptive baseline study; it aims at describing and characterizing cognitive workload in this new educational setting and not testing hypotheses or developing causal relationships.

Even though the single-group design does not have a control or comparison condition, it fits the exploratory objectives of the study to document the general levels of cognitive workloads and which NASA-TLX dimensions were most salient among the participants in the virtual world development [25]. Hart observes that early uses of NASA-TLX in a new area must focus on description and characterization of the area and then move on to hypothesis-based comparative designs [25]. The present study thus offers a reference point to future research that can look at variations across platforms, teaching methods, or student groups.

It is important to acknowledge that the absence of a control group limits causal inferences and comparative conclusions. This research does not claim to establish whether the workload is greater than that of other instructional strategies, and neither does it determine whether certain instructional strategies may decrease cognitive load. These are clearly recognized in Section 5.3, and results are interpreted cautiously, since the workload data are specific to a particular context—a five-week Spatial.io development project with undergraduate computer science students—and should not be extrapolated to other populations without further research. However, these results are adequate to support the main objective of the research, which is to provide empirical evidence about the cognitive accessibility of virtual world creation tasks—a research area where there is currently little empirical evidence [12,15].

3.2. Participants

The sample size was 100 undergraduate students in the Department of Informatics at the University of Piraeus, who were recruited through the Internet Technologies course. The sample was selected among the students enrolled in the spring semester offering of the course because it is a mandatory course for all second-year students in the computer science program. The virtual world development project was a mandatory course activity, but the

NASA-TLX questionnaire completion to conduct research was completely voluntary and not dependent on course evaluation. In the course documentation and orally, students were told that their involvement or lack of involvement in the research component would not affect their academic grades.

Among the 110 enrolled students, informed consent to participate in the research component was given by all. Nevertheless, ten students were later not included in the final analysis: five of them failed to finish the virtual world development project on the Spatial.io platform, and five of them provided incomplete responses to the NASA-TLX questionnaire. Listwise deletion was used to apply these exclusions to all participants, resulting in a final analytical sample of N = 100 participants who had complete data on all study variables. Figure 2 shows how participants were recruited and excluded.

It is important to note that the previous experience with the relevant technologies was evaluated informally during the first week of the module instead of being evaluated using validated psychometric measures; this weakness is discussed in more detail in Section 5.3.

3.2.1. Demographic Characteristics

Out of the 100 respondents who made the final analytical sample, 75 were male and 25 were female. This gender distribution is in line with recorded enrollment trends in computer science programs, where male students are usually the majority [32]. Each participant had taken introductory courses in programming, data structures, and web development, which gave them a technical background on which the virtual world development tasks were based. Table 2 gives full demographic characteristics of the sample.

Table 2. Demographic Characteristics of Study Participants (N = 100).

Characteristic	Category	*n*	%	M (SD)	Range
Gender	Male	75	75.0	—	—
	Female	25	25.0	—	—
Age (years)	—	—	—	20.8 (0.6)	20–22
	20 years	32	32.0	—	—
	21 years	51	51.0	—	—
	22 years	17	17.0	—	—
Academic Year	Second year	100	100.0	—	—
Program	Computer Science	100	100.0	—	—
Course	Internet Technologies	100	100.0	—	—

As all participants were enrolled in a computer science program, they had basic knowledge of programming, including core concepts such as variables, control structures, functions, and fundamental concepts of object-oriented programming. Gaming platform, three-dimensional design tool, and virtual environment experience differed significantly among participants, which were recorded in an informal assessment during the first week of the module (Table 3). Admittedly, this informal measure is not a validated measure of prior experience. This meant that previous experience could not be included as a variable in the statistical analyses, and RQ3 is thus restricted to gender comparisons. This methodological shortcoming is addressed in Section 5.3.

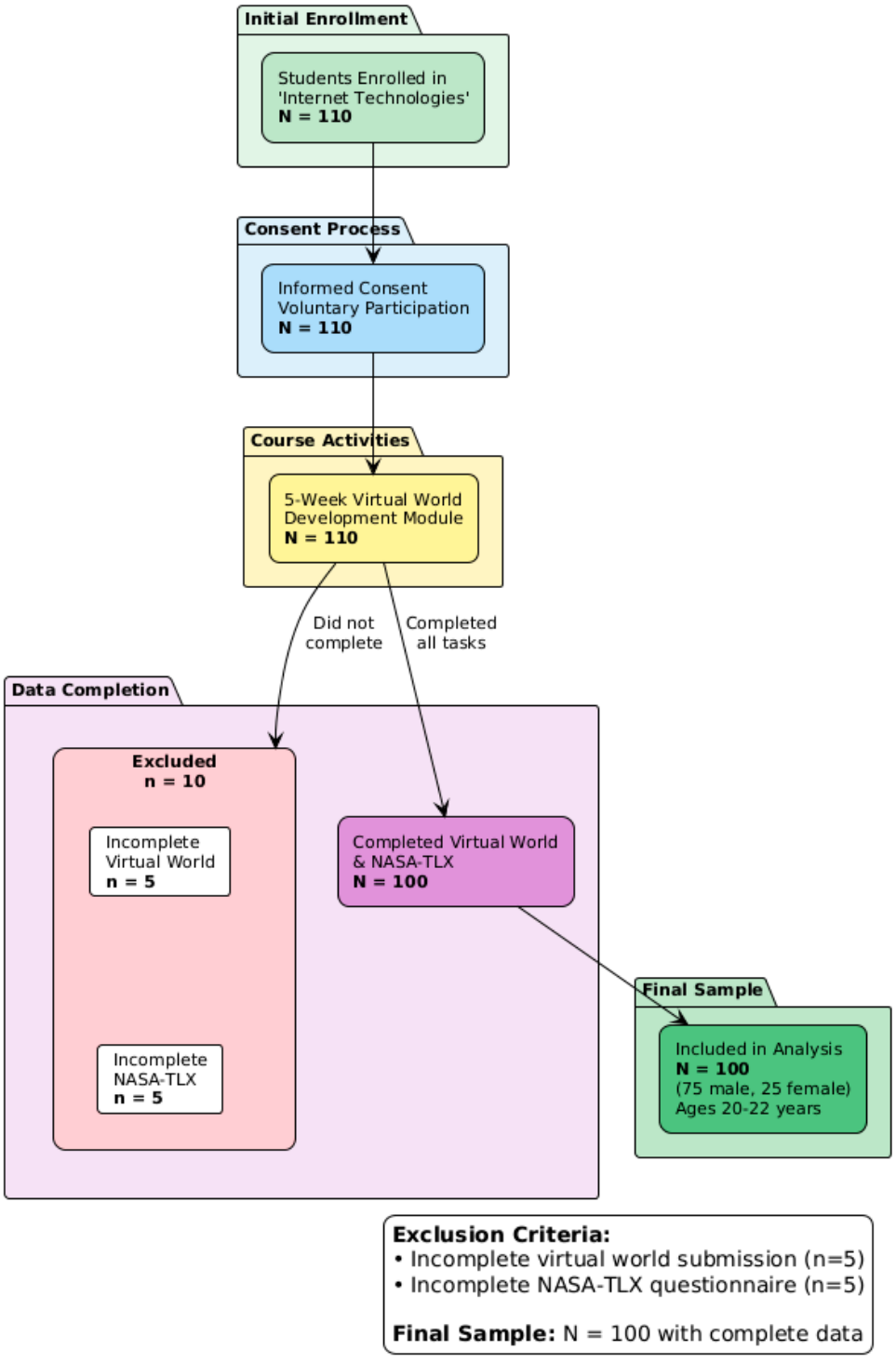


Figure 2. Flowchart of participant recruitment and exclusion. Among 110 students taking the Internet Technologies course, informed consent was given by all. After the five-week virtual world development module, 10 students were dropped because of missing virtual world submissions (n = 5) or missing NASA-TLX submissions (n = 5), which resulted in a final analytical sample of N = 100 (75 male, 25 female), aged 20–22 years.

