

Article

Educational Computational Chemistry for In-Service Chemistry Teachers: A Data Mining Approach to E-Learning Environment Redesign

José Hernández-Ramos ^{1,2} , Lizethly Cáceres-Jensen ^{1,3}  and Jorge Rodríguez-Becerra ^{4,*} 

¹ Physical & Analytical Chemistry Laboratory (PachemLab), Department of Chemistry, Faculty of Basic Science, Universidad Metropolitana de Ciencias de la Educación, Santiago 7760197, Chile; jose.hernandez@umce.cl (J.H.-R.); lyzethly.caceres@umce.cl (L.C.-J.)

² Doctorate in Education Program, Academic Vice-Rector, Universidad Metropolitana de Ciencias de la Educación, Santiago 7760197, Chile

³ Nucleus Computational Thinking and Education for Sustainable Development (NuCES), Center for Research in Education (CIE-UMCE), Universidad Metropolitana de Ciencias de la Educación, Santiago 7760197, Chile

⁴ Escuela de Postgrado, Universidad Tecnológica Metropolitana, Santiago 8940000, Chile

* Correspondence: jorge.rodriguez@utem.cl

Abstract: The use of technology in education has experienced significant growth in recent years. In this regard, computational chemistry is considered a dynamic element due to the constant advances in computational methods in chemistry, making it an emerging technology with high potential for application in teaching chemistry. This article investigates the characteristics and perceptions of in-service chemistry teachers who participated in an e-learning educational computational chemistry course. Additionally, it examines how educational data mining techniques can contribute to optimising and developing e-learning environments. The results indicate that teachers view incorporating computational chemistry elements in their classes positively but that this is not profoundly reflected in their teaching activity planning. On the other hand, generated statistical models demonstrate that the most relevant variables to consider in the instructional design of an e-learning educational computational chemistry course are related to participation in various course instances and partial evaluations. In this sense, the need to provide additional support to students during online learning is highlighted, especially during critical moments such as evaluations. In conclusion, this study offers valuable information on the characteristics and perceptions of in-service chemistry teachers and demonstrates that educational data mining techniques can help improve e-learning environments.

Keywords: educational computational chemistry; in-service chemistry teachers; data mining; technological pedagogical science knowledge (TPASK); e-learning



Citation: Hernández-Ramos, J.; Cáceres-Jensen, L.; Rodríguez-Becerra, J. Educational Computational Chemistry for In-Service Chemistry Teachers: A Data Mining Approach to E-Learning Environment Redesign. *Educ. Sci.* **2023**, *13*, 796. <https://doi.org/10.3390/educsci13080796>

Academic Editors: Shirley Avargil and Jenna Koenen

Received: 23 May 2023

Revised: 12 July 2023

Accepted: 20 July 2023

Published: 3 August 2023



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

1. Introduction

Continuing education and professional development for chemistry teachers are essential for improving the quality of teaching and preparing students to face the challenges of the current and future worlds [1]. This training should include updating pedagogical, scientific, and technological knowledge. In other words, teachers need to develop the necessary expertise to use emerging scientific technology to implement learning environments that promote scientific learning in students [2].

1.1. Educational Computational Chemistry and E-Learning Environments

Computational Chemistry (CC) arises as an update of theoretical chemistry and is based on various computational tools such as mathematical algorithms, statistics, and databases to model systems or calculate properties at the molecular level [3].

From the framework of Technological Pedagogical Science Knowledge (TPASK) [4], CC represents a body of knowledge that emerges from the interaction between technological

knowledge (TK) and science knowledge (SK), that is, technological science knowledge (TSK). In this sense, CC is positioned as a dynamic element since advances in computational methods represent a form of emerging technology that has the potential to be used in teaching chemistry. In this regard, Tuvi-Arad and Blonder [5] highlight the importance of chemical compound databases, which can serve as valuable resources for teaching chemistry. Similarly, Lehtola and Karttunen [6] acknowledge that CC now offers a vast array of open-source and freely available software, enabling its integration with Massive Open Online Courses and thus reaching a wide range of students.

Based on the TPASK framework, Rodríguez-Becerra et al. [7] propose the Technological Pedagogical Chemistry Knowledge, which allows for the direct linking of CC, or Technological Chemistry Knowledge, with Pedagogical Chemistry Knowledge. The intersection of these bodies of knowledge accounts for the educational CC construct, which is understood as the knowledge that a teacher possesses to design learning environments based on activities, resources, and educational strategies that incorporate CC tools and allow for the learning of chemistry in their students.

One way to acquire this emerging knowledge is through e-learning courses, which have become a key tool in modern education, offering a wide range of opportunities for distance education, which gained greater prominence during the COVID-19 pandemic [8]. E-learning environments allow users to access educational content from anywhere and anytime using only the internet. In addition, they enable teachers to design courses using various tools to manage learning by integrating innovative technologies and multimedia content [9].

E-Learning Educational Computational Chemistry Course

The e-learning Educational Computational Chemistry Course (EECCC) took place in January 2023 and lasted five weeks. This free and self-paced course is hosted on <https://mooc.umce.cl> (accessed on 1 January 2023) and was promoted through social media platforms (Facebook) in science teacher discussion groups. It attracted 180 enrollments, of which 108 participants completed at least one activity in the course. Upon entering the EECCC, teachers agreed to an informed consent form, which outlined utilising all information generated during the course for research purposes. This consent form underwent evaluation and approval by an accredited ethics committee in Chile.

