



HAL
open science

Acquiring, Analyzing and Interpreting Knowledge Data for Sustainable Engineering Education: An Experimental Study Using YouTube

Zoe Kanetaki, Constantinos Stergiou, Georgios Bekas, Sebastien Jacques, Christos Troussas, Cleo Sgouropoulou, Abdeldjalil Ouahabi

► To cite this version:

Zoe Kanetaki, Constantinos Stergiou, Georgios Bekas, Sebastien Jacques, Christos Troussas, et al.. Acquiring, Analyzing and Interpreting Knowledge Data for Sustainable Engineering Education: An Experimental Study Using YouTube. *Electronics*, 2022, 11, 10.3390/electronics11142210. hal-03737832

HAL Id: hal-03737832

<https://univ-tours.hal.science/hal-03737832>

Submitted on 25 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Article

Acquiring, Analyzing and Interpreting Knowledge Data for Sustainable Engineering Education: An Experimental Study Using YouTube

Zoe Kanetaki ¹, Constantinos Stergiou ¹, Georgios Bekas ¹, Sébastien Jacques ^{2,*}, Christos Troussas ³, Cleo Sgouropoulou ³ and Abdeldjalil Ouahabi ⁴

- ¹ Laboratory of Mechanical Design, Department of Mechanical Engineering, University of West Attica, 12241 Athens, Greece; zoekanet@uniwa.gr (Z.K.); stergiou@uniwa.gr (C.S.); gb331984@googlemail.com (G.B.)
- ² Research Group on Materials, Microelectronics, Acoustics and Nanotechnology (GREMAN), University of Tours, UMR 7347, CNRS, INSA Centre Val-de-Loire, 37100 Tours, France
- ³ Educational Technology and eLearning Systems Laboratory, Department of Informatics and Computer Engineering, University of West Attica, 12243 Athens, Greece; ctrouss@uniwa.gr (C.T.); csgouro@uniwa.gr (C.S.)
- ⁴ UMR 1253, iBrain, Université de Tours, INSERM, 37000 Tours, France; abdeljalil.ouahabi@univ-tours.fr
- * Correspondence: sebastien.jacques@univ-tours.fr; Tel.: +33-6-667-639-46

Abstract: With the immersion of a plethora of technological tools in the early post-COVID-19 era in university education, instructors around the world have been at the forefront of implementing hybrid learning spaces for knowledge delivery. The purpose of this experimental study is not only to divert the primary use of a YouTube channel into a tool to support asynchronous teaching; it also aims to provide feedback to instructors and suggest steps and actions to implement in their teaching modules to ensure students' access to new knowledge while promoting their engagement and satisfaction, regardless of the learning environment, i.e., face-to-face, distance and hybrid. Learners' viewing habits were analyzed in depth from the channel's 37 instructional videos, all of which were related to the completion of a computer-aided mechanical design course. By analyzing and interpreting data directly from YouTube channel reports, six variables were identified and tested to quantify the lack of statistically significant changes in learners' viewing habits. Two time periods were specifically studied: 2020–2021, when instruction was delivered exclusively via distance education, and 2021–2022, in a hybrid learning mode. The results of both parametric and non-parametric statistical tests showed that "Number of views" and "Number of unique viewers" are the two variables that behave the same regardless of the two time periods studied, demonstrating the relevance of the proposed concept for asynchronous instructional support regardless of the learning environment. Finally, a forthcoming instructor's manual for learning CAD has been developed, integrating the proposed methodology into a sustainable academic educational process.

Keywords: computer-aided design (CAD); educational data mining; engineering education; online and hybrid learning environments; social media analytics



Citation: Kanetaki, Z.; Stergiou, C.; Bekas, G.; Jacques, S.; Troussas, C.; Sgouropoulou, C.; Ouahabi, A. Acquiring, Analyzing and Interpreting Knowledge Data for Sustainable Engineering Education: An Experimental Study Using YouTube. *Electronics* **2022**, *11*, 2210. <https://doi.org/10.3390/electronics11142210>

Academic Editors: Agnieszka Konys and Agnieszka Nowak-Brzezińska

Received: 24 May 2022

Accepted: 13 July 2022

Published: 14 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Today, many higher education institutions are integrating learning activity data analytics into their operations [1]. Universities are recognizing the benefits of information solutions that not only better students at all stages of their education, even the most challenging, but also implement ever more effective educational resources to enhance the learning experience for students and their instructors [2]. Teaching module organizers and instructors focus on analytics results and the use of algorithms to improve their content and flexibility to identify students at risk of academic failure as early as possible and then provide them with more targeted learning solutions [3].

Long before the global health crisis due to the SARS-CoV-2 virus, nearly a decade of student education data could be used to understand key aspects of learner characteristics that would differentiate those who are capable of graduating from those at risk of dropping out. With all educational procedures going online, the acquisition of electronic data from a variety of online sources has led to an increase in all analytical procedures, and their management is a concern for the academic community [4].

It has been more than two years since the learning experience in universities was totally disrupted by the consequences of the COVID-19 pandemic, resulting in a significant decline in student retention and academic performance. Sustainable measures were then gradually put in place to ensure students' virtual presence and support them in their asynchronous tasks. Today, after almost a year and a half of exclusive and mandatory distance learning imposed by the pandemic, the learning processes of higher education have been tested in terms of applicability, feasibility and long-term sustainability. Although the health situation is not yet fully stabilized, the educational models deployed during the most critical period must be evaluated [5–8].

The first semester of the 2021–2022 academic year is undoubtedly a defining moment in the educational community. Although the post-COVID-19 period has not yet arrived, it is now appropriate to speak of the beginning of the “meta-COVID-19 period,” certainly defining a period of disruption, but offering many exciting opportunities in the academic world. Specifically, in the era of digital transformation of the educational procedure [9], the authors of [10] applied the Greek term “meta”, meaning beyond, to describe a promising future for academia, followed by a global health phenomenon. With the increasing implementation of digital learning systems during the health crisis, it is important to ensure that the system design can motivate and support active student engagement to achieve the required educational goals. Therefore, socially oriented technology tools can be incorporated into the design of online learning systems to increase student engagement and improve student performance during the learning process [11]. By adjusting learning tactics and introducing technological features applied during the epidemic into meta-COVID-19 instruction, academic institutions will be able to progress and thrive in sustainable educational models.

The work presented here consists of an experimental study that focuses on instructor monitoring of learner behaviors and engagement throughout the teaching module to assess the effectiveness and sustainability of the applied learning tactic. The methodology applied is based on video analysis and, more precisely, on the processing and interpretation of data coming from the consultation by students of online digital content directly from the reports of a YouTube channel made available as part of a computer-aided mechanical design (CAD) module [12]. The analysis of the experimental data collected, analyzed and interpreted should make it possible to evaluate the benefit of this YouTube channel dedicated to asynchronous pedagogical support in online or hybrid teaching environments. To convince the university community of the sustainable integration of the YouTube channel into the educational process, learners' viewing habits were analyzed in depth from the channel's 37 instructional videos, all in conjunction with the execution of the CAD course. To develop the results of this work, two distinct learning periods are considered: the first refers to exclusively distance learning spaces (i.e., during the most critical period of the health crisis), and the second reflects both face-to-face and mixed learning environments (i.e., mixing face-to-face and distance).

The study proposed here covers a wide range of skills: from the creation of a complete asynchronous educational and social environment based on a YouTube channel where digital observations can be acquired, to the extraction and exploitation of visualization data. The novelty of this study lies in the fact that the use of social media channels in online and hybrid learning spaces has not yet been analyzed in depth, for its sustainable integration in the academic educational procedure.

The development of this work will be articulated as follows: Section 2 will first present a review of the literature related to the objectives of this work. The methodological and

organizational aspects will be presented in Section 3, starting with the creation of the educational YouTube channel, the sources of data exploration, and proceeding to the statistical analysis of the acquired data where the results will be presented. The main results obtained will be presented in Section 4, and a discussion, based on these results, will be conducted in Section 5. Finally, the conclusions and research perspectives will be analyzed in Section 6.

2. Related Work

Long before the emergence of distance learning imposed by the COVID-19 pandemic, institutions around the world incorporated learning management systems (LMS) into their instructional schemes, whether in online or blended learning environments [13]. During the health crisis, several learning tools were used, individually or in combination. At the University of West Attica (Greece), as well as at the University of Tours (France), the Microsoft (MS) Teams learning platform was, for example, widely used for synchronous transmissions, in combination with the E-Class (or Moodle) LMS for asynchronous support. In [14], researchers proved that using a single learning platform (MS Teams), supported by a social media channel such as YouTube for asynchronous support, limited the dropout rate of learners, compared to MS Teams supported by the LMS Moodle.

Just as considering customer preferences is critical to the development of a business, in education, analyzing students' preferences and taking steps to provide them with learning materials in innovative learning spaces could be the key to improving their academic performance [15,16].

Educational data mining (EDM) seems to have a major effect in the field of education [17]. EDM and data analytics promise a better understanding of student learning, as well as new insights into the hidden aspects that influence learner performance [18]. Learning analytics (LA) is an emerging area of learning management systems that tracks and records student activities in online and virtual learning environments [19]. The role of LA is to use the data generated by students as they interact with new technology features to improve the teaching–learning process (TLP) and enable instructors to make better decisions in terms of structuring teaching modules [20]. It is generally accepted that the more data you collect, the more information you get. Therefore, the more information one acquires, the more accurate predictions, forecasts and estimates can be obtained. Data quality is very important, as it is affected by the number of variables and the amount of data acquired, which can lead to information sparsity, especially in cases where the quality of the data appears to be poor [21]. In addition, process analysis allows for the observation of unusual activities and behaviors, which can lead to the detection of “outliers”, alarm objects, and calls for intervention [22]. Given the power of the method, LA can therefore be a major feedback tool for educators and instructional designers to improve the learning experience [23].