3.2.2. Inclusion and Exclusion Criteria

Inclusion and exclusion criteria were defined prospectively before data collection commenced and applied consistently across all participants. Inclusion criteria were intentionally broad, requiring only enrollment in the Internet Technologies course and voluntary consent to participate in the research component. No exclusions were made on the ba-

sis of demographic characteristics, prior experience, or technical skill level, as the study aimed to capture the workload experience of a naturalistic student cohort. Post-hoc exclusions were applied at the data analysis stage using listwise deletion for participants with incomplete data.

Table 3. Prior experience and technical skills of study participants (N = 100). Values marked with ~ are estimates based on informal assessment conducted during Week 1 of the module.

Experience Domain	Level of Experience	*n*	%	Description
Programming Knowledge	Proficient	100	100.0	data structures, algorithms, web development
Gaming Platforms	Experienced	~80	~80.0	Prior use of Roblox, Minecraft, or similar games
	Limited/None	~20	~20.0	Minimal or no gaming experience
3D Modelling/ Design Tools	Some familiarity	~30	~30.0	Exposure to Unity, Unreal Engine, Blender
	No experience	~70	~70.0	No formal training in 3D content creation
Virtual Reality/XR	Limited exposure	~15	~15.0	Experience with VR & AR applications
	No experience	~85	~85.0	No prior VR/AR usage
Spatial.io Platform	No experience	100	100.0	No students had used Spatial.io before

Table 4 summarizes the inclusion and exclusion criteria applied during participant selection and data analysis.

Table 4. Summary of participant inclusion and exclusion criteria.

Stage	Criteria Type	Criteria/Rationale	*n* Affected
Initial Enrollment	Inclusion	Enrolled "Internet Technologies" course, Spring semester	110
	Inclusion	Voluntary consent for data collection	110
	Exclusion (a priori)	No demographic, experience, or skill-based exclusions	0
Data Analysis	Exclusion (post-hoc)	Incomplete virtual world submission on Spatial.io	5
	Exclusion (post-hoc)	Incomplete NASA-TLX questionnaire responses	5
Final Sample	—	Complete data on all measures	100

Note: All 110 enrolled students met all inclusion criteria simultaneously. Post-hoc exclusions reduced the analytical sample to N = 100.

3.2.3. Ethical Considerations

This research was conducted in accordance with the principles of the Declaration of Helsinki and the General Data Protection Regulation (GDPR).

The research was conducted after approval by the ethics committee of the University of Piraeus. Throughout the study, ethical principles were carefully observed and implemented as described below.

Participants were clearly informed that completion of the NASA-TLX questionnaire was voluntary and that non-participation would have no impact on their academic standing or course grade.

The following measures were taken to protect data. A pseudonymization approach was utilized, whereby the numerical coding systems were used to avoid identification of individual participants. All data of the research were stored in a secure encrypted format in password-protected university servers, and only the research team had access to them. The research dataset did not include any directly identifiable information (e.g., names, student identification numbers, or contact details). The participants were well informed about the exact research purposes that the data would be used and stored and the confidentiality

that would be provided to the responses. The data management and protection procedures were applied in accordance with GDPR requirements [33].

3.3. Setting: The 5-Week Educational Module

The virtual world development project was carried out in the spring semester of the Internet Technologies course at the University of Piraeus as part of the laboratory portion of the course, where students were tasked with designing, developing, and testing a functional virtual world using the Spatial.io platform. The weekly structure of the module is summarized in Table 5.

Table 5. Structure of the five-week virtual world development module conducted during the spring semester Internet Technologies course.

Week	Duration	Activities	Deliverables
Week 1	5 h (2-h lecture + 3-h lab)	Introduction to metaverse, Web 3.0; Spatial.io platform demonstration; navigation and basic tools tutorial;	Practice exercises completed; platform account created
Week 2	Independent work (approximately 6–8 h per week)	Access to educational materials via eClass platform	Design planning document (informal)
Week 3	Independent work (approximately 6–8 h per week)	Continued development; office hours available for technical support	Work-in-progress submission (optional)
Week 4	Independent work (approximately 6–8 h per week)	Refinement and multimedia integration; troubleshooting	Near-complete draft
Week 5	Independent work (approximately 6–8 h per week)	Final adjustments and submission; NASA-TLX administration	Final virtual world submission; NASA-TLX questionnaire completion

Note: Weeks 2–5 Students were working individually or in small groups during the time outside of the regular classes. The university eClass learning management system made instructional materials available. Weekly office hours were used to provide technical support. The NASA-TLX questionnaire was conducted at the end of the final virtual world submission at the end of Week 5.

The only formally planned face-to-face teaching in the module was the laboratory session in Week 1. Weeks 2–5 were designed as independent work time, and students were supposed to spend around 6–8 h per week working on their projects in the virtual world. All learning materials such as video tutorials, written instructions, and template environments were accessible via the eClass learning management system of the university. There were optional office hours once a week to offer technical and design assistance, but no formal attendance was recorded. The NASA-TLX questionnaire was completed electronically through Google Forms when the final virtual world project was submitted at the end of Week 5, as per Hart and Staveland's recommendation to measure workload immediately after completing a task [24].

3.4. Materials

3.4.1. Spatial.io Platform

The 2023 free educational workspace level of the Spatial.io platform was used in this study. Spatial.io is a WebXR-compatible, browser-based platform that allows users to create, customize, and share three-dimensional virtual spaces without the need to use high-performance hardware or install special software [34].

During the five weeks of development, students used the following features of Spatial.io: three-dimensional scene composition and layout; asset library access (360-degree photos, three-dimensional objects, videos, and audio); custom asset import in .glb, .fbx, .png, or .mp4 formats; object interaction implementation (click events and proximity triggers); spatial audio; multiplayer functionality; environmental controls (lighting Project

completion did not need custom coding or sophisticated SDK integration, so all participants, no matter their level of programming experience, were on the same technical footing.

Students used their personal computers—either laptops or desktops—which might have added variability to platform performance and, by extension, the cognitive load experience [15]. The minimum system requirements were 8 GB RAM, a stable internet connection (minimum 5 Mbps download/2 Mbps upload; recommended 20 Mbps download/5 Mbps upload), a modern web browser (Chrome, Firefox, or Edge), and a latency of less than 70 ms. VR headsets were not mandatory, with all participants completing their projects through desktop browser interfaces.

3.4.2. NASA Task Load Index (NASA-TLX)

The NASA Task Load Index (NASA-TLX) was used to measure cognitive workload; it is a multidimensional tool that has been validated and is detailed in Section 1.3. NASA-TLX is a measure of six dimensions: Mental Demand (cognitive effort required), Physical Demand (physical activity required), Temporal Demand (time pressure experienced), Performance (perceived task completion success), Effort (intensity of work invested to achieve performance), and Frustration (level of stress and irritation experienced). Each of the six dimensions is rated on a 21-point scale (0 being very low/none and 20 being very high).

The full weighted form of NASA-TLX was used in this study, given in two phases. During the first stage, the participants rated the six NASA-TLX dimensions on a 21-point scale as they considered their experience in developing a virtual world. The second stage involved participants responding to 15 pairwise comparison questions (e.g., Which contributed more to workload: Mental Demand or Temporal Demand?), where they chose one dimension at a time. The weighting coefficients were assigned to each dimension (0 to 5) based on the number of times the dimension was chosen in the 15 pairwise comparisons, which showed the relative significance of each dimension to the total workload. The total weighted workload score was calculated by multiplying the rating of each dimension by the weighting coefficient, adding the products of these, and dividing by 15, resulting in a final score between 0 and 100 [24].