The instructional design of the EECCC has been published by our research group [9] and consists of three specific modules (Table 1):

- Module I: CC for Science Education is based on using research-grade computational chemistry software (RGCCS) to edit, create, and test 3D chemical structures linked to drug design.
- Module II: Use databases of chemical compounds, visualisers, and virtual screening to solve problem scenarios.
- Module III: Development of experiences using the Problem-Based Learning (PBL) methodology, problem scenarios, and pedagogical aspects.

The closing evaluation of the course reported on the preparation and planning of learning activities framed in a PBL environment. For this, it was suggested to consider the following aspects:

- (a) Description of the stages of the PBL methodology for future implementation.
- (b) Approach to a problem scenario, which should contemplate a socio-scientific theme that could be solved through CC, main background, and research question.
- (c) Identification of the educational sector in which it will be applied, indicating subjects and learning objectives according to the local curricula, the number of students, technical requirements, application time, etc.
- (d) All the necessary information must be provided to help an instructor implement the elaborated activity and the evaluations that she will use during the process.
- (e) It must incorporate all the necessary reference material to complement the activity: videos, websites, computer programmes, articles, book chapters, and other documents.

Table 1. E-learning Educational Computational Chemistry Course design.

Module	Specific Topics/Contents	Activity Type	Evaluation
I Introduction to computational chemistry for science education.	• Introduction to computational chemistry tools for teaching sciences. • Molecular editing and visualisation with Avogadro software. • Discovery and design of new drugs.	• Draw a structure with Avogadro software. • Identify hydrogen donors/acceptors. • Lipinski rule application.	Identification of potential drugs for the treatment of COVID-19.
II Virtual screening, visualisers, and molecular editors in pedagogical contexts.	• Online databases of protein structures. • Introduction to molecular docking. • Molecular docking with Autodock software and MGLTools. • Visualisation of results with Autodock and Discovery Studio software.	• Molecular docking – File preparation. – Results. – Analysis of results.	
III Fundamentals of PBL.	• Introduction to PBL. • PBL design. • Role of teacher and students.	• PBL scenario example “Requirement of the World Health Organization in Chagas disease”. – Stages of development of PBL. – Expected results.	Planning of learning activities framed in a PBL environment.

An example of a learning activity plan is presented in Supplementary Materials (Tables S1–S5).

Each module included various activities as well as interactive challenges and evaluations. The course aimed to provide teachers with comprehensive knowledge and practical skills for incorporating CC into their teaching practises. Offering a flexible and accessible learning environment allowed participants to engage in self-directed learning, fostering their professional development in the field of educational computational chemistry and contributing to the development of their TPASK.

1.2. Educational Data Mining in E-Learning Environments

Data mining is a discipline that uses computational tools to analyse massive datasets to identify patterns, trends, and relationships that can be used for decision-making. This discipline has been highly relevant due to the high rate of information generation in all areas of society, encompassing economic, governmental, medical, and educational aspects, among many others.

Educational Data Mining (EDM) in e-learning environments has generated increasing interest due to the large amount of information generated in learning management systems, including participant characteristics, learning outcomes, and event records [10,11]. Analysing this information can improve the quality of online education by providing valuable information about the effectiveness of educational strategies, predictive performance models, student satisfaction, or technology tools [12]. There are various methods and tools for EDM [13], but they can be classified into three types:

1. Classification is a procedure that consists of grouping individual elements into categories based on the analysis of quantitative information on one or more characteristics inherent to these elements using a training set composed of previously labelled components. Predicting student performance or retention/dropout in a particular course is possible from these categories. Some of the most commonly used classification algorithms are K-Nearest Neighbours, Decision Tree (DT), Naïve Bayes, Support Vector Machine (SVM), and Random Forest (RF).
2. Clustering is a technique that classifies students based on their learning and interaction patterns. This technique has recurrent applications in various fields, such as resource recommendation, understanding learning processes, and preventing academic failure, especially in the university environment. Some of the most commonly used clustering algorithms are Hierarchical clustering and K-means.

3. Regression is a technique that allows predicting a range of numerical values from a specific dataset. The regression analysis has been used in various applications, including predicting student academic performance and how accurately they will answer particular questions. Additionally, regression has been applied to model user learning behaviour, making it a valuable tool for understanding the cognitive processes associated with knowledge acquisition.

Regarding the student success rate, Gaftandzhieva et al. [12], using a database of the performance of 105 students in an e-learning programming course, reported a significant correlation between the final grades of the students and their activity in the e-learning course, as well as between the final marks and their participation in virtual classes. The RF algorithm showed a prediction accuracy of 78%, a practical approach to predicting which students might fail after a given time. In line with Al-Kindi and Al-Khanjari's [14] proposal, using a database of a course supported in Moodle with more than 250,000 records, we carried out a comparative study between the Naive Bayes algorithm and RF and determined that the latter has an accuracy of 97% of correctly predicted instances, thus determining that it is an ideal algorithm to predict the performance of course participants.