Researchers in [24] developed a technology acceptance model (TAM) to examine which factors of social networking sites such as YouTube and TikTok can support and facilitate online knowledge acquisition. In this study, data collection was conducted using an online questionnaire on four external factors: content richness, innovativeness, satisfaction, and enjoyment. The results showed that both social networking sites contribute to knowledge sharing and acquisition. Although reports from YouTube channels were not retrieved and considered in this research, the authors concluded that to increase acceptance, the focus should be on uploaded video content.

Although the authors of [25] conducted a systematic review of the literature describing the sources and use of educational data, data analyses from the YouTube channel were unfortunately not considered. Additionally, in [26], the researchers examined the influence of instructor-generated video content on student engagement and participation in a course using the number of posts per week and the number of characters per post as parameters. They conducted an independent-sample *t*-test to compare student evaluation of the course by dividing the population into two groups: those who had been exposed to instructor-

generated videos and those who had not. The test showed statistical significance between the two groups.

In [27], the authors discussed the use of YouTube analytics both to assess student attendance during lectures and to measure the impact of lectures on the student learning experience. Going further, the authors in [28] planned campaigns by processing analytics data from the tools offered by social media channels.

The remarkable popularity of social media applications can be attributed to the encryption technology used, which ensures user privacy and limits access to personal information [28]. Social media is thought to offer benefits such as enhancing human interaction through the use of electronic media, increasing creativity, creating a sense of affiliation and acceptance, encouraging engagement and cooperative learning, reducing restrictions in terms of space and social or economic position within a community, increasing interaction and communication among members, and improving users' technological expertise [29]. With technological tools now widely available to people under the age of 20, it is possible to access social media sites with one hand, thanks to mobile technology and the use of inexpensive electronic devices such as tablets and smartphones [30,31].

Sharing videos or their URL links has become easier for instructors with the adoption of virtual learning environments (VLE). Content in teaching modules, discussion forums, and targeted sequences in online courses can be shared via YouTube video links embedded in the assessment features of learning platforms [32]. The analytics of YouTube channels can provide valuable information about learners' viewing habits, as well as measures of how students engage in the learning process with videos [33]. With YouTube analytics, instructors can thus track video viewing behaviors on supportive tasks to better understand their usefulness [34]. Previous studies in this area have shown that students do not follow a video in its entirety. They play, stop, rewind, and replay the educational content in order to review segments of the video recordings. This specific learner behavior is intended to recall the part of the newly acquired knowledge that was not clearly defined [33,35]. In [35], the authors investigated the use of pausing and searching in videos of course recordings provided by the channel interface, and how these two features relate to students' learning tactics and performance in a specific curriculum. Before the COVID-19 pandemic imposed restrictions on academia, researchers studied how learners in traditional learning environments and flipped classrooms interacted with the videos, as well as the nature of those interactions through analysis of processing data [32]. Instructors recorded short or long videos of portions of their lectures and uploaded them to VLE [32].

In transforming the traditional learning space into a virtual one, one of the most significant problems has been the loss of contact with the engineering environment itself, which is a key aspect of engineering education [36,37]. In addition, the authors of [10] associated educational data mining, processing, and analysis with the term "sustainability", which is present and promoted in most aspects of everyday life, from business operations to manufacturing and the environment [38]. Data mining, the processing of data and eventual interpretation of the results, is a fundamental process that allows researchers to establish the relationship between raw digital data and the assessment of real world conditions [39]. With digital data now available by tracking activities, analysts can get lost in unnecessary information. To facilitate the process, researchers should be able to set their boundaries, creating controlled environments for data production, i.e., targeted to the goals of their research area.

In this study, the data collected and analyzed provides insight into students' video content viewing behaviors from a dedicated YouTube channel. These behaviors were analyzed over two distinct time periods: the first in 2020–2021, i.e., during the most critical period of restrictions due to the COVID-19 pandemic, and the second in 2021–2022, during the early hours of the meta-COVID-19 period. The objective is to study the similarities between the delivery of engineering training modules in exclusively online mode and in mixed mode. This objective had already been set well before the implementation of the new pedagogical environments, whether online or hybrid; the social communication

channel was then used as an asynchronous pedagogical support allowing the collection of information directly related to students' behaviors, while avoiding the "noise" due to irrelevant details.

Given the results already available and the gaps identified by the literature review, this work is guided by the following five research questions:

- Can a YouTube channel provide adequate and quality asynchronous support on student tasks in online learning environments?
- Can we provide a guide for future instructors and organizers of CAD modules, willing to implement pedagogical methods supported by technology and social media sites such as YouTube, which would benefit the learning process in online and hybrid spaces?
- Can learners' viewing habits be revealed in online and hybrid learning spaces through information provided by YouTube?
- Do learners' behaviors and visualization patterns follow the flow of the module?
- Can educational data mining (EDM) from social media sites like YouTube provide a solid foundation for addressing student needs?

The answers to the above research questions should help demonstrate the relevance of educational data from social media channels such as YouTube to the academic community, and help institutions implement their strategies for sustainable digital transformation in higher education.

The research objective of this study is not limited to evaluating a specific YouTube channel as a tool to support asynchronous teaching. It also aims to provide feedback to instructors and suggest steps and actions to implement in their teaching modules to ensure students' access to new knowledge while promoting their engagement and satisfaction, regardless of the learning environment, i.e., face-to-face, fully remote, and hybrid.

3. Methodological and Organizational Aspects

3.1. Foreword

The methodology described in this section was deployed in a 12-week "computer-aided mechanical design (CAD I)" module, a teaching module within the Department of Mechanical Engineering of the University of West Attica (Greece). Prior to the health crisis, this module was divided between traditional mechanical design in a room equipped with drawing boards and computer-aided design with Autodesk Inventor software in a computer lab.

During the most critical hours of the COVID-19 pandemic, videos (directly downloadable into the MS Teams environment available to students) were integrated into flipped classrooms to provide asynchronous instructional support to complement the online courses delivered synchronously via the MS Teams platform. All of this work was done, including the integration of the following three tasks:

- Provide asynchronous support to students and incorporate it into assessments of individual tasks.
- Guarantee students the robustness of the learning process, including avoiding any disruption caused by multiple platforms or LMS.
- Create a virtual lab to connect distance learning spaces with real engineering environments related to students' future work [36,37,40].

During the period when pandemic restrictions were relaxed, the 12-week CAD module was divided into two main stages. For the first 4 weeks, all classes were conducted face-to-face in the classroom equipped with drawing boards. The objective of this phase was to teach students to represent views of a three-dimensional object by freehand drawings (sketches). In the second stage, conducted in the computer lab, students enrolled in activities to create three-dimensional mechanical objects in different views (top, side, and cross-section) using CAD tools. To complete these activities, learners had the option of participating face-to-face or online (the class was streamed live by the instructor via

MS Teams). In both stages, asynchronous support videos were attached to the students' assigned task. The YouTube channel "MCAD I UNIWA" was created to provide learners with all the video support needed to complete their learning independently. This YouTube channel, and all of its video content (see Appendix A), is managed by a CAD instructor with over twenty years of CAD experience. The administrator of the YouTube channel was also responsible for coordinating all activities of the various instructors in the teaching module. The MCAD I UNIWA's YouTube channel policy is public. In order to target specific tasks and associate them with their asynchronous support video links, the tasks were signed as "assignments" on the MS Teams communication platform. Each task was announced in the students' MS Teams dashboard, where instructions were provided in text form [36]. The YouTube channel links for each task were uploaded as reference material targeting the specific task to avoid confusion when searching the 37 videos to find the one relevant to the task, as shown in Figure 1. The instructor was motivated to attach the URLs of the videos to each task after considering that most viewers of today's social media channels do not easily subscribe to the channels. In this way, we were able to reach both non-subscribing and subscribing students, with the latter being immediately notified of new videos.

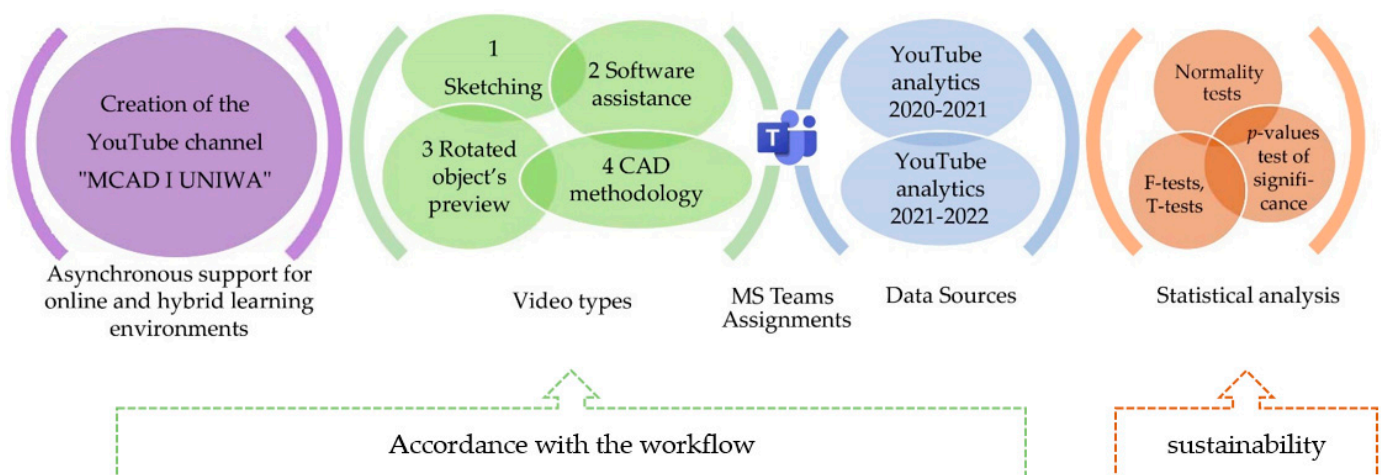


Figure 1. Flow chart illustrating the data mining and processing methodology applied.