The NASA-TLX questionnaire was introduced in Google Forms, where each of the six rating scales was introduced as a 21-point slider with verbal anchors at both ends and 15 pairwise comparison items. Each subscale had concise contextual definitions to aid in understanding by the participants. Table 6 provides the NASA-TLX dimensions as operationalized in the context of the virtual world development.

Table 6. NASA-TLX dimensions operationalized for the virtual world development context.

Dimension	Scale Anchors	Operational Definition in Context
Mental Demand	Low (0)–High (20)	Cognitive effort required to learn Spatial.io tools, plan spatial layouts, make design decisions, and troubleshoot technical problems
Physical Demand	Low (0)–High (20)	Physical effort from mouse/keyboard manipulation, prolonged sitting, and repetitive actions during extended development sessions
Temporal Demand	Low (0)–High (20)	Time pressure to complete virtual world within five-week deadline while managing other academic responsibilities
Performance	Perfect (0)–Failure (20)	Self-assessed success in creating a functional, aesthetically pleasing virtual world meeting assignment requirements
Effort	Low (0)–High (20)	Intensity of work needed to achieve performance level, including learning curve, iteration, and refinement efforts
Frustration	Low (0)–High (20)	Stress, irritation, or discouragement from technical difficulties, platform limitations, or gaps between intentions and outcomes

Pilot Testing: One week before the actual study, a pilot test was done on five students who had done similar projects in virtual worlds. The questionnaire was easy to understand and comprehend by the pilot participants, but some of them needed to clarify some of the terms, including Temporal Demand. Based on pilot feedback, concise contextual definitions and sample prompts were added to each subscale (e.g., "How time-pressured were you to deliver?"). The average time to complete was eight minutes, which proved the administration process to be feasible and that the questionnaire did not overburden the respondents.

Administration Timing: The NASA-TLX was conducted directly after the last virtual world submission in Week 5, which is in line with the recommendation of Hart and Staveland to measure workload right after task completion to reduce retrospective bias to the lowest extent possible [24]. The questionnaire was distributed to students via a Google Forms link shared on the eClass platform.

3.5. Procedure

The research process was combined with the normal coursework of the Internet Technologies course. In Week 1, students were introduced to the concepts of the metaverse, Web 3.0 technologies, and the pedagogical situation of creating a virtual world in a two-hour lecture. It was succeeded by a three-hour laboratory session during which students were demonstrated how to navigate the Spatial.io interface, add and move objects, and add simple interaction capabilities. Students were given introductory practice tasks and made their own Spatial.io accounts. Additional learning resources—tutorial videos, quick-start instructions, and template virtual environments—were provided via the eClass learning management system of the university.

Weeks 2–5 were self-directed, where students worked on their virtual world projects on their own. Instructional materials and technical documentation were available to students on the eClass platform, and students were supposed to spend about 6–8 h a week in virtual world development. There was no formal tracking of actual time spent on development, which is a methodological weakness of the interpretation of the temporal demand findings. Office hours were offered weekly to offer technical and design assistance, but attendance was not registered.

In Week 5, students completed and submitted their final virtual world projects. The final deliverable was a fully operational published virtual world on the Spatial.io platform, available through a shared URL, with the design and multi-media components created during the five-week process. Students were instructed to fill out the NASA-TLX questionnaire online using a Google Forms link shared on the eClass platform when submitting it. The questionnaire was completed voluntarily and required an average of eight minutes to complete, as determined in pilot testing. All the answers were anonymous and were gathered with the help of numerical participant codes, and no personal information was noted.

3.6. Data Analysis

All statistical analyses were conducted using IBM SPSS Statistics, version 28.0 (IBM Corp., Armonk, NY, USA), with an alpha level of 0.05. Data were filtered before primary analyses on completeness, outliers, and distributional properties.

Descriptive Statistics: The overall weighted and unweighted NASA-TLX scores, and the six subscales, were computed using means, standard deviations, and 95% confidence intervals. The subscales were ranked in terms of mean score to determine the main contributors of overall workload. Demographic variables were also computed and put into perspective using descriptive statistics and published NASA-TLX ranges in educational technology settings where possible [25].

Reliability Analysis: The Cronbach alpha coefficient was used to determine the internal consistency of the NASA-TLX in this sample, and the value of 0.70 and above was regarded as acceptable in research [35]. Correlations between items and total were also analyzed to determine whether there were subscales that showed weak correlations with the overall workload score.

Distributional Assumptions: The distributional properties of NASA-TLX subscale scores were tested with Shapiro–Wilk tests, with additional visual inspection of Q-Q plots and histograms. Since the sample size (N = 100) and the non-normality of subjective workload data [25], non-parametric alternatives were ready to conduct inferential tests in which normality was not met [36].

Correlational Analysis: The relationships between NASA-TLX subscale scores were analyzed with the help of the Spearman rank correlation coefficient (rho), which was chosen due to its ability to work with ordinal rating scale data and its ability to withstand normality violations [36]. The correlation matrix was created between all the subscale pairs (6 × 6). To adjust Type, I error inflation due to the fact that 15 pairwise correlations were tested, Bonferroni correction was used (0.05/15 = 0.0033). The effect sizes were interpreted based on the guidelines of Cohen [35]: small (below 0.30), medium (between 0.30 and 0.50), and large (above 0.50).

Subgroup Comparisons: To examine potential differences in NASA-TLX scores between gender subgroups, Mann–Whitney U tests were used to compare overall and subscale workload scores between male and female participants, as the ordinal nature of the rating scale data and unequal group sizes (n = 75 male, n = 25 female) make them the most suitable non-parametric alternative [28]. Cohen's d [35] was used to calculate the effect sizes, and the effect sizes of 0.20, 0.50, and 0.80 were interpreted as small, medium, and large effects, respectively. It should be noted that with n = 25 in the female subgroup, statistical power to detect medium effects (d = 0.50) at α = 0.05 is limited. Results of gender comparisons should be taken with caution, and the lack of statistically significant differences cannot be interpreted to mean that there is no difference.

Missing Data: As outlined in Section 3.2, listwise deletion was used in the cleaning of data, which left a final analytical sample of N = 100 participants with complete data on all study variables. There was no need to impute since the participants who had incomplete data were removed before the analysis.

Table 7 summarizes the statistical analysis plan and its alignment with the research questions, variables, statistical methods, and effect size indicators used in the present study.

Table 7. Statistical analysis plan aligned with research questions.

Research Question	Variable(s)	Statistical Method	Rationale	Effect Size/CI
RQ1: What levels of cognitive load do students experience during virtual world development?	Overall NASA-TLX (weighted, unweighted); six subscales	Descriptive statistics (M, SD, 95% CI); contextualizes findings against published NASA-TLX ranges from educational technology contexts	Establishes baseline workload levels [25]	95% CI for means
RQ2: Which NASA-TLX dimensions contribute most significantly to overall workload?	Six NASA-TLX subscales (Mental, Physical, Temporal, Performance, Effort, Frustration)	Ranking by mean scores; Spearman correlations between subscales; Cronbach's alpha	Identifies primary workload sources; examines dimensional relationships [24,25]	Spearman's ρ with Bonferroni correction (α = 0.0033)

Table 7. *Cont.*

Research Question	Variable(s)	Statistical Method	Rationale	Effect Size/CI
RQ3: Are there statistically significant differences in NASA-TLX workload scores between male and female students?	Gender (male/female); overall and subscale NASA-TLX scores	Mann–Whitney U tests for gender differences	Examines gender-based variations in workload; power limitations acknowledged due to unequal group sizes (n = 75 male, n = 25 female) [36]	Cohen's d for group differences
Additional Analyses	Internal consistency of NASA-TLX	Cronbach's alpha; item-total correlations	Validates instrument reliability in this context [35]	α coefficient
	Distributional properties	Shapiro–Wilk tests; Q-Q plots; histograms	Informs choice of parametric vs. non-parametric tests [36]	—

4. Results

4.1. Descriptive Statistics: Overall Cognitive Load

The full sample (N = 100) was calculated to obtain both weighted and unweighted overall NASA-TLX workload scores. The weighted NASA-TLX score indicated moderate cognitive load (M = 48.42, SD = 12.18, 95% CI [45.99, 50.85]), while the unweighted score was somewhat lower (M = 40.42, SD = 10.76, 95% CI [38.28, 42.56]). The difference between the weighted and unweighted scores indicates the difference in the level of importance that the participants gave to each dimension in their overall workload experience. Table 8 shows the descriptive statistics and workload category distributions of the overall NASA-TLX score.