Concerning the learning environment, Davies et al. [11], using a database of an e-learning course on information systems together with the K-means clustering algorithm, identified the strategies used by the students in their participation in the course, observing that during the simplest activities, they did not use resources such as videos or explanatory texts. However, during the most complex activities, they used all available tools. These findings can serve as a basis for the development of new orientations in the optimisation of the instructional design of the course.

Based on the above, this research focuses on a case study that reports on the results obtained in an e-learning training instance for in-service chemistry teachers that focused on using Educational Computational Chemistry and whose design can be optimised using data mining techniques and the participant's perception of the course. The research questions for this study are as follows:

RQ1. How do in-service chemistry teachers integrate the design principles that guide them to use CC tools in planning learning activities wherein they "teach with" rather than "from" CC?

RQ2. How can data mining techniques assist in the design of e-learning environments for in-service chemistry teachers?

2. Methods

2.1. Sample Characterisation

The characterisation of the sample was conducted by applying an initial questionnaire completed by the course participants. The questionnaire was designed by the researchers and consisted of four main sections: (i) personal and demographic information, (ii) academic information, (iii) employment information, and (iv) information regarding the use of Information and Communication Technologies, based on the proposal by Esfijani and Zamani [15] (for the whole questionnaire, review Supplementary Materials Table S6).

- Demographic information

The 108 teachers who participated in the EECCC course are geographically located on the American continent and come from the countries of Chile, Colombia, the United States, Ecuador, Uruguay, and Venezuela. Of these countries, Chile hosts 93.5% of the participants, followed by Ecuador with 1.9%, and the remaining countries with 0.9% each.

- Gender, age, and education

The sample is predominantly composed of female teachers, accounting for 67.6%. In terms of age range, the entire sample falls within the 25- to 40-year-old bracket (91.6%). Regarding educational aspects, 54.6% have not pursued any postgraduate programmes after completing their undergraduate studies, while the remaining 45.4% have undertaken

diploma programmes (19.4%), postgraduate studies (6.5%), or master's degrees (19.4%). Concerning years of service, 73.1% have less than ten years of professional experience.

- Knowledge, use, and technological access

Regarding their initial training, 46.3% of the participants have not completed any courses related to the use of technology. However, during their professional practise, 47.2% reported having participated in courses related to the use of technology. In terms of general knowledge about the characteristics of the computer they regularly use, such as memory capacity or speed, 70.3% of the participants claimed to be familiar with them.

All participants had 100% access to a computer and the internet at home, while 91.6% reported having access at their workplace. As for other peripherals, such as printers or scanners, 81.2% reported having access to them.

Regarding the most frequent use of technologies, participants predominantly mentioned the use of email and file sharing (84.2%), chat platforms and social networks (80.6%), and slideshow presentations (80.2%).

2.2. In-Service Chemistry Teachers' Perception (RQ1)

A content analysis was conducted on the learning activity plans generated by the teachers who participated in the EECCC. Furthermore, a methodological triangulation was performed with the teachers' perceptions regarding the constructs associated with TPASK framework and the integration of CC. A survey and a focus group were used to assess these dimensions:

1. The survey consists of 48 questions evaluated on a five-point Likert scale, aiming to obtain balanced responses and avoid the possibility of neutral answers, thereby compelling participants to take definite positions when responding. Using the Likert scale allowed for quantitatively measuring attitudes, opinions, and perceptions [16].

An example of one of the questions included in the survey focused on evaluating Technological Pedagogical Knowledge (TPK). The question was: "Can I adapt the digital technologies I am learning to different teaching activities, such as designing and planning activities, classroom management and organisation, and evaluation processes, among others?" In response to this question, participants were required to indicate their level of agreement or disagreement based on constructs related to the TPASK framework. The complete set of questions and their corresponding constructs is presented in Supplementary Materials Table S7.

This instrument was developed based on the proposals of Habibi et al. [17], Lin et al. [18], and Schmidt et al. [19], and presented the following Cronbach's Alpha coefficient values for each construct: 0.840 (Pedagogical Knowledge, PK), 0.826 (SK), 0.915 (TK), 0.860 (Pedagogical Science Knowledge, PSK), 0.967 (TSK), 0.950 (TPK), and 0.968 (TPASK). Additionally, the instrument presented a weighted Cronbach's alpha of 0.977, indicating high internal consistency.

The results were compiled from the responses for each indicator and organised into the previously mentioned constructs.

2. The focus group was conducted to gather information about participants' perceptions of the e-learning module on educational computational chemistry and its impact on learning and future teaching endeavours. The focus group allows for an in-depth and detailed exploration of the topic, as discussions and idea exchanges among participants can provide new perspectives, perceptions, and nuances. The focus group included open-ended and closed-ended questions related to the TPASK framework. For example, regarding knowledge in the sciences, a question used was, "Did you acquire or develop scientific skills during the completion of the module? If so, what were those skills?" [9].

This focus group was conducted virtually once the EECCC concluded and was moderated by an independent researcher. The proceedings were recorded on video and meticulously transcribed to ensure the precision and reliability of the collected data.

The research participants comprise a cohort of 108 in-service teachers who participated in the EECCC. A subsample of participants was selected to conduct the focus group. Participants were chosen based on the following criteria: participants with over three years of teaching experience, participants with a limited level of knowledge associated with elements of CC [9], and participants who completed at least 90% of the activities proposed in the course. Ultimately, the focus group consisted of six teachers identified using a code (Id_i; i = 1–6).