What we seek to highlight in this manuscript is the presence of a significant relationship between learners' listening habits and their behaviors, particularly in the acquisition of knowledge and skills necessary for graduation. To do so, we draw on feedback from the CAD I module, looking in depth at data collected from the "MCAD I UNIWA" YouTube channel. To achieve this objective, we implemented the methodology described in Figure 1. The first step was the creation of a YouTube channel in which most of the instructional video content was created from screen recordings and audio recordings intended exclusively for student use. We did not use the raw recording of the full laboratory lecture because we wanted to target specific tasks, i.e., fundamental to the mechanical engineering profession, to help students perform them outside of class [41].

The uploaded videos were divided into four categories, based on the learning objectives to be achieved at each stage of the teaching process. Video categories 1 and 3 were devoid of audio, primarily to invite the instructor to explain the content presented at their own pace and in their own style. Specifically, the sketch videos showed the instructor's sketchbook, complete with pencil, as he or she drew freehand views of the object. The videos showing the model of the object being studied in three dimensions were intended to help students design the geometric shapes in all views around the object. The third and fourth categories were generated by a combination of screen recordings and audio recordings from the software modeling environment.

The YouTube channel administrator examined viewing patterns since the first video was uploaded to progressively interpret learners' needs and determine if these screen

and audio recordings met the demand for asynchronous support. During learning phases conducted exclusively online, visualization patterns could be identified by the end of the first semester. When the health regulations were relaxed, i.e., when university educational spaces were allowed to transform into hybrid learning environments (i.e., combining online and face-to-face learning), a challenge arose: to analyze the visualization patterns of learners in hybrid educational spaces and correlate the results with those of spaces conducted exclusively at a distance. Once all the data was collected, the method then consisted of filtering the data from the social media channel reports corresponding to the two distinct time periods: the first corresponding to the first semester of the 2020–2021 academic year in the exclusively distance learning environments and the second, the first semester of the 2021–2022 year in the face-to-face, virtual, and mixed learning environments. Thus, two years were necessary to collect sufficient data and to ensure the results that will be discussed in the rest of the manuscript.

The final step in the methodology is to perform statistical tests on the defined variables. First, normality tests were performed to assess the normality of the distribution. In the case where the distribution is Gaussian, parametric tests were chosen. In the opposite case, non-parametric tests were performed [42]. For determining if the variances analyzed for the two separate academic years' time periods are equal or unequal, a series of *F*-tests were performed. These tests can eventually be completed by *t*-tests depending on the equality or inequality of the variances generated by the *F*-tests.

3.2. Participant Demographics

Although not such an easy task, analysis of YouTube channel reports provides detailed information about the demographics of the participants and gives a “typical user profile” while revealing repetitive viewing behaviors. In our case, the term “user” refers to engineering students attending the computer-aided mechanical design module. Since one of the research questions focuses on developing a manual for future instructors, it is essential that module coordinators observe student activities outside the classroom and ultimately develop a profile of the typical student, which each instructor must consider, before taking steps to improve module delivery. For this reason, information about the status of the YouTube channel subscription, the type of viewing device used, as well as the preferred operating system, is collected, in addition to the standard demographic data. Note that in the YouTube studio, this type of data is available in separate tabs for each of the variables, after filtering for the two observation periods (i.e., 2020–2021 and 2021–2022). The goal of this strategy is to provide a clear picture with enough numerical data to allow for the most accurate comparisons and conclusions possible.

Accumulating demographic characteristics directly from the source, as opposed to self-reported responses in questionnaires, can increase the validity of the acquired data. It should be noted that many studies, including [43], have relied exclusively on demographic attributes to assess and even predict students' academic performance with high accuracy.

Table 1 provides a summary of the key demographic characteristics of the mechanical engineering students who participated in the study during the two time periods noted above. This summary was compiled from data directly extracted from YouTube channel reports, content available in YouTube Studio's advanced analysis mode, and by specifying a custom time period. Each metric (gender, age, geography, etc.) was exported from the YouTube Studio view to Google Sheets and converted to an MS Excel file [44].

The number of students who took the online module was 212 in the winter semester of the 2020–2021 year. In 2021–2022, Table 1 shows a slightly higher number of students (i.e., 230) who took the module in a blended learning mode. Regardless of the two time periods considered, 90% of the learners were male. This percentage reflects the true majority of male students enrolled in the Department of Mechanical Engineering at the University of West Attica, as verified by the student registry. In the hybrid learning environments, 97.6% of the students were between the ages of 18 and 24, and the total number of participants was located in Greece. Considering that, in the exclusive online learning environments,

60.9% of the students were between the ages of 18 and 24 and 37.7% were between the ages of 25 and 34, it can be inferred that the online learning spaces offered a unique opportunity for older students to attend their courses without their physical presence. In addition, the exclusive online learning spaces allowed a small number of international resident students to take the module from another country.

Table 1. Demographic information of mechanical engineering students who participated in the study.

Demographics	2020–2021 (n = 212)	Viewing Time (Hours)	2021–2022 (n = 230)	Viewing Time (h)
Gender of the participant and age distribution				
Female	10.0%		10.0%	
Male	90.0%		90.0%	
Age of the participant				
18–24	60.9%		97.6%	
25–34	37.7%		2.4%	
35–44	0.9%		0.0%	
45–54	0.5%		0.0%	
Geography				
National	97.8%	261.6	100%	174.6
International	2.2%	0.1	0%	0.0
Subscription status				
Not subscribed	75.9%	193.6	82.1%	151.4
Subscriber	24.1%	70.9	17.9%	30.2
Type of viewing device used				
Computer	75.0%	210.39	70.7%	137.4
Mobile phone	23.6%	52.91	27.7%	38.0
Tablet	0.9%	0.81	1.3%	5.6
TV	0.5%	0.32	0.3%	0.2
Operating system				
Windows	78.1%	209.4	67.2%	133.4
Android	17.7%	46.0	22.2%	31.1
iOS	3.7%	7.8	6.4%	12.5
Macintosh	0.4%	1.0	2.6%	4.4
Smart TV	0.1%	0.02	0.0%	0.0
PlayStation	0.0%	0.0	0.4%	0.2
Linux	0.0%	0.0	0.4%	0.1
WebOS	0.0%	0.0	0.4%	0.02
Xbox	0.0%	0.0	0.4%	0.01

The registration status of viewers is already an interesting factor: Only 24.1% of viewers are registered on the educational channel, indicating an initial trend that most students watch the videos repeatedly, without finding a reason to subscribe to the YouTube channel. This suggests that students have a similar attitude toward the educational channel, as probably with other social media sites.

Although personal computers were widely used in both periods studied, they are losing ground to mobile devices each year. Tablet use increased in the second period. Finally, television as a viewing device decreased in 2021–2022. Finally, the most used operating system was Microsoft Windows, but there is a 10.9% decline between 2020–2021 and 2021–2022, which is explained by the increasing use of tablets and the doubling of iOS android devices.

3.3. Reports on Students' Views, Comparative Analysis and Discussion

A total of thirty-seven videos were uploaded to the YouTube channel. Since there was no preparation time available before the universities closed, the timing of the posting of

each video was scheduled in parallel with the running of the laboratory module. The name of each video corresponds to the title of the assigned task. All videos can be distinguished by content and learning objective into the following four categories, as shown in Table 2:

1. Sketch and make freehand drawings in two dimensions of views of given objects.
2. Assist the CAD software with the basic commands needed to accomplish the assigned tasks.
3. Preview of the object to be studied that was modeled in the three-dimensional CAD modeler. This specific type of video was created not only to help students perceive shapes and geometric entities, generate multiple representations of them (views), but also to develop their optimized design principle [45].
4. Support on tasks by screen and sound recordings of the drawing procedure in the CAD environment.

Table 2. Analysis of data from the “MCAD I UNIWA” YouTube channel.

	Number of Videos	Number of Views	Average Viewing Time (s)	Average Percentage of Viewing
Video category no. 1	11	3758	727	37.4
Video category no. 2	5	2107	444	28.5
Video category no. 3	9	4777	383	67.1
Video category no. 4	12	5576	1644	29.9
Total	37	16,218		

In order to understand the actual size of the learners, the unique users’ metric was filtered from the reports provided by the YouTube channel. This specific metric provides a clearer picture of the estimated number of views during the two different time periods (i.e., during the pandemic period and during the start-up period of meta-COVID-19). The metrics for specific aspects related to viewing by video categories are presented in Figures 2 and 3. Specifically, the fourth category, relating to the methodology of carrying out computer-aided design tasks, is the most viewed and, as expected, has the longest viewing time. The videos with the highest average viewing percentage also belong to this category. The third category of videos, showing an overview of filmed objects, comes next in the students’ preferences. It should be noted that this type of video has the shortest duration, limited to ten to fifty seconds. Furthermore, three of the nine videos of this type were generated in the second period, i.e., in hybrid learning environments, after taking into account the students’ requests. The second category of videos contains the smallest number of videos due to the fact that most of the software support tools were integrated in the fourth category. This type of video has a considerable number of views in relation to the number of views. This suggests that these videos are aimed at a specific audience, who will persist in watching the content to better understand the use of the software. Finally, the first category, drawing without sound, has a considerable number of views, duration of viewing and average percentage of viewing that mainly reflect the first period, when the educational process was carried out exclusively online.