Table 8. Overall NASA-TLX workload scores and distribution statistics (N = 100).

Measure	M	SD	95% CI	Range	Interpretation	
Weighted NASA-TLX	48.42	12.18	[45.99, 50.85]	18.50–72.33	Moderate workload [25]	
Unweighted NASA-TLX	40.42	10.76	[38.28, 42.56]	16.67–68.33	Moderate workload (lower bound)	
Difference (Weighted—Unweighted)	8.00	3.82	[7.24, 8.76]	0.00–18.50	Reflects dimensional weighting effects	
Statistic	Value	Category	Score Range	*n*	%	Color Suggestion
Mean	48.42	Low Workload	0–35	23	23%	Green
Median	47.67	Moderate Workload	36–60	68	68%	Yellow
SD	12.18	High Workload	61–80	9	9%	Orange
Min	18.50	Very High Workload	81–100	0	0%	Red
Max	72.33					
95% CI Lower	45.99					
95% CI Upper	50.85					

The weighted score of 48.42 is in the moderate range of workload. According to the guidelines of NASA-TLX interpretation, scores lower than 40 indicate low workload, 40–60 moderate workloads, and higher than 60 high workloads [25]. Properly designed immersive learning activities usually result in mean NASA-TLX scores of 35–55, based on the difficulty of the task and the expertise of the learner [23]. The 48.42 score is in line with the fact that the creation tasks of the virtual world require more cognitive load compared to passive navigation, which has a range of 35–45 [8,23].

Figure 3 **shows the distribution of weighted NASA-TLX scores.** The distribution is normally distributed with a small positive skew, which is supported by the Shapiro–Wilk test (W = 0.982, p = 0.187). Most of the participants (68) were in the moderate range of workload (36–60); 23% were in the low range of workload (<36) and 9% in the high range of workload (>60). None of the participants had very high scores on workload (>80), indicating that cognitive load was manageable to most of the participants throughout the five-week development period.

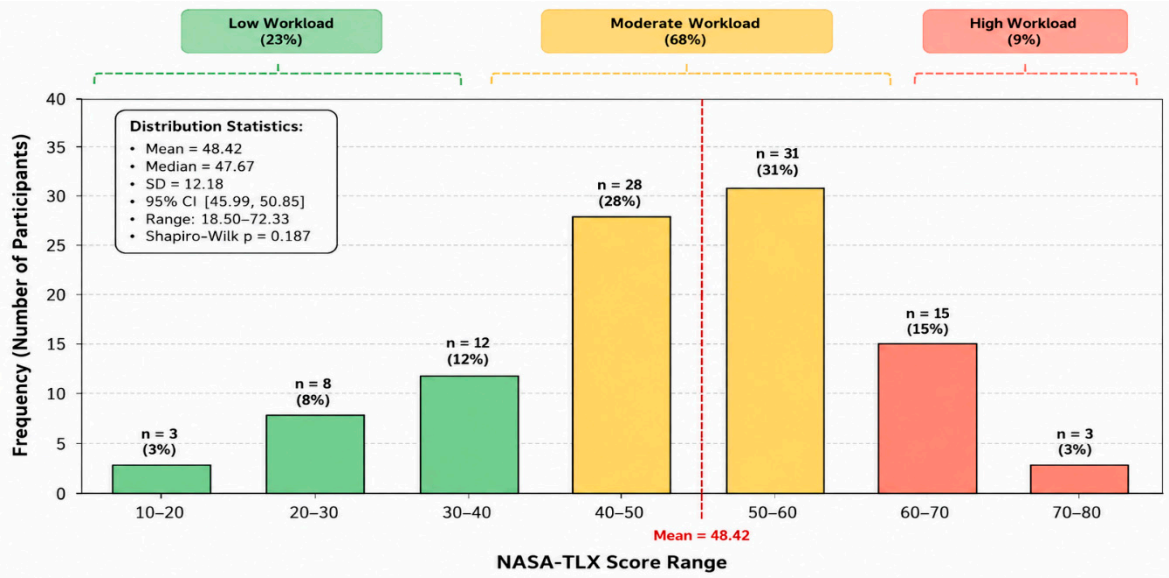


Figure 3. Weighted NASA-TLX overall workload scores frequency distribution (N = 100). The distribution is normally distributed (Shapiro–Wilk W = 0.982, p = 0.187) with a minor positive skew. Most participants (68) fell in the moderate range of workload (36–60), with 23 percent falling in the low range (<36) and 9 percent falling in the high range (>60). The broken vertical line represents the sample mean (M = 48.42).

The weighting process increased the scores by an average of 8.00 points (SD = 3.82) compared to the unweighted scores, which means that the participants rated some dimensions higher than others. This observation is in line with NASA-TLX validation studies that have shown that the weighting process enhances sensitivity by considering the relative significance of the task requirements [24,25]. Since the data of the rating scale is ordinal and the subgroups are relatively small to conduct demographic analyses, non-parametric tests were used to conduct correlational and comparative analyses, as explained in Section 3.6 [36]. Section 4.2 shows the subscale analysis, which revealed the dimensions that made the most significant contribution to the overall workload.

4.2. NASA-TLX Subscale Analysis

The six NASA-TLX subscale analyses showed that there were unique trends in the cognitive load profile of the development of a virtual world. Table 9 provides descriptive statistics of each subscale in descending order of mean score, and mean scores with 95% confidence intervals are shown in Figure 4.

Temporal Demand was the highest-rated subscale (M = 14.32, SD = 3.84, 95% CI [13.56, 15.08]), which was the main source of workload in the de-development of the virtual world. Notably, 78% of participants rated time pressure at or above the scale midpoint, indicating that participants experienced considerable time pressure throughout the development period. This observation is probably due to the demands of other coursework at the same

time, the learning curve of learning the Spatial.io platform, and the iterative process of design and refinement [15,22].

Table 9. Descriptive statistics for NASA-TLX subscales arranged in descending order of mean score (N = 100).

Rank	Dimension	M	SD	95% CI	Range	Median	Weighted Contribution *
1	Temporal Demand	14.32	3.84	[13.56, 15.08]	5–20	15.00	High
2	Mental Demand	13.68	3.52	[12.98, 14.38]	6–20	14.00	High
3	Effort	12.95	3.28	[12.30, 13.60]	5–19	13.00	Moderate-High
4	Frustration	9.84	4.12	[9.02, 10.66]	2–19	10.00	Moderate
5	Performance	8.76	3.68	[8.03, 9.49]	2–18	9.00	Moderate
6	Physical Demand	5.23	2.94	[4.65, 5.81]	0–14	5.00	Low

* Weighted contribution reflects the relative importance assigned by participants during pairwise comparisons.