The data analysis was conducted using qualitative content [20], categorising the expressions of the participants based on the research question. The analysis process's validity and reliability were confirmed by having another researcher repeat the categorisation process using the interrater reliability method. The interrater reliability was calculated as the average Kappa value among three reviewers (J.H.-R., L.C.-J. and J.R.-B.). An interrater reliability value above 0.8 is considered a strong agreement [21].

2.3. Compilation of Records and Database

This study used all the records from teachers' interactions with the Learning Management System (Moodle). These records incorporate information related to the frequency of access to the platform, interactions with each section, and learning results. This information was compiled, generating a database of 9046 entries.

Likewise, additional variables were incorporated that address aspects related to the sample characterisation described in Section 2.1.

This methodology generated a comprehensive database containing detailed information about the participants' characteristics and interactions during the course (Table 2).

Table 2. List of the variables used in this study.

Variable	Data Type	Description
V1	Numeric	Final grade of the course
V2	Numeric	Participation in course activities
V3	Categorical	Gender, Female = 1, Male = 0
V4	Numeric	Age
V5	Categorical	No further study, Yes = 1, No = 0
V6	Categorical	Diploma studies, Yes = 1, No = 0
V7	Categorical	Postgraduate studies, Yes = 1, No = 0
V8	Categorical	Master's studies, Yes = 1, No = 0
V9	Categorical	Years in service (0 a 2), Yes = 1, No = 0
V10	Categorical	Years in service (3 a 5), Yes = 1, No = 0
V11	Categorical	Years in service (6 a 10), Yes = 1, No = 0
V12	Categorical	Years in service (>10), Yes = 1, No = 0
V13	Categorical	Access to computer at home, Yes = 1, No = 0
V14	Categorical	Access to internet at home, Yes = 1, No = 0
V15	Categorical	Access to computers at work, Yes = 1, No = 0
V16	Categorical	Access to internet at work, Yes = 1, No = 0
V17	Categorical	Hardware knowledge, Yes = 1, No = 0
V18	Categorical	Technology courses (pre-service), Yes = 1, No = 0
V19	Categorical	Technology courses (in-service), Yes = 1, No = 0
V20	Numeric	Total courses taken
V21	Numeric	Computer use
V22	Numeric	Event log (Total)
V23	Numeric	Event log (M1) *
V25	Numeric	Event log (M1.1) *
V25	Numeric	Event log (M1.2) *
V26	Numeric	Event log (M2) *
V27	Numeric	Event log (M2.1) *
V28	Numeric	Event log (M2.2) *
V29	Numeric	Event log (M2.3) *
V30	Numeric	Event log (M2.4) *

Table 2. Cont.

Variable	Data Type	Description
V31	Numeric	Event log (M3) *
V32	Numeric	Event log (M3.1) *
V33	Numeric	Event log (M4) *
V34	Categorical	Performance, If scores ≥ 5.50 , High; Low otherwise

* “Event Log (M.X)” gives an account of the activity/assessment number X belonging to each module (M).

2.4. Educational Data Mining (RQ2)

The database indicated in the previous point was processed using the R statistical software package (v.4.1.0) [22] together with the R Analytical Tool To Learn Easily (Rattle) [23] and RGtk2 [24] packages, which incorporate a series of functions that allowed data mining computations through a graphical interface.

The database was loaded into the Rattle package interface, followed by an initial verification to determine the variables to consider in the modelling stage. Specifically, the variables that presented constant values were discarded, and the variable V34, which reflects the success of the course, was selected as the objective variable. In addition, the database was partitioned into three sets using a random strategy: a training set that covered 70%, a test set that represented 15%, and a validation set with 15% remaining.

Subsequently, the algorithms were used: (i) Classification and regression trees (CART), (ii) RF, and (iii) SVM, to generate models that allow predicting the level of achievement of the participants as well as reveal guidelines that contribute to improving the instructional design of the EECCC.

3. Results and Discussion

At the end of the course, the group of participants with the lowest results and who left the course early was made up of teachers with an average age of 39 years, of whom 44.4% have more than ten years of teaching experience and 55.6% have not continued higher education than their initial training. This group is made up mostly of women, who represent 88.9%.

The group of participants who passed the course are mainly young teachers with an average age of 32 years, of whom 44.4% have less than five years of teaching experience and 55.6% have not studied beyond their initial formation. Regarding Gender, 55.6% of those approved were men. In this last group, their planning of learning activities was analysed.

3.1. Integration of CC in the Planning of Learning Activities (RQ1)

3.1.1. Examination of Teaching Plans

The problem scenario presented by the participating teachers in the course mainly revolved around socio-scientific issues related to health topics, such as diseases caused by HIV, the Dengue virus, and other infectious agents. Consistent with the findings previously reported by our research group, this highlights that the health field provides an important source of real-world problems with significant educational potential [2]. In these scenarios, teachers contextualise the problem and invite their students to develop potential solutions.

For example, one of the questions proposed by a participating teacher to guide the learning activity was:

“Are there other candidate compounds to block the binding of HIV to the CCR5 receptor? As a working group, generate scientific evidence that answers this research question.”