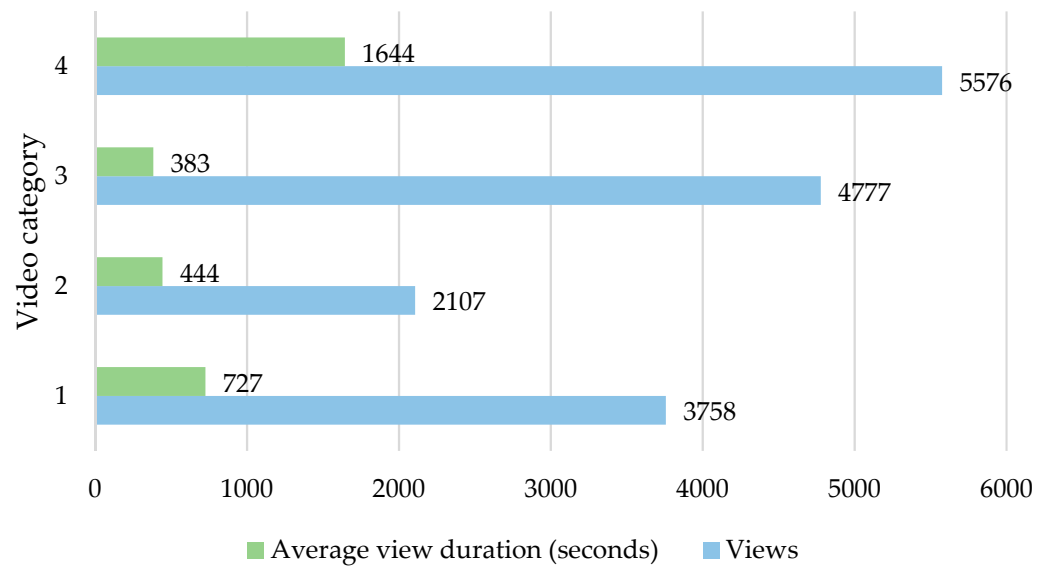


Figure 2. Average viewing time and number of views by video category.

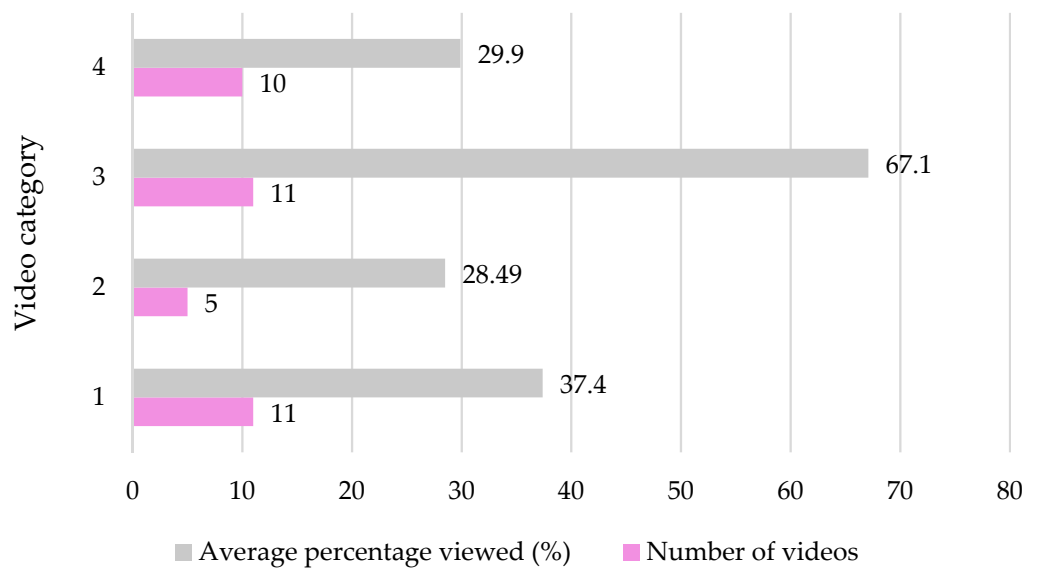


Figure 3. Average percentage of views and number of videos by video category.

4. Main Results

4.1. The Number of Views and Unique Viewers: Two Major Variables in the Foreground

Table 3, whose data is best depicted in Figure 4, summarizes the number of views per week for the two time periods studied (in 2020–2021, when distance learning was exclusive, and in 2021–2022, when hybrid learning was the norm) based on the sequence of course modules. The interest of Table 3 and Figure 4 is also to examine the variations between the two periods studied. It is important to note that the number of views decreased in 2021–2022, which may be due to the fact that students were attending classes in face-to-face mode and asynchronous support was not as necessary as in distance learning environments.

Table 3. Variation of the number of views per week for two time periods per module step.

Year	Number of Weeks												
	Sketching (First Stage)			CAD Drawings (Second Stage)							Final Exam (Third Stage)		
	1	2	3	4	5	6	7	8	9	10	11	12	13
2020–2021	54	137	137	84	154	184	224	186	116	72	110	190	202
2021–2022	27	72	109	134	120	154	177	161	143	96	47	97	143
Variation	27	65	28	−50	34	30	47	25	−27	−24	63	93	59

Separate colors were applied in the two periods according to Figure 4.

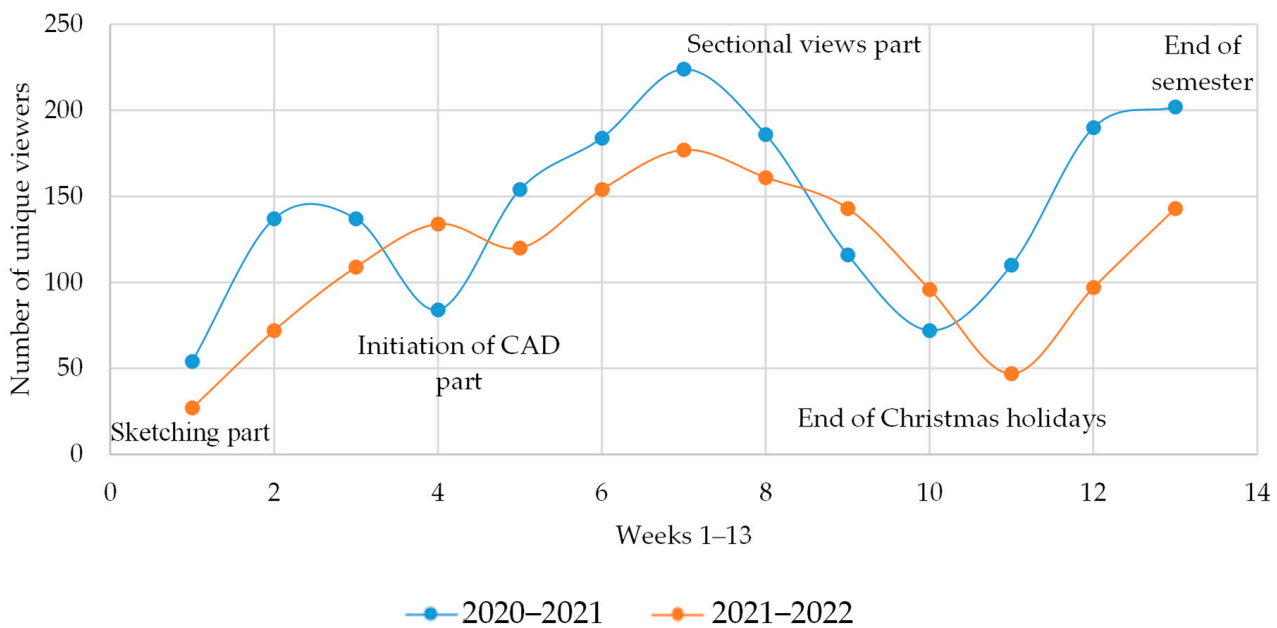


Figure 4. Chart of unique viewers per week. Separate colors were applied in the two periods according to Table 3.

Figure 4 shows two trend curves for the number of unique viewers as a function of CAD module length (in number of weeks) for the two time periods studied (2020–2021 and 2021–2022). Both curves show similar trend dynamics, but only from the fourth course onward. From the beginning of the teaching until the fourth laboratory lecture, the trends are slightly different and can be explained by the fact that the module was delivered exclusively online in 2020–2021 and face-to-face in the first semester of 2021–2022. Higher values of unique viewers in both time periods, can be noticed in week seven. At this point in the educational process, students enroll in the third stage of the CAD module from the second stage of the module workflow. The seventh laboratory lecture is where the most difficult part of the new knowledge has been delivered, referring to the sectional views. It should be noted that at this stage, students who have not yet assimilated the layout of the plans, have difficulties in completing their weekly tasks. Finally, new knowledge was accumulated in the first seven lessons, combining the use of software, the rules of mechanical drawing and the perception of object views, therefore asynchronous task support was necessary in this specific period. From the 8th to the 10th week, the Christmas vacations took place and a clear downward trend can be observed in both lines of the time series. Upward trends are again observed from week 10 onwards, when students resume module attendance after the Christmas break. High values of unique viewers can also be distinguished in week 13, which can be interpreted as the fact that the laboratory lectures have reached their final point and the exams are only one week away.

These first graphically illustrated results must now be rigorously demonstrated by statistical data, which will be analyzed in the following sections.

4.2. Selected Statistical Variables

The statistical study proposed in this section is conducted on the following six variables: “Impressions”, “Unique viewers”, “Viewing time”, “Impressions Click-Through Rate (CTR)”, “Number of views” and “Average viewing time”. In particular, we will use YouTube Analytics terminology to describe these variables. The term “Impressions” is used to express the number of times one of the video thumbnails from the YouTube channel appears on the participant’s screen. Therefore, the “Impressions Click-Through Rate” (CTR) indicates how many impressions were converted into views. Its calculation (expressed as a percentage) is based on the ratio between the number of clicks and the number of impressions. This metric aims to reveal how many people saw the thumbnail, found it interesting and clicked on it. The “Views” metric indicates the number of times a video has been viewed. The term “viewers” refers to the number of individuals who watch a certain piece of content. A viewer may click more than once on a specific piece of content. This is why the “Unique viewers” metric is used, which is a more accurate and relevant variable since it only counts one person, even if they click on the same video content multiple times or use multiple devices or browsers. “Viewing time” is expressed in YouTube terminology as “watch time”. This audience retention metric, expressed here as an average value, counts the hours that a specific video is watched by individual viewers. For each of the six variables defined above, two pairs of datasets were compared:

- Dataset 1: data from the academic period 2020–2021;
- Dataset 2: data from the academic period 2021–2022.

4.3. Normality Test Results

First, for each variable, it was necessary to verify the normality of the data distribution. This normality test is very important insofar as it determines the choice to carry out a test, either parametric or non-parametric. In the remainder of this subsection, we will only discuss the main results of the procedure. The basis and steps of the Shapiro–Wilk test, along with detailed results, are presented in Appendix B.