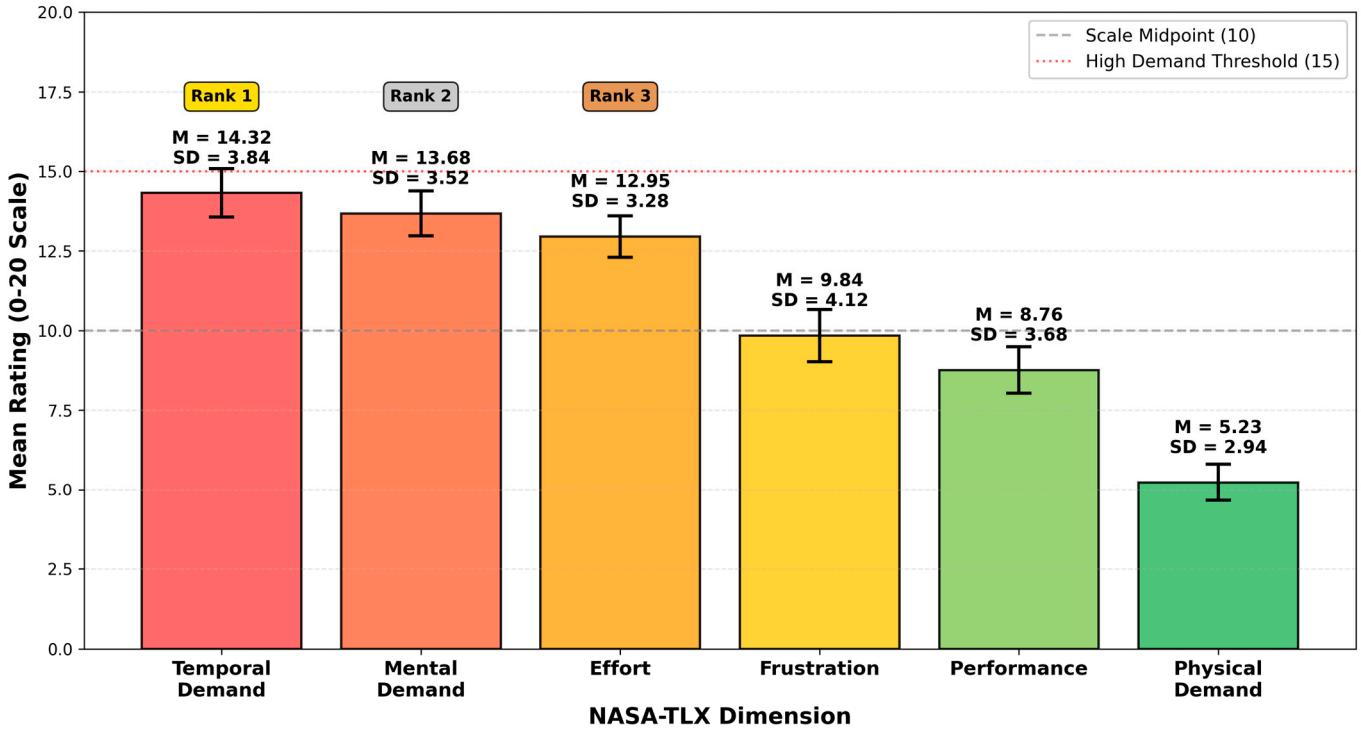


Figure 4. Mean NASA-TLX subscale scores with 95% confidence intervals (N = 100). Subscales are arranged in descending order of mean rating on the 0–20 scale. The dashed grey line indicates the scale midpoint (10), and the dotted red line indicates the high demand threshold (15). Error bars represent 95% confidence intervals. Rank labels (Rank 1, Rank 2, Rank 3) indicate the three highest-rated dimensions.

Mental Demand was the second highest-rated subscale (M = 13.68, SD = 3.52, 95% CI [12.98, 14.38]), which revealed that participants used a lot of cognitive effort in the decision-making process of spatial design, technical problem-solving, and multimedia integration. These results are in line with the CLT hypothesis that complex creative tasks produce high intrinsic load due to the interactivity of elements at the same time [18,19]. The relatively small standard deviation (SD = 3.52) indicates that mental demand was uniformly experienced by participants, which means that the cognitive challenge was task-based, but not reliant on individual differences.

The third highest-rated subscale was effort (M = 12.95, SD = 3.28, 95% CI [12.30, 13.60]) which indicated the level of effort needed to accomplish the design and development tasks. Frustration was the fourth (M = 9.84, SD = 4.12, 95% CI [9.02, 10.66]), with the larger

standard deviation indicating more variability in the emotional reactions to technical issues and platform constraints. Performance was the fifth (M = 8.76, SD = 3.68, 95% CI [8.03, 9.49]), and the participants tended to see their performance in terms of moderately successful completion of the task. The lowest-rated subscale was Physical Demand (M = 5.23, SD = 2.94, 95% CI [4.65, 5.81]), as the Spatial.io platform is browser-based and does not entail any physical activity beyond the use of a standard computer.

4.3. Reliability Analysis

The NASA-TLX had good internal consistency in this sample (0.82, 95% CI [0.77, 0.86]) [27]. It is necessary to add that internal consistency does not mean construct unity; the NASA-TLX is a multi-dimensional tool that is designed to represent a specific facet of workload [24]. The alpha of 0.82 indicates that the instrument was reliable in this sample and context and not that the six subscales are a single latent construct.

The item-total correlations ensured that all six subscales had a positive contribution to the total workload score, with the lowest correlation being r = 0.48 with Physical Demand and the highest correlation of r = 0.71 with Temporal Demand. The item-total correlation of Physical Demand is relatively lower, which is in line with the browser-based nature of the development task, which does not demand much physical activity, and the low mean score of this subscale (M = 5.23). The highest item-total correlations were found in Temporal Demand (r = 0.71) and Mental Demand (r = 0.68), which is not surprising since they were the highest-rated subscales.

4.4. Correlational Analyses

The Spearman correlation analysis revealed a number of statistically significant associations between NASA-TLX subscales after Bonferroni correction (0.0033; 15 pairwise comparisons). Physical Demand correlations were not statistically significant after Bonferroni correction, but all other reported correlations passed the corrected threshold.

Temporal Demand showed strong positive relationships with Mental Demand (rho = 0.58, $p < 0.001$, large effect), Effort (rho = 0.62, $p < 0.001$, large effect), and Frustration (rho = 0.45, $p < 0.001$, medium effect), indicating that students with higher time pressure also reported higher cognitive effort and emotional strain.

Mental Demand was significantly correlated with Effort (0.54, $p = 0.001$, large effect) and Frustration (0.62, $p = 0.001$, large effect), meaning that the higher the cognitive demand, the more effort was invested and the more the person felt frustrated.

Most cognitive and emotional dimensions were moderately to largely correlated with frustration, indicating that emotional tension was strongly associated with cognitive difficulties and time pressure during the development process. Physical Demand had weak, non-significant correlations with all other subscales (maximum: $\rho = 0.12$, $p = 0.24$), which confirms that physical exertion was mostly independent of the cognitive and emotional aspects of workload in this browser-based development scenario.

4.5. Exploratory Subgroup Analyses

Mann–Whitney U tests indicated that there were no statistically significant differences between gender groups in overall weighted NASA-TLX scores (male: Mdn = 47.50, female: Mdn = 49.33, U = 872, $p = 0.58$, d = 0.15). Likewise, individual subscale analyses did not show any statistically significant gender differences after Bonferroni correction, but female participants did report statistically higher scores on Frustration (Mdn = 11.00 vs. 9.50, U = 795, $p = 0.18$) and Mental Demand (Mdn = 14.50 vs. 13.00, U = 810, $p = 0.22$), representing small-to-medium effect sizes (d = 0.33 and d = 0.28, respectively) that did not reach statistical significance following Bonferroni correction.