To address this question, the teacher suggests the use of various computational chemistry tools, such as databases of chemical compounds like PubChem (<https://pubchem.ncbi.nlm.nih.gov>, accessed on 1 January 2023) and DrugBank (<https://go.drugbank.com>, accessed on 1 January 2023), as well as 3D structure visualisation software (Avogadro, <https://avogadro.cc>, accessed on 1 January 2023) and molecular docking (AutoDock, <https://autodock.scripps.edu>, accessed on 1 January 2023). In terms of supporting information for future teachers applying this material, elements related to potential objectives that

can be pursued are incorporated, along with a comprehensive molecular docking protocol that indicates the use of protein structures extracted from the Protein Data Bank website (<https://www.rcsb.org>, accessed on 1 January 2023).

In general, the participating teachers suggested in their lesson plans that they use the same CC tools originally introduced in the EECCC. While it is interesting to apply these tools to new problem scenarios, it also highlights the lack of knowledge among teachers regarding new tools. Therefore, it becomes evident that teachers need to understand the computational tool they intend to use. They should experiment with different datasets, explore additional functionalities, and only employ them with their students when they feel adequately prepared.

3.1.2. TPASK Survey

The teachers' perception of the development of the constructs associated with the TPASK framework is shown in Table 3, which shows the total percentage of adherence for each construct based on the responses of the EECCC participants.

Table 3. Perception of teachers in the TPASK framework.

Knowledge Type	Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
PK	27.8%	61.1%	9.7%	1.4%	0.0%
TK	42.2%	42.2%	11.1%	4.4%	0.0%
SK	31.1%	55.6%	13.3%	0.0%	0.0%
PSK	26.4%	59.7%	12.5%	1.4%	0.0%
TPK	25.0%	68.1%	6.9%	0.0%	0.0%
TSK	20.6%	28.6%	25.4%	25.4%	0.0%
TPASK	30.2%	58.7%	11.1%	0.0%	0.0%

Regarding the survey described in Section 2.2, teachers perceive extensive dominance around PK, TK, and SK, expressing substantial unanimity in the two categories with the best scores, with an approval rate of 88.9%, 84.4%, and 86.7%, respectively. Following the characterisation mentioned in Section 2.1, in-service teachers generally present an important perception in accordance with the elementary constructs of the TPASK framework, such as PK, TK, and SK, separately.

Concerning the PSK, this construct presents an approval level of 86.1%. The relationship between pedagogical and scientific aspects is established directly, as evidenced in the sample analysis, where the presence of teachers with experience in the educational system is verified. Furthermore, this correlation is supported by information documented in the literature, which maintains that PSK improves with teaching experience. In this sense, it is observed that more experienced teachers have a stronger PSK compared with novice teachers [25].

Regarding the TPK, various authors have highlighted its importance for developing digital skills in education. According to Rodríguez-Becerra et al. [7], “TPK is a knowledge about the possibilities and challenges implied on different ways to teach and learn”. In this sense, it is relevant to highlight the high level of assessment observed in this category, with 93.1% approval in its two highest Likert categories. This value agrees with the findings reported by teachers in post-pandemic service because the need to reorganise traditional classes into online classes motivated teachers to incorporate and develop elements linked to TPK that enhanced their skills in incorporating technologies in their teaching practise [26].

On the other hand, according to the TSK, there is evidence of an approval level of 49.2%, accompanied by 25.4% of disapproval. This result suggests the existence of a low perception regarding the use of technologies that allow the development of different aspects of scientific knowledge. It is pertinent to highlight that, despite the participation of teachers in the instruction instance and their approval of the course, they still express the

feeling of requiring greater support in terms of technological aspects associated with the teaching objectives.

Regarding the development of the TPASK of teachers, it is evident the importance that is given to the integration of different skills and knowledge to optimise the teaching and learning process in the scientific field mediated by technologies. In this sense, the high degree of approval of 88.7% that is evidenced in teachers' assessments of the TPASK is significant. This indicates that, for the most part, teachers perceive that they have the necessary skills and knowledge to use methodologies or pedagogical strategies that allow the integration of knowledge of science, the use of digital technologies, and scientific computing.

3.1.3. Focus Group

Regarding the integration of CC tools in the development of the EECCC, the authors categorised the following text passages in the TPASK framework. This categorisation presents a Kappa value of 0.814, which indicates excellent agreement between the authors.

- TK and TPK

It is evident that several participants presented difficulties when trying to install one of the computer programmes used: *"(...) When I had to install the programs, some complexities arose and I had to ask the teacher for help". (Id_5). "(...) It happened to me too". (Id_2) and (Id_4).* However, the teachers did not try to solve the problem independently but requested technical support through internal messaging. Finally, the problem accounted for the absence of a dynamic-link library (dll) that could be easily solved based on TK. This evidence contrasts with what was reported through the survey, which accounted for a high percentage of perceptions of this body of knowledge.