The results in Appendix B show that, regardless of the two time periods studied (i.e., 2020–2021 and 2021–2022), only the two variables “Impressions” and “Number of views” have p -values greater than the 5% threshold, which does not allow us to reject the null hypothesis and thus to consider that each sample in question is normally distributed. For these variables, an F -test was performed to evaluate whether the variances of the variables (whose distributions can be considered as normal according to the Shapiro–Wilk test) are equal or not. The final objective is to reject or not the null hypothesis of the existence of statistically significant differences between the academic years 2020–2021 and 2021–2022 [36]. These F -tests will be complemented by Student’s t -tests to compare the means of the two data distributions considered.

For the other four variables (i.e., “Unique viewers”; “Watch time”; “Impressions CTR”; and “Average viewing time”) with non-normal distributions, a Mann–Whitney–Wilcoxon test (which is a non-parametric method), was performed to assess whether there is a statistically significant difference between the two periods of 2020–2021 and 2021–2022.

4.4. Hypothesis Testing Results

For each of the two normally distributed variables mentioned above (i.e., “Impressions” and “Number of views”), the Fisher–Snedecor test or F -test was performed. As with the results of the normality tests presented earlier, here we analyze only the results obtained (see Table 4). However, the foundations and main steps of the method are recalled in Appendix C. The results in Table 4 show that for the “Impressions” variable, there is no equality of variances between the 2020–2021 and 2021–2022 data, since the p -value is below the 5% threshold. As for the variable “Number of views”, there is equality of variances

as long as the p -value is greater than the 5% threshold. For this variable in particular, this allows us to conclude that for the two periods considered, since the variances are equal, learners' need for assistance in completing their tasks from viewing asynchronous educational content does not depend on the learning environment.

Table 4. Results of the F -test for the two normally distributed variables "Impressions" and "Number of views".

	2020–2021 "Impressions"	2021–2022 "Impressions"	2020–2021 "Number of Views"	2021–2022 "Number of Views"
p-value	0.01	0.01	0.12	0.12

The F -tests were supplemented with t -tests, as shown in Table 5. The results of these tests reflect the comparison of the means of the data sets of the variables "Number of Views" and "Impressions". As the p -values in Table 5 are above the 5% threshold, the null hypothesis cannot be rejected. Therefore, no statistically significant difference was observed between 2020–2021 and 2021–2022.

Table 5. Results of the t -test for the two normally distributed variables "Impressions" and "Number of views".

	"Impressions"	"Number of Views"
Assumption	Unequal variance	Equal variance
p-value	0.055	0.08

For each of the four remaining variables (i.e., "Unique viewers"; "Viewing time or Watch time"; "Impressions CTR"; and "Average viewing time") whose distributions are not Gaussian for the two periods studied (i.e., 2020–2021 and 2021–2022), a non-parametric Mann–Whitney–Wilcoxon test was performed.

The results in Table 6 show that there is no statistically significant difference between the two periods studied for the variables "Viewing time" and "Unique viewers" because their respective p -values are above the defined risk of 5%. However, this is not the case for the other variables (i.e., "Impressions CTR" and "Average viewing time"). Indeed, their respective p -values are below the 5% risk, confirming the statistically significant difference between the 2020–2021 and 2021–2022 data.

Table 6. Mann–Whitney–Wilcoxon test results for the four variables "Unique viewers"; "Watch time"; "Impressions CTR"; and "Average viewing time".

	Unique Viewers (10 Weeks)	Viewing Time or Watch Time (h)	Impressions CTR (%)	Average Viewing Time (s)
p-value	0.443	0.122	0.000	0.000
Null hypothesis rejected or not rejected	Null hypothesis not rejected	Null hypothesis not rejected	Null hypothesis rejected	Null hypothesis rejected

All the statistical results described above, whether parametric or non-parametric, confirm the preliminary results established in Section 4.1. in that the two variables "Number of views" and "Unique viewers" do not depend on the learning environment (i.e., exclusively remote or hybrid environment).

5. Discussion

Beyond the analysis of the statistical tests proposed above, Figure 4 allows us to draw some major conclusions that we will review and discuss in this section.

The total number of unique viewers in the first semester of the 2020–2021 academic year was 719, and in the same period of the 2021–2022 year was 570. By calculating the percentage of unique viewers for each of the two periods compared to the total number, a percentage difference of 11.55% was calculated. The two curves in Figure 4 representing the data series for both time periods show similar trends in student viewing habits for both time periods, leading us to conclude that learners' need for assistance in completing their tasks is not dependent on the learning environment. This observation is made when comparing online and hybrid modes of instruction. When comparing the visual behavior of students during the first four laboratory lectures taught in the face-to-face mode in 2021–2022 with the same period in 2020–2021 in the online spaces, the curve does not show similar trends. This specific observation can be explained by the fact that when teaching the online module, all twelve lectures were delivered exclusively at a distance, whereas in the hybrid learning spaces, the first four courses were delivered exclusively in face-to-face mode. In addition, the theme of the first four lectures was "sketching", which involves freehand drawings. For the non-computer tasks, videos were only used for object representation, and their contribution was limited to categories 1 and 3, with tasks related to the first category being repeated face-to-face.

The variable "Impressions CTR" was one of the viewing measures that showed statistically significant differences between the two time periods in the non-parametric tests conducted. By comparing the proportion of reduction between the "Number of views" (11.86%) and "Impressions CTR" (1.76%), we can conclude that even though the "Number of views" decreased in the hybrid learning spaces, the "Impressions CTR" variable showed a very low percentage of reduction, which proves the positive attitude of students clicking on the thumbnails of the videos to watch them.

Although the results of this study revealed viewing patterns for both time periods examined, there are still some limitations, based on the circumstances in which the module was delivered during each time period. The learning experience at universities was affected by the consequences of the pandemic, resulting in a significant decline in student retention and academic performance. Sustainable measures, such as those implemented in this study, had to be taken to first ensure students' virtual presence, as well as support for asynchronous tasks.

In the exclusive online instruction, the lack of physical contact between learners and their educators allowed the former to engage in the asynchronous support channel, which allowed instructors to track their viewing activities and analyze their viewing behaviors through data analysis by retrieving high-precision information. In the hybrid learning modes, specifically in the face-to-face delivered modules, learners had the opportunity to physically communicate with their instructors and get support for their tasks.

In synthesis of the above, the YouTube channel created and used in this work as an asynchronous tutoring tool has been integrated into the educational process. It provides quality asynchronous support when needed and is part of a long-term viability and sustainability approach. These new tools for supporting individual tasks outside of the classroom can benefit pedagogical practices and ultimately the learning process in very implicit ways and primarily by being "masked" by popular social media sites like YouTube. The channel analysis provided valuable information about learners' visualization habits that can serve as guidelines for future instructors and developers of instructional module structure. The log of visualization measures revealed viewing patterns indicating that students' visualization behaviors follow the flow of the module, and especially regardless of their mode of attendance and teaching space. The ability to follow the content of a module through educational videos at one's own pace and preference contributes to the development of senses of quality and equity [19]. The EDM and its statistical analysis showed that the foundation of the YouTube channel met the needs of students regardless of the learning environment on which this tool was applied.

Moving forward, the methodology applied in this study provides direct feedback for future CAD instructors and instructional module developers. Our proposed recommendations and action plan are as follows:

- Recognize module gaps throughout the course. Recognize the needs of learners, for the promotion of sustainable engineering education. Reorganize the flow of tasks according to the learning objectives of the modules and determine the critical points of difficulty according to the knowledge to be acquired.
- Do not focus on a specific task, but on the unit by creating categories of units. Each unit may include several tasks, but the focus of the knowledge introduced is not the technical instruction itself, but the concept of the learning methodology. At the stage where the new knowledge is accumulated, call it the “pick of the curve”. Determine the points at which the new knowledge needs to be performed, as well as its nature in terms of asynchronous support. Do not be afraid to combine new knowledge with entertainment by using user-friendly digital tools, such as YouTube channels.
- Create direct access to certain parts of the course to redirect learners if necessary. Make access clear to avoid confusion.
- Analyze, evaluate and reconstruct the course based on the results. This means questioning the actions taken, modifying them if they fail, and adapting them according to the nature of the learning environments.
- Never neglect the social aspects: learners must be prepared for their professional future. Technical aspects can be learned through training, but methodology is the expertise of higher education instructors.

6. Conclusions and Future Work

In 2020–2021, during the most critical hours of the COVID-19 pandemic, higher education instructors, specifically those at the University of West Attica (Greece), created a social media channel (in this study, using YouTube), as part of a mechanical engineering CAD module, to provide students with asynchronous support for their teaching tasks. To provide learners with the most direct access, links to the 37 videos were attached to each assessed task. One year later, at the beginning of the meta-COVID-19 period, the same asynchronous task support technique was applied, but this time in face-to-face and blended learning spaces.

The experimental analysis proposed in this manuscript is based on the process of processing and interpreting acquired knowledge data extracted directly from the reports of a YouTube channel; this YouTube channel having been created and used by an instructor and containing educational videos of different categories based on the learning objectives of a CAD module. The main challenge here was to investigate the potential of the data as acquired from the YouTube channel and whether the raw material downloaded in the form of previews could be processed and reveal valuable information about how students use this form of asynchronous digital educational material; YouTube having been hijacked from its primary function, i.e., entertainment.

The shift from exclusively online to hybrid learning environments first showed that the use of asynchronous task support decreased. YouTube analytics were the most appropriate tool in terms of efficiency and accuracy for expressing student retention beyond physical, online, or hybrid engineering labs, as they not only recorded the number of times a video was viewed, but also differentiated between users who watched the same video multiple times. Specifically, we defined and used variables and measures that are very common on social media sites, expressing the level of audience acceptance, to examine the contribution of video to the learning process. The following six statistical variables were selected as primary measures expressing audience retention in YouTube channels: “Impressions”, “Unique viewers”, “Viewing time or watch time”, “Impressions Click-Through Rate (CTR)”, “Number of views” and “Average viewing time”.