The absence of statistically significant gender differences indicates that there was no significant difference in cognitive load in the development of the virtual world between male and female participants in this sample, but this result should be viewed with caution instead of being taken as a sign of similarity. The small subsample of females (n = 25) meant that there was little statistical power to identify small-to-medium effects. A post-hoc power analysis showed that at n = 25 in the smaller group, the Mann–Whitney U test has a power of about 40–50 to reject the null hypothesis with a medium effect (d = 0.50) at 80 = 0.05—far less than the traditionally regarded power of 80. The lack of statistically significant gender differences, in turn, cannot be viewed as a sign of equivalence, and future research should include gender-balanced samples to allow making more conclusive findings.

Together, the findings in this section form a descriptive foundation of cognitive workload in the context of student-led development of a virtual world, answering the three research questions in Section 1.4.

5. Discussion

5.1. Interpretation of Principal Findings

This study provides one of the first empirical evaluations of cognitive workload among students when developing a virtual world, which fills a gap in the literature on immersive educational technologies [11,12]. The NASA-TLX evaluation of 100 undergraduate computer science students working on a five-week virtual world development project on Spatial.io produced three main results. Being a descriptive baseline study, interpretations are inherently exploratory and must be understood in the context of this cohort and instructional setting.

The first principal finding is that students had moderate overall cognitive load (weighted M = 48.42, SD = 12.18), which places the development of virtual worlds in the range of workload that is normally linked with successful educational technology activities [25]. This result indicates that, in this particular teaching situation, the development tasks of the virtual world did not place an undue cognitive burden on most of the participants. It is important to note that 9% of the respondents indicated high workload (scores > 60), and no respondents indicated extreme overload (scores > 80), which indicates that cognitive load was not exceeded in levels that are typically linked to effective learning in educational technology settings [16,19]. To compare, Radianti et al. [8] found mean NASA-TLX scores of 35–55 in immersive VR learning tasks, which mainly involved navigation and interaction, whereas the current study found a weighted mean of 48.42 in creation tasks—which is within the upper range of this range and indicates that creation is a relatively more demanding task than consumption.

The second principal finding concerns the dominance of Temporal Demand (M = 14.32, SD = 3.84) as the main component of workload, as 78 percent of the participants rated time pressure as high or higher than the midpoint of the scale. The high temporal demand is probably a concurrent coursework demand, the iterative nature of creative design work that needs numerous cycles of refinement, and the learning curve of learning to use an unfamiliar platform [15,22]. Research on consumption, but not creation, of virtual worlds has always found Mental Demand to be the dominant workload com-ponent [8,23]. The difference between the temporal demand profile in this case and the mental demand profiles in consumption-oriented VR studies is a theoretically significant difference. When students are fed ready-made virtual environments, the main cognitive task is to process spatial and informational information—a mental task. When students develop virtual environments, the difficulty lies in time management in a long iterative design process and, at the same time, building platform competency. The implications of this temporal-cognitive change on the way educators plan and sequence virtual world development projects are significant.

The third main result is about Mental Demand (M = 13.68, SD = 3.52), which was second in rank and was relatively low in the variation among the participants. This regular trend of mental demand—including spatial design decision-making, technical problem-solving, and multimedia integration—indicates that the mental challenge was more task-based than individual differences. These processes are cognitively challenging and seem to be within the working memory capacity due to the moderate mean score [18,19], which is in line with Cognitive Load Theory perspective of intrinsic load as a factor of element interactivity [18].

5.2. Theoretical and Practical Implications

The findings contribute to both theoretical understanding of cognitive load in immersive learning environments and practical guidance for educators implementing virtual world development curricula.

Theoretical Contributions: This study applies Cognitive Load Theory and multimedia learning concepts to a largely understudied field of student-created immersive content [18,22]. There is a lack of empirical evaluations of cognitive load in the process of creating digital artifacts in immersive settings [13], and the current results start to fill this gap. The exceptionally low Physical Demand (M = 5.23) in comparison with the high Temporal and Mental Demand scores indicates a unique cognitive load pattern where physical activity is not a significant factor. These results are in line with the postulation of the constructionist theory that active knowledge construction involves working memory more than passive information intake [13,22]. The fact that the temporal, as opposed to the mental, demand is predominant implies that time management, and not cognitive overload as such, may be the main issue in the context of student-led virtual world creation.

Mapping the NASA-TLX findings onto the tripartite framework of CLT to gain additional theoretical understanding. The large Temporal Demand (M = 14.32) is probably due to extraneous cognitive load—pressure not due to the complexity of the design task itself but to simultaneous academic obligations and time constraints of the five-week project schedule [18,19]. The closest correlation to intrinsic load is Mental Demand (M = 13.68), which is the element of interactivity of simultaneous spatial reasoning, multimedia integration, and technical decision-making [18]. The moderate Effort scores (M = 12.95) and the low Physical Demand scores (M = 5.23) indicate that the germane load—the productive cognitive work of schema formation—was engaged without overloading working memory capacity [19,22]. The low Physical Demand validates that browser-based creation tools like Spatial.io are successful in removing motor-physical extraneous load, redirecting cognitive resources toward design-relevant processing.

The findings also offer empirical evidence of the use of constructionist pedagogy in immersive settings. Constructionism by Papert assumes that more profound learning takes place when shareable artifacts are created [13], and the moderate cognitive load (M = 48.42) indicates that the creation of virtual worlds sustains cognitive demands within the working memory capacity and maintains adequate challenge to facilitate meaningful learning [19]. In a CTML sense [22], the simultaneous control of several design factors—spatial structure, media, in-interactive functionality, and aesthetic integrity—is consistent with the dual-channel processing requirements. The moderate Mental Demand scores indicate that the available interface design of Spatial.io could have minimized extraneous processing, enabling students to allocate cognitive resources to coherent design instead of interface navigation, which is in line with the principles of coherence and signaling of CTML [22].

Practical Implications: To educators planning to develop virtual world development courses, the temporal demand finding has the most direct practical implications. A project time longer than five weeks is advisable; the temporal demand was high (M = 14.32) which

indicates that the five-week schedule placed a lot of time pressure, probably enhanced by the demands of coursework at the same time. In particular, a seven-to-eight-week schedule would enable enough iteration cycles without time pressure overwhelming the workload profile. Scheduled milestone dates—a design plan at Week 2, a prototype at Week 3, and a near-complete draft at Week 4—would spread cognitive load over the project timeline instead of clustering it around the submission deadline [15,16]. Giving clear time allocation instructions in every stage of development (planning, building, testing, and refinement) would also help students to cope with the iterative nature of the creative process [22].

The moderate Mental Demand score indicates that the available drag-and-drop interface at Spatial.io might have minimized extraneous cognitive loads related to technical complexity, which instead channeled cognitive resources to design-relevant tasks—spatial design thinking, creative problem-solving, and multimedia integration [22]. This observation justifies the use of available platform interfaces instead of pro-professional tools with higher learning curves in introducing students to the development of a virtual world in the first instance.

The lack of statistically significant gender differences in cognitive load is indicative that the development of virtual worlds can be similar between the genders in this sample, but due to the low statistical power of the gender comparison, no conclusive results can be drawn. Teachers must, however, not make assumptions about students having prior gaming or technical backgrounds, as the variability of backgrounds in students has been documented [24], and should offer fair instructional support to meet the needs of students with varying levels of prior experience.

5.3. Limitations and Future Directions

This study's findings must be interpreted within the context of several methodological limitations that constrain generalizability and causal inference.