On the one hand, there has been an excellent reception of the technological resources used in the educational field, such as videos mediated by artificial intelligence and chemical modelling. These tools have been positively received by teachers, who recognise their potential to enrich the teaching and learning processes: *"(...) I loved the use of artificial intelligence in the videos because it brings not only us but also the students closer to an artificial intelligence where interpretation is sought in the videos, which was also very in vogue (...) a little while ago, I don't know if today or yesterday I reminded myself that there was a presenter in the United States based on artificial intelligence, so each time it is going to be inserted more into our society and what better way to do it through education". (Id_3). "(...) I believe that the most essential skills that I can take away from the course are the use of models, the use of digitisation itself (...) is to put a new learning modality in front of a computer and front of the student, that it is not only in the notebook, in the book or even in the laboratory (...) now with digitisation everything is easier." (Id_3).*

However, a certain reservation has also been identified regarding content creation due to the time required to plan and prepare the materials supporting a possible implementation of similar activities. It is important to emphasise that the quality and relevance of the pedagogical resources play a fundamental role in the success of an implementation: *"(...) it is relevant to have the time available to generate all the materials such as the video tutorial to use the software, the didactic material to be used or the evaluation instruments (...) so I feel that these are like the difficulties that could have". (Id_2).*

"(...) as difficulties, I think it is time to design a module (...) the fact that one has to do all these activities from scratch and plan them, it could be a bit exhausting (...) there is a lot of work behind the implementation of a PBL environment". (Id_4)

In this context, the level of development of the TPK becomes a determining factor for the generation of effective pedagogical resources. Although the initial survey results indicated that teachers have adequate mastery of this construct, it is essential to continue promoting its further development due to the dynamic nature of technological tools.

- PSK

Concerning these categories, the teachers realise the multiple application options of the aspects evidenced in the development of the course: “(...) I was thinking of moving it to sciences for citizenship course, since they are cross-cutting issues, the use of medicines and then emphasising chemistry”. (Id_4). In the same way, the linkage of aspects linked to the course’s instructional design is evident, such as the methodology that incorporated some aspects of PBL, as was the work with learning scenarios based on socio-scientific problems: “(...) in science for citizenship course, it works perfectly because it works with active learning methodologies, and that is where the PBL is doing super well”. (Id_6).

In addition, teachers have expressed their positive perception of the interconnection of the contents addressed in the course, which are intrinsically linked through a socio-scientific problem. This perspective generates greater value and meaning in instruction since it allows students to understand the relevance and application of the knowledge acquired in real-world situations: “(...) The module had intermolecular forces, organic chemistry with the visualisation. It also had a topic of stoichiometry and how to calculate molar mass with the program. So, although perhaps it is content we see regularly at school, integrating or linking them to teach everything together was very enriching”. (Id_2). “(...) For me, the Lipinski rule was fascinating (...) One knows that drugs go through certain stages before being administered orally or in another way, but I had no idea about this rule (...) For me, it was a novelty”. (Id_3).

The articulation of contents is revealed as a fundamental element to ensure adequate course development. By intertwining different concepts and themes, a deeper and more holistic understanding of the content is fostered for students. In addition, this integration promotes critical thinking and the ability to solve complex problems by addressing situations in which it is necessary to apply knowledge from various areas.

- TSK

On the use of RGCCS, the participants give an account of their knowledge and use of these tools: “(...) I only knew the chemsketch, and now I realised that it was completely outdated with respect to all the ranges of programs that exist in general”. (id_2). “(...) The different stages of the modules will allow us to do different things, such as acquiring knowledge of technology with this autodock software or the use of these databases” (id_3). “(...) I liked working with real data, having to search in some specialised chemistry database, seeing the subject of Lipinski’s rule and all that”. (id_5). “(...) We are using technologies that are already available, we are using data that is available published on web pages, specialised in scientific dissemination with real data, data that at some point other scientists obtained in various ways and published” (id_4).

Similarly, there is notable recognition of free and open-access software, which is of significant value in the educational context. This strategic choice is based on the elimination of economic barriers and the possibility of accessing emerging technologies without the need to incur additional costs to acquire specific programmes:

“(...) all the software was open access (...) anyone could access the specific scientific information, totally free, they did not require registration or anything (...) it was downloaded from the official page and used immediately (...) even, the issue of databases could be done directly by accessing the internet (...) it can even be done with a mobile phone”. (Id_4)

Using technological tools such as Avogadro, Autodock, and Databases of chemical compounds expands the possibilities of exploration and experimentation, promoting critical thinking and creativity in learning chemistry: “(...) I think that the most useful thing was the visualisation using Avogadro (...) I had already used it a long time ago, but I couldn’t remember”. (Id_2). “(...) I think that where I learned the most was in the docking section, using Autodock (...) now, I have the confidence to use Autodock and follow the entire flow of the docking process”. (Id_1). “(...) search for data in a specialised chemistry database. Searching for a particular data of a compound and obtaining its properties no longer has any complexity. It is the minimum complexity of using a browser or a web page (...). When one uses scientific software, you can develop many skills that sometimes go beyond the specific contents of chemistry or biology, and it has to do with the use of technologies, which can be enriching”. (Id_4).

These findings are relevant since, in previous investigations, our research group has reported the importance of the development of TSK [7,27] since it is directly related to the development of scientific skills such as the interpretation of computational models, the use of scientific protocols, the work with databases and the visualisation in 3D, these tools allow to understand abstract knowledge and visualise through a complex screen process such as molecular docking.

- TPASK

The projection towards new formative instances in the particular contexts of the teachers is evident, mainly from the TSK to the PSK: “(…) Now I can project the visualisation of molecules in a class and explain a content” (Id_2). “(…) Since we are all teachers, we are all thinking about how to modify this, according to the technological tools or the resources we have in each school to implement it in the future” (Id_3). “(…) The fact of being able to complement my knowledge with the use of different software to be able to teach will be exciting for the students. They will be able to interact with some models and understand abstract concepts” (Id_5). “(…) In my class, I would start with the topic of databases and the calculation of properties using the software” (Id_4).