The comparative analysis of the YouTube reports showed similar trends in student viewing habits over the two time periods studied (i.e., at the most critical time of the global

health crisis and at the beginning of the meta-COVID-19 period), with a slight decrease of nearly 12% in viewing due to a return to traditional learning environments, where most students solved their tasks in class and did not require asynchronous support. In contrast to face-to-face and 100% distance learning, the analysis showed that trends in learner viewing habits are similar during online and hybrid learning spaces. In particular, reports from the social media channel showed that educational videos followed the weekly stream trend, resulting in an increase in active viewers, which was directly related to the increase in workload and workload accumulation.

After testing the normality of each of the above six variables, a series of hypothesis tests (i.e., parametric and non-parametric based on normality tests) were performed to accept or reject the null hypothesis of this study, which concerns the absence of statistically significant changes in learners' listening habits over the two periods studied (i.e., 2020–2021 and 2021–2022). Of the six variables analyzed, only two—"Impressions CTR" and "Average viewing time"—show display metrics with statistically significant differences between the two periods studied. For both of these statistically significant variables, it is appropriate to focus on impressions CTR, which expresses the number of times viewers click to watch a video after seeing its thumbnail. Although there was a decrease in views and impressions CTR between the two time periods, the percentage reduction was not proportional for either variable.

Although we have answered the five research questions listed in Section 2, the current study has some limitations. Although many universities have begun to use data and analytics, there is still a long way to go before these tools can fully prove their potential in terms of improving the learning experience. This is especially true today, due to the unstable health conditions caused by the COVID-19 outbreak, although overall they seem to be gradually normalizing.

Future work will involve processing analyses of a second semester CAD module Computer Aided Mechanical Design (CAD II) and performing similar statistical tests to improve the reliability of the results. Due to the increase in the number of students and institutions participating in online learning and using digital tools over the past two years, there is now a plethora of data available that may not have been available before. Institutions of higher education may want to start using this data with an eye toward serving students ever better in the years to come.

Author Contributions: Conceptualization, Z.K. and C.S. (Constantinos Stergiou); methodology, Z.K., C.S. (Constantinos Stergiou), G.B. and S.J.; software, G.B.; validation, G.B.; formal analysis, Z.K. and G.B.; investigation, Z.K. and C.T.; resources, Z.K. and C.S. (Constantinos Stergiou); data curation: Z.K.; writing—original draft preparation: Z.K.; writing—review and editing: Z.K., C.S. (Constantinos Stergiou), G.B., S.J. and A.O.; visualization: Z.K., C.S. (Constantinos Stergiou), G.B. and S.J.; supervision, C.S. (Constantinos Stergiou) and C.S. (Cleo Sgouropoulou); project administration, C.S. (Constantinos Stergiou) and S.J.; funding acquisition, S.J. and A.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This study involves the analysis of data sets obtained from reports on a YouTube channel created and administered by a member of the authors' team. Therefore, all rights to the reports are reserved to the authors.

Informed Consent Statement: Informed consent was obtained from all study participants at the time of initial data collection.

Data Availability Statement: Not applicable.

Acknowledgments: These research activities are currently supported by the University of West Attica and more particularly by its Department of Mechanical Engineering, as well as by the University of Tours. The authors of this manuscript would like to thank their colleagues at the following institutions, as well as the students, who contributed greatly to the success of this work.

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

Abbreviations

The following abbreviations are used in this paper:

CAD	Computer-aided design
CTR	Impressions Click-Through Rate
EDM	Educational data mining
LA	Learning analytics
LMS	Learning management system
MS	Microsoft
TAM	Technology acceptance model
TLP	Teaching—learning process
VLE	Virtual learning environments

Appendix A. URL of Each Educational Video Content of the YouTube Channel “MCAD I UNIWA”

Video	Video Title	URL
Total		
1	CAD 06A	https://www.youtube.com/watch?v=Pp2KWt28haM (accessed on 23 May 2022).
2	CAD 07 PART 13B	https://www.youtube.com/watch?v=N01Bul3DCAU&t=426s (accessed on 23 May 2022).
3	CAD 07 PART 13A	https://www.youtube.com/watch?v=ezeYLUihbn4 (accessed on 23 May 2022).
4	CAD 06B ASK 11B	https://www.youtube.com/watch?v=-VrzgqM-x5Q (accessed on 23 May 2022).
5	CAD 08 ASK 15	https://www.youtube.com/watch?v=HT-oJB2JZ9k (accessed on 23 May 2022).
6	CAD 06B ASK 11A	https://www.youtube.com/watch?v=il4Nyba1dn8 (accessed on 23 May 2022).
7	CAD 05D	https://www.youtube.com/watch?v=rARN13OfGig (accessed on 23 May 2022).
8	CAD 10 PART 20	https://www.youtube.com/watch?v=v40Dg4ftOXw (accessed on 23 May 2022).
9	CAD 07 DIMENSION	https://www.youtube.com/watch?v=xBQ_qv-szIQ (accessed on 23 May 2022).
10	CAD 09 PART 17	https://www.youtube.com/watch?v=MxraNn92Wfo (accessed on 23 May 2022).
11	CAD 07 PART 13	https://www.youtube.com/watch?v=AleUy0Q_1v8 (accessed on 23 May 2022).
12	ASK 03A	https://www.youtube.com/watch?v=A7UeOz1wluE (accessed on 23 May 2022).
13	CAD 08 PART 14	https://www.youtube.com/watch?v=g2eXTIYJm0Y (accessed on 23 May 2022).
14	CAD 05A	https://www.youtube.com/watch?v=HtUdjt8w89Y&t=8s (accessed on 23 May 2022).
15	CAD 07 PART 12	https://www.youtube.com/watch?v=qSENFhmbZfw (accessed on 23 May 2022).
16	CAD 11	https://www.youtube.com/watch?v=Kp-vX2eBpLQ (accessed on 23 May 2022).
17	ASK 03B	https://www.youtube.com/watch?v=c7CaOIUXO0E (accessed on 23 May 2022).
18	ASK 06C	https://www.youtube.com/watch?v=l8LVvr5uIz0&t=17s (accessed on 23 May 2022).
19	ASK 06A	https://www.youtube.com/watch?v=3BP7jxdpt7c (accessed on 23 May 2022).
20	CAD 04B	https://www.youtube.com/watch?v=Dlc8CvUG5wk (accessed on 23 May 2022).
21	CAD 09 PART 18	https://www.youtube.com/watch?v=vopGLakGO_g (accessed on 23 May 2022).
22	CAD 08 LIBRARY	https://www.youtube.com/watch?v=RWY5EuLB9nQ (accessed on 23 May 2022).
23	ASK 03A TOP	https://www.youtube.com/watch?v=knAEEWcHqLA (accessed on 23 May 2022).
24	CAD 04A	https://www.youtube.com/watch?v=9y2bkS7T3Wc (accessed on 23 May 2022).
25	ASK 06B	https://www.youtube.com/watch?v=FjmdZa2ix2o&t=11s (accessed on 23 May 2022).
26	ASK 04A	https://www.youtube.com/watch?v=W5FzH5-RF3w (accessed on 23 May 2022).
27	ASK 04B	https://www.youtube.com/watch?v=4ft1ZpLzj8 (accessed on 23 May 2022).
28	ASK 03A FRONT LEFT	https://www.youtube.com/watch?v=cRUM-BOB8jM&t=105s (accessed on 23 May 2022).
29	CAD 05B	https://www.youtube.com/watch?v=0m-9ofdEgJs (accessed on 23 May 2022).
30	CAD 05C	https://www.youtube.com/watch?v=gWf-1wpQnAA (accessed on 23 May 2022).
31	CAD 11 PART 22	https://www.youtube.com/watch?v=TPNZibk1CJs (accessed on 23 May 2022).
32	CAD 07 PART 006	https://www.youtube.com/watch?v=p5YSKekgDmw (accessed on 23 May 2022).
33	CAD 07 PART 007	https://www.youtube.com/watch?v=5hUMxgh3AzE (accessed on 23 May 2022).
34	ASK 03A RIGHT	https://www.youtube.com/watch?v=gDi0qexf1KI (accessed on 23 May 2022).
35	CAD 10 TEXT	https://www.youtube.com/watch?v=tG39K1aFc8I (accessed on 23 May 2022).
36	01 SKETCH	https://www.youtube.com/watch?v=v2BTZmS6YIk (accessed on 23 May 2022).
37	AUTOCAD DESIGN CENTER	https://www.youtube.com/watch?v=3g3CqBEB7MQ (accessed on 23 May 2022).

Appendix B. Summary of the Basics and Main Steps of the Shapiro–Wilk Normality Test

The Shapiro–Wilk normality test is designed to detect all deviations from normality. In particular, this test rejects the normality hypothesis when the *p*-value is less than or equal to a threshold value (usually 5%). The null hypothesis is that there is no difference between the distribution studied and a normal distribution. The alternative hypothesis is that there is a difference. If the *p*-value at the end of the test is less than the set threshold (usually 5%), then the null hypothesis can be rejected and the data are not considered normal. In this case, a series of non-parametric tests can be applied. Conversely, if the null hypothesis cannot be rejected, the data are considered normal and parametric tests can be implemented [46,47]. Note that if only one of the data sets does not meet the set threshold and the other data set does, the variable of interest is considered a non-parametrically valued variable.

In addition to the *p*-value and to decide the normality of each distribution, we will focus on two metrics in particular: skewness and kurtosis of each of the variables examined for the two datasets (i.e., variables referring to a sample of 37 YouTube videos, with Dataset 1 containing observations from the 2020–2021 academic year and Dataset 2 from the 2021–2022 academic year). While skewness focuses on the overall shape, kurtosis focuses on the tail shape. The normal distribution is characterized by a zero skewness coefficient and a zero kurtosis coefficient. Concerning the skewness, a positive coefficient indicates a left asymmetry, while a negative coefficient indicates a right asymmetry. With respect to kurtosis, a negative value indicates that the distribution is “platykurtic”, i.e., more flattened than a normal density. A positive kurtosis coefficient indicates that the distribution is “leptokurtic”, i.e., less flattened.