Limitations of the Study Prediction: The single-group descriptive design was suitable to the exploratory objectives of the current study, but it does not allow making comparative or causal inferences. The study will not be able to establish the comparative workload levels with other teaching methods, other platforms, or other project times without a control group or comparison condition. The moderate work-load score observed (M = 48.42) cannot be characterized as optimal, for example, until comparative data from other platforms (e.g., Mozilla Hubs, Roblox Studio) or consumption-focused virtual world activities are available.

Study Sample Details: The sample is homogeneous (100 computer science students at one institution) and thus cannot be generalized to students in other fields and levels of education. Students of computer science probably have programming skills and technology-related self-efficacy that are not necessarily representative of students in other fields, and this could lead to a reduced cognitive load than would otherwise be the case with non-technical groups. The gender imbalance (75:25 male:female ratio), although mirroring the reported enrollment trends in computer science [24], led to a low statistical power of gender comparisons. A post-hoc power analysis revealed that at $n = 25$ in the female subpopulation, the Mann–Whitney U test has a power of about 40–50 percent to reject the null hypothesis in the case of a medium effect (d = 0.50) at an 80 percent significance level—far lower than the traditional 80 percent level [35]. Future research ought to use gender-balanced samples or use prospective power analysis to make sure that gender comparisons are adequately sensitive. The small age group (2022 years) is also a result of the cohort under investigation and does not allow generalizing to older or non-traditional students, who might have other obligations like employment or family life and have less time.

Platform Specificity: The findings are platform-specific to Spatial.io, a browser-based platform chosen due to its ease of access and low barrier to entry. The development of

virtual worlds with more professional tools that have higher learning curves, like Unity or Unreal Engine, would probably cause significantly greater cognitive load and qualitatively different workload profiles. Further studies are needed to determine whether the workload patterns here are replicated in platforms of different complexity.

Uncontrolled Variables: A set of potentially influential variables was not re-recorded or measured. The personal devices (laptops or desktop computers) of the students were of different technical specifications and internet connection speeds, which could have brought variability in platform performance and, therefore, frustration and effort ratings. The real time spent on development was not recorded in an official manner, and it was not possible to tell whether the high time requirement was due to the lack of time allocation, poor time management, or both. The first week of the module involved an informal self-report assessment of prior experience with gaming platforms, three-dimensional design tools, and virtual environments instead of validated psychometric measures. This meant that prior experience could not be included as a variable in the statistical tests, which restricted RQ3 to gender comparisons only. The next round of research must utilize validated measures to quantify previous experience in a systematic way to be able to test whether gaming background, experience in three-dimensional design, or familiarity with VR influences cognitive workload when developing a virtual world.

NASA-TLX does not separate the three types of loads of CLT (intrinsic, extraneous, and germane); thus, this research is unable to conclude how much Mental Demand is a measure of task difficulty, quality of instructional design, or both. The one post-task administration of NASA-TLX does not allow the analysis of the changes in cognitive load during the five-week development. Longitudinal designs with repeated NASA-TLX at various time points would allow establishing whether cognitive load is greatest in the early stages of learning the platforms and declines as students become familiar with the development tools.

Future Research Directions: The limitations and results of this study present several priority directions. To begin with, comparative research on workload in various platforms (e.g., Mozilla Hubs, Roblox Studio, and Unity), project timeframes, and scaffolding would determine whether the moderate workload profile observed in this case is unique to Spatial.io or can be applied to other creation environments. Second, longitudinal designs with repeated NASA-TLX administration would allow monitoring cognitive load variations during project phases and determining when peak demand occurs, which could be supported with specific instructional help. Third, mixed-method designs that would integrate NASA-TLX data with qualitative interviews would give a more detailed understanding of the exact causes of the temporal demand that students experience. Fourth, objective validation of subjective workload ratings would be provided by complementing self-reported workload data with physiological data, including heart rate variability or eye tracking. Fifth, evidence-based course design would be informed by intervention studies that test the effects of structured milestone deadlines and time management scaffolding on temporal demand. Sixth, the generalizability of these baseline findings would be tested by replication studies with diverse populations, such as students of non-technical disciplines, different educational levels, and different cultural contexts.

6. Conclusions

This study provides one of the first empirical evaluations of cognitive workload in students when developing a virtual world, which fills a gap in the literature on immersive educational technologies. This study uses the NASA-TLX to collect baseline data that defines the cognitive load involved in the implementation of virtual world development as a pedagogical activity by administering the NASA-TLX to 100 undergraduate computer

science students who completed a five-week virtual world development project on Spatial.io. Since it is a descriptive baseline study, such findings are inherently exploratory; no causal inferences or generalizations can be made without replication across a variety of populations and instructional contexts.

In relation to the three research questions, the study observed that (RQ1) students had moderate overall cognitive workload (M = 48.42, SD = 12.18), which is within the range of effective educational technology tasks [22]; (RQ2) Temporal Demand was the most pronounced workload dimension (M = 14.32), then Mental Demand (M = 13.68) and Effort (M = 12.95); and (RQ3) no statistically significant gender differences in workload were observed, though the limited statistical power of this comparison warrants cautious interpretation.

The moderate overall cognitive load has positive implications for the incorporation of virtual world development assignments in post-secondary education. This observation is in line with the constructionist pedagogy focus on learning by creating [13], which justifies the idea that the cognitive load of three-dimensional design, multimedia construction, and technical problem-solving is kept within a manageable scope under the guidance of the right constructionist pedagogical design [19,22].

Dimensional analysis showed a significant trend: Temporal Demand (M = 14.32) became the most significant workload component, unlike the previous studies on the consumption of the virtual world, where Mental Demand was always the most significant workload component [8,23]. The temporal demand is probably due to the demands of other coursework at the same time, the iterative design process, and the learning curve of an unknown platform and not an unreasonable project schedule per se [15,22]. The implication of this difference directly relates to course design: the educator must extend project timelines beyond five weeks, introduce structured milestone deadlines to spread cognitive load over the development period, and give clear time allocation guidance to each development stage: planning, building, testing, and refinement.

The good internal consistency of NASA-TLX in this regard (=0.82) justifies its applicability in workload assessment research on immersive content creation studies, which empowers enabling educators to transition to evidence-based assessment of the cognitive load of their course designs.

Future studies are needed to investigate workload in various platforms, project duration, and student groups; monitor cognitive load over time in project phases; and experiment with the impact of structured scaffolding interventions on temporal demand. The current research provides empirical baseline data on the cognitive accessibility of virtual world creation tasks and offers evidence-based recommendations to teachers who may want to use immersive technologies and balance cognitive demands within the range that facilitates effective learning.

Author Contributions: Conceptualization, D.M. and S.M.; methodology, D.M.; software, D.M.; validation, D.M., S.M. and C.D.; formal analysis, D.M.; investigation, D.M.; resources, D.M.; data curation, D.M.; writing—original draft preparation, D.M.; writing—review and editing, D.M.; visualization, D.M.; supervision, S.M.; project administration, C.D.; funding acquisition, C.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been partly supported by University of Piraeus Research Center (UPRC).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the University of Piraeus (No. 32/Meeting 2/2025–2026/13.01.2026).

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

Data Availability Statement: There are no other data other than those mentioned in the text.

Acknowledgments: Many thanks to the students for their participation in this study.