Finally, the integration of the bodies of knowledge is evidenced, as is how these interactions enable enriched learning environments: “(…) This was to connect technology with the class in a super coherent line of deepening” (Id_6). “(…) It helped me to learn computational chemistry, software and databases, and then move from that to teaching chemistry with computational tools” (Id_2).

Furthermore, one of the participants reports a significant contrast relating the classical practises of chemistry with the use of new emerging technologies, hitting the nail on the head: “(…) We totally got out of the paradigm of the scientist in the laboratory, but we made them do laboratory activities with a computer, and it is totally different from science class, with the teacher who wears the white coat here it was the same. Finally, we are using a computer that was our laboratory material, the computer” (Id_4). A temperature sensor in the laboratory was once cutting-edge technology. Nowadays, the use of RGCCS is at the forefront. Its use, together with active learning methodologies, is undoubtedly the path towards which the professional development of teachers should be directed.

3.1.4. Final Section Considerations

Based on the three inputs presented in Section 3.1 (survey, focus group, and lesson planning), it is possible to observe the integration of CC tools in the planning of learning activities in a concrete manner and through the teachers’ discourse.

In this regard, the PSK is significantly developed among the teachers. The selection of socio-scientific problems and adherence to a PBL methodology, with all its challenges (scenario generation, formulation of challenging questions, tool selection, among others), indicate that the teachers have educational experience. This assertion is supported by the high percentage presented by the construct in Table 3 and the highlighted passages of text in Section 3.1.3.

On the other hand, the use of computational chemistry tools is still at an early stage. While technological advancements have greatly contributed to the development of the field, these advancements are not aligned with the usability of computational tools related to science in educational environments. Currently, teachers appreciate the accessibility and the potential technological, pedagogical, and scientific contributions of computational chemistry. However, their usage is minimally evident in the planning of learning activities.

3.2. Model Generation Using Educational Data Mining (RQ1)

Models were generated from the three algorithms used for each model. The results obtained are described as follows:

3.2.1. Classification and Regression Trees

The model generated by the CART algorithm presented the following rules, where each rule corresponds to a path through the decision tree (Figure 1).

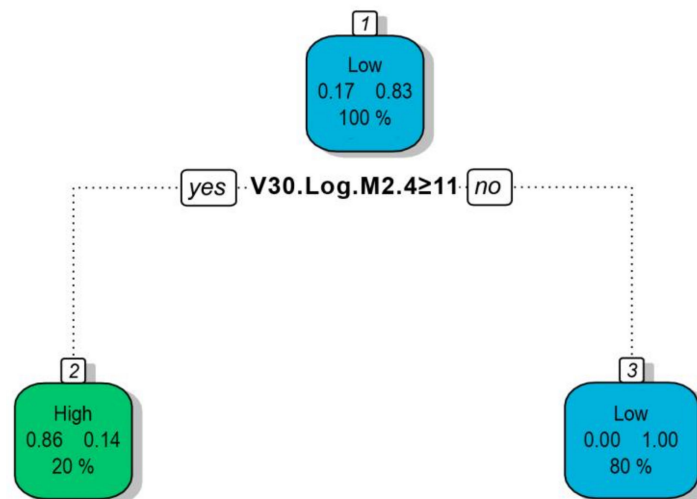


Figure 1. Decision Tree was made using the CART algorithm. This algorithm selected the variable V30 as the most important factor that determines the division of the data, and therefore its value determines the level of success in the course.

- Rule number: 3 [V34.perf = Low cover = 28 (80%) prob = 1.00] V30.Log.M2.4 < 10.5.
- Rule number: 2 [V34.perf = High cover = 7 (20%) prob = 0.14] V30.Log.M2.4 ≥ 10.5.

The CART algorithm accounts for a relevant stage in obtaining success in the development of the course. It is evident that the variable V30, which corresponds to the closing evaluation of modules I and II of the EECCC, is a key factor in the success of the course. This evaluation corresponds to developing a PBL module whose problem scenario is positioned towards the “identification of potential drugs for the treatment of COVID-19” through CC techniques. In this regard, the final evaluation of the course shows that only 10.7% of the participants completed this activity. Along the same lines, the events log indicates that 73.1% of all events recorded in this evaluation correspond to interactions carried out by teachers who completed the course.

The decision rules derived from the analysis indicate that success will be low for 100% of the sample (probability = 1.0) when the variable V30 has less than 10.5 points. Furthermore, this rule covers 80.0% of the training set, with a total of 28 observations that satisfy this condition. Consequently, it is necessary to generate more content and various supports or guides before the evaluation so the participants can solve it correctly.

3.2.2. Random Forest and Support Vector Machine

The model generated by the RF algorithm used 500 DT and presented an estimated error rate of 2.86%. Additionally, a confusion matrix that registers the disagreement between the final model’s predictions and the training set observations’ results was recorded. It can be observed that the model and the training dataset coincide in high performance for five comments and in low performance for 29 observations. However, there is one observation where the model predicts low performance and the observation reports high performance, resulting in an error rate of 16.7% of the observations. This contrasts with the 0% error rate when predicting high performance (Table 4).