The following tables summarize the *p*-values, skewness and kurtosis of the six variables defined in this study for the two periods studied (i.e., 2020–2021 and 2021–2022).

	2020–2021 “Impressions”	2021–2022 “Impressions”	2020–2021 “Unique Viewers” (10 Weeks)	2021–2022 “Unique Viewers” (10 Weeks)
<i>p</i> -value (Shapiro–Wilk test)	0.675	0.120	0.561	0.038
Skewness	−0.102	0.150	0.138	−0.029
Kurtosis	−0.688	−0.917	−0.884	−0.946
Type of test allowed	Parametric		Non-parametric	

	2020–2021 “Watch Time (h)”	2021–2022 “Watch Time (h)”	2020–2021 “Impressions CTR (%)”	2021–2022 “Impressions CTR (%)”
<i>p</i> -value (Shapiro–Wilk test)	0.000	0.000	0.310	0.000
Skewness	1.280	1.717	0.550	1.772
Kurtosis	0.289	3.320	0.358	3.591
Type of test allowed	Non-parametric		Non-parametric	

	2020–2021 “Number of Views”	2021–2022 “Number of Views”	2020–2021 “Average Viewing Time (s)”	2021–2022 “Average Viewing Time (s)”
<i>p</i> -value (Shapiro–Wilk test)	0.479	0.086	0.000	0.000
Skewness	0.177	0.447	−0.323	1.086
Kurtosis	−0.730	0.844	0.924	1.369
Type of test allowed	Parametric		Non-parametric	

Appendix C. Summary of the Basics and Main Steps of and *t*-Tests for Parametric variables, and the Mann–Whitney–Wilcoxon Test for Non-Parametric Variables

As stated in [48,49], the hypothesis is an interpretation of the reasons for a certain phenomenon. Two types of hypotheses can be defined in a scientific approach:

- The first is the research hypothesis, which aims to state the subject of the research. If it is well defined, it will include the factors being studied and their expected relationship.
- The second is the statistical hypothesis, which converts the research hypothesis into a mathematical complex and a statistically testable statement about the presumed value of the variable being studied in the population.

Therefore, the null hypothesis must be tested. In our case, it concerns the lack of significant difference between the variances of the six variables examined, when comparing the samples of variables from the two academic semesters mentioned above [50].

Parametric statistical tests (based on the hypothesis that the sample under consideration is drawn from a population following a distribution belonging to a given family, i.e., the normal distribution), when their use is well justified, they generally have greater statistical power than non-parametric tests (i.e., without a distribution). More precisely, they are likely to detect a significant effect when it actually exists. Normality is tested here by the Shapiro–Wilk test mainly because, for a given significance level, the probability of rejecting the null hypothesis (i.e., a sample is drawn from a normally distributed population) if it is false is higher than for other tests of normality [42].

Appendix C.1. Fisher–Snedecor or F-Test for Parametric Variables

For the two variables examined (i.e., “Impressions” and “Number of views”), a sample size of 37 was set, referring to the number of videos uploaded to the YouTube channel. The degrees of freedom are determined by subtracting one from the sample size. In statistical calculations, degrees of freedom measure the mathematical complexity of a calculated parameter. As mentioned earlier, the probability that the tested parameters have statistical significance is tested by setting a threshold for their statistical significance.

The Fisher–Snedecor test or *F*-test consists in comparing the resultant value of the test with the critical value (F_{critical}) of the Fisher–Snedecor distribution for the risk sought (the risk is equal to 5% here); this critical value is determined from a table. If the resulting value of the test (*p*-value) is higher than the critical value (5%), then the null hypothesis (i.e., the two variances are equal) is rejected. Otherwise, it is not rejected [51]. Note that the more dissimilar the variances are, the more the *p*-value tends to zero.

*Appendix C.2. Student’s *t*-Test for Parametric Variables*

The Student’s *t*-test, or *t*-test, is a popular statistical test used to measure the differences between the means of two groups or a group compared to a standard value. It is based on a probability distribution called Student’s *T* distribution. Performing this test is used to understand whether the differences are statistically significant.

The 2-sample *t*-test, which is the most standard and classical analysis technique, aims at comparing the means of two independent populations to identify a significant difference. To run a Student’s *T* distribution, the following must be available: the difference between the mean values of the data sets; the variance of each sample; the number of data in each group; and the acceptable error threshold (usually 5%). In this study, for a variable under consideration, the null hypothesis is that the two populations (of 2020–2021 and 2021–2022) are identical and that there is no significant difference between them. At the end of the test, if the *p*-value is lower than the set threshold (usually 5%), the null hypothesis can be rejected. Otherwise, it cannot be rejected.

Appendix C.3. Mann–Whitney–Wilcoxon Test for Non-Parametric Variables

The Mann–Whitney–Wilcoxon test is used to test the hypothesis that the distributions of each of the two groups of data are close. Like any statistical test, it consists in highlighting an event whose probability distribution is known (at least its asymptotic form) from what is observed. The *p*-value obtained, if it is unlikely according to this law, will suggest rejecting the null hypothesis. More precisely, if the *p*-value is greater than the fixed risk (here 5%), then the null hypothesis cannot be rejected. Otherwise, it can be rejected.

References

1. Chango, W.; Lara, J.A.; Cerezo, R.; Romero, C. A review on data fusion in multimodal learning analytics and educational data mining. *WIREs Data Min. Knowl. Discov.* **2022**, *e1458*, 1–19. [[CrossRef](#)]
2. Wang, Q.; Mousavi, A.; Lu, C. A scoping review of empirical studies on theory-driven learning analytics. *Distance Educ.* **2022**, *43*, 6–29. [[CrossRef](#)]
3. Hantoobi, S.; Wahdan, A.; Al-Emran, M.; Shaalan, K. A Review of Learning Analytics Studies. In *Recent Advances in Technology Acceptance Models and Theories*; Al-Emran, M., Shaalan, K., Eds.; Studies in Systems, Decision and Control; Springer International Publishing: Cham, Switzerland, 2021; pp. 119–134, ISBN 978-3-030-64987-6.
4. Gutierrez-Bucheli, L.; Kidman, G.; Reid, A. Sustainability in engineering education: A review of learning outcomes. *J. Clean. Prod.* **2022**, *330*, 129734. [[CrossRef](#)]
5. Hodges, C.; McCullough, H. The Adjacent Possible for Higher Education: The Digital Transformation of Faculty. Available online: <https://er.educause.edu/articles/2021/9/the-adjacent-possible-for-higher-education-the-digital-transformation-of-faculty> (accessed on 21 November 2021).
6. Jacques, S.; Ouahabi, A.; Lequeu, T. Remote Knowledge Acquisition and Assessment during the COVID-19 Pandemic. *Int. J. Eng. Pedagog. IJEP* **2020**, *10*, 120–138. [[CrossRef](#)]
7. Jacques, S.; Ouahabi, A.; Lequeu, T. Synchronous E-learning in Higher Education during the COVID-19 Pandemic. In Proceedings of the 2021 IEEE Global Engineering Education Conference (EDUCON), Vienna, Austria, 21–23 April 2021; pp. 1102–1109.
8. Shloul, T.; Javeed, M.; Gochoo, M.; Alsuhibany, S.; Ghadi, Y.; Jalal, A.; Park, J. Student's Health Exercise Recognition Tool for E-Learning Education. *Intell. Autom. Soft Comput.* **2022**, *35*, 149–161. [[CrossRef](#)]
9. Lutfi, A.; Alsyouf, A.; Almaiah, M.A.; Alrawad, M.; Abdo, A.A.K.; Al-Khasawneh, A.L.; Ibrahim, N.; Saad, M. Factors Influencing the Adoption of Big Data Analytics in the Digital Transformation Era: Case Study of Jordanian SMEs. *Sustainability* **2022**, *14*, 1802. [[CrossRef](#)]
10. Kanetaki, Z.; Stergiou, C.; Bekas, G.; Jacques, S.; Troussas, C.; Sgouropoulou, C.; Ouahabi, A. Grade Prediction Modeling in Hybrid Learning Environments for Sustainable Engineering Education. *Sustainability* **2022**, *14*, 5205. [[CrossRef](#)]
11. Orji, F.A.; Vassileva, J. A Comparative Evaluation of the Effect of Social Comparison, Competition, and Social Learning in Persuasive Technology on Learning. In Proceedings of the Intelligent Tutoring Systems, Virtual Event, 7–11 June 2021; Cristea, A.I., Troussas, C., Eds.; Springer International Publishing: Cham, Switzerland, 2021; pp. 369–375.
12. Jackman, W.M. YouTube Usage in the University Classroom: An Argument for its Pedagogical Benefits. *Int. J. Emerg. Technol. Learn. IJET* **2019**, *14*, 157–166. [[CrossRef](#)]
13. Cabrera, I.; Villalon, J.; Chavez, J. Blending Communities and Team-Based Learning in a Programming Course. *IEEE Trans. Educ.* **2017**, *60*, 288–295. [[CrossRef](#)]
14. Kanetaki, Z.; Stergiou, C.; Bekas, G.; Troussas, C.; Sgouropoulou, C. The impact of different learning approaches based on MS Teams and Moodle on students' performance in an on-line mechanical CAD module. *Glob. J. Eng. Educ.* **2021**, *23*, 185–190. [[CrossRef](#)]
15. Kanetaki, Z.; Stergiou, C.; Bekas, G.; Troussas, C.; Sgouropoulou, C. Creating a Metamodel for Predicting Learners' Satisfaction by Utilizing an Educational Information System During COVID-19 Pandemic. *Nov. Intell. Digit. Syst.* **2021**, *338*, 127–136. [[CrossRef](#)]
16. Baker, R.S.J.D. International Encyclopedia of Education. In *Data Mining*, 3rd ed.; McGaw, B., Peterson, P., Baker, E., Eds.; 2010; pp. 112–118.
17. Govindarajan, M.; Govindarajan, M. Educational Data Mining Techniques and Applications. Available online: <https://www.igi-global.com/gateway/chapter/www.igi-global.com/gateway/chapter/272957> (accessed on 8 May 2022).
18. Ihantola, P.; Vihavainen, A.; Ahadi, A.; Butler, M.; Börstler, J.; Edwards, S.H.; Isohanni, E.; Korhonen, A.; Petersen, A.; Rivers, K.; et al. Educational Data Mining and Learning Analytics in Programming: Literature Review and Case Studies. In Proceedings of the 2015 ITiCSE on Working Group Reports, Vilnius, Lithuania, 4–8 July 2015; Association for Computing Machinery: New York, NY, USA, 2015; pp. 41–63.
19. Krishnan, R.; Nair, S.; Saamuel, B.S.; Justin, S.; Iwendi, C.; Biamba, C.; Ibeke, E. Smart Analysis of Learners Performance Using Learning Analytics for Improving Academic Progression: A Case Study Model. *Sustainability* **2022**, *14*, 3378. [[CrossRef](#)]
20. Siemens, G.; Baker, R.S.J.d. Learning analytics and educational data mining: Towards communication and collaboration. In Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Vancouver, BC, Canada, 29 April 2012–2 May 2012; Association for Computing Machinery: New York, NY, USA, 2012; pp. 252–254.
21. Muttakin, F.; Wang, J.-T.; Mulyanto, M.; Leu, J.-S. Evaluation of Feature Selection Methods on Psychosocial Education Data Using Additive Ratio Assessment. *Electronics* **2022**, *11*, 114. [[CrossRef](#)]
22. Nowak-Brzezińska, A.; Łazarz, W. Qualitative Data Clustering to Detect Outliers. *Entropy* **2021**, *23*, 869. [[CrossRef](#)] [[PubMed](#)]
23. Dooley, L.; Makasis, N. Understanding Student Behavior in a Flipped Classroom: Interpreting Learning Analytics Data in the Veterinary Pre-Clinical Sciences. *Educ. Sci.* **2020**, *10*, 260. [[CrossRef](#)]
24. Al-Marouf, R.; Ayoubi, K.; Alhumaid, K.; Aburayya, A.; Alshurideh, M.; Alfaisal, R.; Salloum, S. The acceptance of social media video for knowledge acquisition, sharing and application: A comparative study among YouTube users and TikTok Users' for medical purposes. *Int. J. Data Netw. Sci.* **2021**, *5*, 197–214. [[CrossRef](#)]
25. de Oliveira, C.F.; Sobral, S.R.; Ferreira, M.J.; Moreira, F. How Does Learning Analytics Contribute to Prevent Students' Dropout in Higher Education: A Systematic Literature Review. *Big Data Cogn. Comput.* **2021**, *5*, 64. [[CrossRef](#)]