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

References

1. Mikropoulos, T.A.; Natsis, A. Educational Virtual Environments: A Ten-Year Review of Empirical Research (1999–2009). *Comput. Educ.* **2011**, *56*, 769–780. [CrossRef]
2. Pellas, N.; Fotaris, P.; Kazanidis, I.; Wells, D. Augmenting the Learning Experience in Primary and Secondary School Education: A Systematic Review of Recent Trends in Augmented Reality Game-Based Learning. *Virtual Real.* **2019**, *23*, 329–346. [CrossRef]
3. Bell, M.W. Toward a Definition of "Virtual Worlds". *J. Virtual Worlds Res.* **2008**, *1*, 1–5. [CrossRef]
4. Dalgarno, B.; Lee, M.J.W. What Are the Learning Affordances of 3-D Virtual Environments? *Br. J. Educ. Technol.* **2010**, *41*, 10–32. [CrossRef]
5. Allcoat, D.; von Mühlenen, A. Learning in Virtual Reality: Effects on Performance, Emotion and Engagement. *Res. Learn. Technol.* **2018**, *26*, 2140. [CrossRef]
6. Mystakidis, S. Metaverse. *Encyclopedia* **2022**, *2*, 486–497. [CrossRef]
7. Merchant, Z.; Goetz, E.T.; Cifuentes, L.; Keeney-Kennicutt, W.; Davis, T.J. Effectiveness of Virtual Reality-Based Instruction on Students' Learning Outcomes in K-12 and Higher Education: A Meta-Analysis. *Comput. Educ.* **2014**, *70*, 29–40. [CrossRef]
8. Radianti, J.; Majchrzak, T.A.; Fromm, J.; Wohlgenannt, I. A Systematic Review of Immersive Virtual Reality Applications for Higher Education: Design Elements, Lessons Learned, and Research Agenda. *Comput. Educ.* **2020**, *147*, 103778. [CrossRef]
9. Mystakidis, S.; Berki, E.; Valtanen, J.-P. Deep and Meaningful E-Learning with Social Virtual Reality Environments in Higher Education: A Systematic Literature Review. *Appl. Sci.* **2021**, *11*, 2412. [CrossRef]
10. Mystakidis, S.; Christopoulos, A.; Pellas, N. A Systematic Mapping Review of Augmented Reality Applications to Support STEM Learning in Higher Education. *Educ. Inf. Technol.* **2022**, *27*, 1883–1927. [CrossRef]
11. Hew, K.F.; Cheung, W.S. Use of Three-Dimensional (3-D) Immersive Virtual Worlds in K-12 and Higher Education Settings: A Review of the Research. *Br. J. Educ. Technol.* **2010**, *41*, 33–55. [CrossRef]
12. Duncan, I.; Miller, A.; Jiang, S. A Taxonomy of Virtual Worlds Usage in Education. *Br. J. Educ. Technol.* **2012**, *43*, 949–964. [CrossRef]
13. Papert, S.; Harel, I. Situating Constructionism. In *Constructionism*; Ablex Publishing: Norwood, NJ, USA, 1991; pp. 1–11.
14. Mikropoulos, T.A. Presence: A Unique Characteristic in Educational Virtual Environments. *Virtual Real.* **2006**, *10*, 197–206. [CrossRef]
15. Checa, D.; Bustillo, A. A Review of Immersive Virtual Reality Serious Games to Enhance Learning and Training. *Multimed. Tools Appl.* **2020**, *79*, 5501–5527. [CrossRef]
16. Sweller, J.; van Merriënboer, J.J.G.; Paas, F. Cognitive Architecture and Instructional Design. *Educ. Psychol. Rev.* **1998**, *10*, 251–296. [CrossRef]
17. Fernandez, M. Augmented Virtual Reality: How to Improve Education Systems. *High. Learn. Res. Commun.* **2017**, *7*, 1–15. [CrossRef]
18. Sweller, J. Cognitive Load During Problem Solving: Effects on Learning. *Cogn. Sci.* **1988**, *12*, 257–285. [CrossRef] [PubMed]
19. Sweller, J.; van Merriënboer, J.J.G.; Paas, F. Cognitive Architecture and Instructional Design: 20 Years Later. *Educ. Psychol. Rev.* **2019**, *31*, 261–292. [CrossRef]
20. Sweller, J.; Ayres, P.; Kalyuga, S. *Cognitive Load Theory*; Springer: New York, NY, USA, 2011. [CrossRef]
21. Cierniak, G.; Scheiter, K.; Gerjets, P. Explaining the Split-Attention Effect: Is the Reduction of Extraneous Cognitive Load Accompanied by an Increase in Germane Cognitive Load? *Comput. Hum. Behav.* **2009**, *25*, 315–324. [CrossRef]
22. Mayer, R.E. Cognitive Theory of Multimedia Learning. In *The Cambridge Handbook of Multimedia Learning*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2014; pp. 43–71. [CrossRef]
23. Makransky, G.; Petersen, G.B. The Cognitive Affective Model of Immersive Learning (CAMIL): A Theoretical Research-Based Model of Learning in Immersive Virtual Reality. *Educ. Psychol. Rev.* **2021**, *33*, 937–958. [CrossRef]
24. Hart, S.G.; Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology*; Hancock, P.A., Meshkati, N., Eds.; Elsevier: Amsterdam, The Netherlands, 1988; Volume 52, pp. 139–183. [CrossRef]
25. Hart, S.G. NASA-Task Load Index (NASA-TLX); 20 Years Later. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*; SAGE Publications: Los Angeles, CA, USA, 2006; Volume 50, pp. 904–908. [CrossRef]
26. Parong, J.; Mayer, R.E. Learning Science in Immersive Virtual Reality. *J. Educ. Psychol.* **2018**, *110*, 785–797. [CrossRef]
27. Vygotsky, L.S. *Mind in Society: The Development of Higher Psychological Processes*; Harvard University Press: Cambridge, MA, USA, 1978; Volume 86.

28. Tlili, A.; Huang, R.; Shehata, B.; Liu, D.; Zhao, J.; Metwally, A.H.S.; Wang, H.; Denden, M.; Bozkurt, A.; Lee, L.H.; et al. Is Metaverse in Education a Blessing or a Curse: A Combined Content and Bibliometric Analysis. *Smart Learn. Environ.* **2022**, *9*, 24. [CrossRef]
29. Clark, R.C.; Nguyen, F.; Sweller, J. *Efficiency in Learning: Evidence-Based Guidelines to Manage Cognitive Load*; Pfeiffer: San Francisco, CA, USA, 2006.
30. De Witte, B.; Reynaert, V.; Kieken, D.; Jabbour, J.; Demarey, C.; Dumoulin, A.; Possik, J. Immersive virtual reality learning and cognitive load: A multiple-day field study. *Comput. Hum. Behav.* **2025**, *176*, 108853. [CrossRef]
31. Mayer, R.E.; Fiorella, L. Principles for Reducing Extraneous Processing in Multimedia Learning: Coherence, Signaling, Redundancy, Spatial Contiguity, and Temporal Contiguity Principles. In *The Cambridge Handbook of Multimedia Learning*, 3rd ed.; Cambridge University Press: Cambridge, UK, 2021; pp. 151–172. [CrossRef]
32. Sax, L.J.; Lehman, K.J.; Jacobs, J.A.; Kanny, M.A.; Lim, G.; Monje-Paulson, L.; Zimmerman, H.B. Anatomy of an Enduring Gender Gap: The Evolution of Women's Participation in Computer Science. *J. High. Educ.* **2017**, *88*, 258–293. [CrossRef]
33. European Commission. General Data Protection Regulation (GDPR). *Off. J. Eur. Union* **2016**, *L119*, 1–88.
34. Spatial Systems Inc. Spatial.io: Collaborate in Virtual Spaces. Available online: https://www.spatial.io (accessed on 1 November 2025).
35. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; Lawrence Erlbaum Associates: Hillsdale, NJ, USA, 1988.
36. Field, A. *Discovering Statistics Using IBM SPSS Statistics*, 5th ed.; SAGE Publications: London, UK, 2018.

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