26. Draus, P.J.; Curran, M.J.; Trempus, M.S. The Influence of Instructor-Generated Video Content on Student Satisfaction with and Engagement in Asynchronous Online Classes. *J. Online Learn. Teach.* **2014**, *10*, 240–254.
27. Rodrigo, M.M.T.; Ladrido, E.M.M. Promoting Equity and Assuring Teaching and Learning Quality: Magisterial Lectures in a Philippine University during the COVID-19 Pandemic. *Educ. Sci.* **2022**, *12*, 146. [[CrossRef](#)]
28. Isaenko, E.; Makrinova, E.; Rozdolskaya, I.; Matuzenko, E.; Bozhuk, S. Research of social media channels as a digital analytical and planning technology of advertising campaigns. In Proceedings of the IOP Conference Series: Materials Science and Engineering, St. Petersburg, Russian, 27–29 August 2020; Volume 986, p. 012014. [[CrossRef](#)]
29. Al-rahmi, W.M.; Othman, M.S.; Yusuf, L.M. Using Social Media for Research: The Role of Interactivity, Collaborative Learning, and Engagement on the Performance of Students in Malaysian Post-Secondary Institutes. *Mediterr. J. Soc. Sci.* **2015**, *6*, 536. [[CrossRef](#)]
30. Dahlstrom, E. *ECAR Study of Undergraduate Students and Information Technology, 2012*; EDUCAUSE Center for Applied Research: Louisville, CO, USA, 2012.
31. Kanaki, K.; Kalogiannakis, M.; Poulakis, E.; Politis, P. Employing Mobile Technologies to Investigate the Association Between Abstraction Skills and Performance in Environmental Studies in Early Primary School. *Int. J. Interact. Mob. Technol. IJIM* **2022**, *16*, 241–249. [[CrossRef](#)]
32. Walsh, J.N.; O'Brien, M.P.; Slattery, D.M. Video Viewing Patterns Using Different Teaching Treatments: A Case Study Using YouTube Analytics. *Res. Educ. Learn. Innov. Arch.* **2019**, *22*, 77–95. [[CrossRef](#)]
33. McGowan, A.; Hanna, P.; Anderson, N. Teaching Programming: Understanding Lecture Capture YouTube Analytics. In Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education, Arequipa, Peru, 9–13 July 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 35–40.
34. Schwarzenberg, P.; Navon, J.; Nussbaum, M.; Pérez-Sanagustín, M.; Caballero, D. Learning experience assessment of flipped courses. *J. Comput. High. Educ.* **2018**, *30*, 237–258. [[CrossRef](#)]
35. Le, A.; Joordens, S.; Chrysostomou, S.; Grinnell, R. Online lecture accessibility and its influence on performance in skills-based courses. *Comput. Educ.* **2010**, *55*, 313–319. [[CrossRef](#)]
36. Kanetaki, Z.; Stergiou, C.; Bekas, G.; Troussas, C.; Sgouropoulou, C. Analysis of Engineering Student Data in Online Higher Education During the COVID-19 Pandemic. *Int. J. Eng. Pedagogy IJEP* **2021**, *11*, 27–49. [[CrossRef](#)]
37. Binkley, M.; Erstad, O.; Herman, J.; Raizen, S.; Ripley, M.; Miller-Ricci, M.; Rumble, M. Defining Twenty-First Century Skills. In *Assessment and Teaching of 21st Century Skills*; Griffin, P., McGaw, B., Care, E., Eds.; Springer Netherlands: Dordrecht, The Netherlands, 2012; pp. 17–66, ISBN 978-94-007-2324-5.
38. Konys, A. An Ontology-Based Knowledge Modelling for Sustainable Entrepreneurship Domain. In Proceedings of the 55th Hawaii International Conference on System Sciences, Hawaii, HI, USA, 4–7 January 2022; pp. 5316–5325.
39. Petrilă, L.; Goudenhooff, G.; Gyarmati, B.F.; Popescu, F.-A.; Simuț, C.; Brihan, A.-C. Effective Teaching during the COVID-19 Pandemic? Distance Learning and Sustainable Communication in Romania. *Sustainability* **2022**, *14*, 7269. [[CrossRef](#)]
40. Zabidi, N.; Wang, W. The Use of Social Media Platforms as a Collaborative Supporting Tool: A Preliminary Assessment. *Int. J. Interact. Mob. Technol. IJIM* **2021**, *15*, 138–148. [[CrossRef](#)]
41. Kanetaki, Z.; Stergiou, C.; Bekas, G.; Troussas, C.; Sgouropoulou, C. Evaluating Remote Task Assignment of an Online Engineering Module through Data Mining in a Virtual Communication Platform Environment. *Electronics* **2022**, *11*, 158. [[CrossRef](#)]
42. Rus-Arias, E.; Palos-Sanchez, P.R.; Reyes-Menendez, A. The Influence of Sociological Variables on Users' Feelings about Programmatic Advertising and the Use of Ad-Blockers. *Informatics* **2021**, *8*, 5. [[CrossRef](#)]
43. Batool, S.; Rashid, J.; Nisar, M.W.; Kim, J.; Mahmood, T.; Hussain, A. A Random Forest Students' Performance Prediction (RFSP) Model Based on Students' Demographic Features. In Proceedings of the 2021 Mohammad Ali Jinnah University International Conference on Computing (MAJICC), Karachi, Pakistan, 15–17 July 2021; pp. 1–4.
44. Anantharamaiah, K.B. YouTube Analytics Using Google Data Studio 2020. Available online: <https://doi.org/10.2139/ssrn.3655551> (accessed on 23 May 2022).
45. Chatzopoulos, A.; Kalogiannakis, M.; Papadakis, S.; Papoutsidakis, M. A Novel, Modular Robot for Educational Robotics Developed Using Action Research Evaluated on Technology Acceptance Model. *Educ. Sci.* **2022**, *12*, 274. [[CrossRef](#)]
46. Camuffo, A.; Gerli, F. Modeling management behaviors in lean production environments. *Int. J. Oper. Prod. Manag.* **2018**, *38*, 403–423. [[CrossRef](#)]
47. Zhang, L.; Lu, W.; Liu, X.; Pedrycz, W.; Zhong, C. Fuzzy C-Means clustering of incomplete data based on probabilistic information granules of missing values. *Knowl.-Based Syst.* **2016**, *99*, 51–70. [[CrossRef](#)]
48. Sakai, T. *t*-Tests. In *Laboratory Experiments in Information Retrieval: Sample Sizes, Effect Sizes, and Statistical Power*; Sakai, T., Ed.; The Information Retrieval Series; Springer: Gateway East, Singapore, 2018; pp. 27–41, ISBN 9789811311994.
49. Booth, T.; Doumas, A.; Murray, A.L. Null Hypothesis. In *Encyclopedia of Personality and Individual Differences*; Zeigler-Hill, V., Shackelford, T.K., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 3267–3270, ISBN 978-3-319-24612-3.
50. Saxon, E. Defining the null hypothesis. *BMC Biol.* **2015**, *13*, 68. [[CrossRef](#)] [[PubMed](#)]
51. F Distribution—An Overview | ScienceDirect Topics. Available online: <https://www.sciencedirect.com/topics/mathematics/f-distribution> (accessed on 23 May 2022).