Table 4. Confusion matrix for the Random Forest model.

Performance	High	Low	Class.Error
High	5	1	0.1666667
Low	0	29	0.0000000

Additionally, a classification was generated to evaluate the importance of the variables (Table 5). Each input variable has four importance measures, with higher values reflecting the relatively greater importance of the variable. Table 5 is ordered by the precision of importance measure, where the variable “Mean decrease accuracy” expresses how much accuracy the model loses by excluding each variable, while the variable “Mean decrease Gini” is a measure of how each variable contributes to the homogeneity of the nodes and leaves in the resulting RF.

Table 5. Variable importance based on Mean Decrease Accuracy for the ten best-scored variables (If scores ≥ 5.50 , High; Low otherwise).

Variable	High	Low	Mean Decrease Accuracy	Mean Decrease Gini
V2	8.43	8.27	8.46	1.18
V32	8.23	6.73	7.64	0.83
V33	7.88	5.25	7.33	0.68
V31	6.78	5.25	7.04	0.62
V30	6.03	4.81	5.91	0.51
V1	5.11	4.89	5.48	0.90
V22	4.33	2.84	3.91	0.34
V28	3.13	3.33	3.63	0.25
V29	3.60	1.76	3.49	0.32
V26	2.97	1.37	2.69	0.18

For the full table, review Supplementary Materials Table S8.

Concerning the model generated by the RF algorithm, a high precision was observed (97.1%), which made it possible to determine the relative importance of the variables in the study. According to the results obtained, the five most relevant variables are the general participation in the course (V2), the final evaluation corresponding to the planning of a PBL environment (V32), the final perception of the teachers (V33), the complete PBL module (V31), and the V30 that the DT had previously identified. This finding suggests that the active participation of the students, the final evaluation, and the teachers’ perception are crucial factors for the success of the e-learning courses.

The result of the SVM algorithm indicates the creation of 22 support vectors and a general prediction error of 0.0% on the training dataset. All predictions made for variable V34 in the training dataset are accurate.

3.2.3. Model Evaluation and Guidelines for EECCC Redesign

The models are evaluated using the dataset that was not considered in constructing the prediction models. Since the prediction was based on a binary variable (Performance), the evaluation of the models presents a two-by-two matrix of the predicted and actual values for the indicated variable (Supplementary Materials Table S9).

The overall prediction error for the decision tree-based model was 28.5%, and the average error of the variable categories (High/Low) was 20.0%. The error matrix shows that the “High” category prediction was accurate, while the “Low” category prediction was vague. On the other hand, the evaluation of the models proposed by RF and SVM did not present prediction precision errors for both variable categories (0.0% error).

The evaluation of the models generated by the RF and SVM algorithms revealed the best performance in predicting the participants’ success levels. This finding agrees with what was proposed by Bulut and Yavuz [28], who account for these algorithms’ high reliability and precision.

Based on these generated models, it is possible to establish guidelines that optimise the EECCC based on the trends reported by the algorithms. In this regard, the most important variables (see Table 5) are mainly related to event logs consistent with evaluations. This may seem trivial, but it is essential in an e-learning course because, unlike face-to-face or synchronous courses, this modality relies on participants' self-learning, where their self-efficacy, time management, motivation, and persistence become crucial. These results are consistent with previous studies that have found a positive relationship between student engagement and success in online learning [29]. In addition, they highlight the need to provide additional support to students in online learning, especially at critical moments such as assessments [30,31]. For example, regarding event logs, participants who completed the course had an average of 342 logs. In contrast, the average for students who were unsuccessful in the course was 145 logs, meaning they achieved approximately a 42% success rate.

To prevent this amount and promote success in the course's development, it is necessary to strengthen the instances before evaluations. Although the course design includes intermediate assessments (such as identifying compounds with pharmacological properties using databases of chemical compounds), it is necessary to reinforce their completion and ensure participants review the feedback generated through the internal messaging system to create a sense of accompaniment. Additionally, reviewing all evaluation instances is fundamental to determining specific difficulties in their completion.

4. Conclusions

The quality of student learning is related to the training and experience of teachers, which highlights the importance of continuous training and professional development to improve the quality of teaching. Therefore, it is necessary to consider practical strategies to promote teachers' participation in constant education and professional development, such as e-learning courses.

The present study focused on an online learning experience aimed at young teachers who lack experience in the school system and have a limited level of continuous training. It was observed that these teachers exhibited low usage of science-related technology tools, instead opting for office-related technological tools such as email and presentation slide applications.

Upon completion of the course, participants were surveyed about the constructs associated with Technological Pedagogical Science Knowledge, revealing significant development in terms of the potential integration of technological tools and research-grade computational chemistry software, as well as an interest in the potential use of problem-based learning as a methodology to enhance the quality of teaching and students' comprehension of scientific concepts. However, it is acknowledged that the practical application of this knowledge remains a challenge for many teachers. It is important to note that the mere use of these tools does not necessarily indicate a deep understanding of the potential of technology in teaching and learning. Therefore, it is crucial for teachers to acquire technical skills and a critical understanding of the use of technology in the classroom and its impact on student learning. In this context, computational chemistry emerges as a discipline that offers a wide range of valuable tools and resources for teaching chemistry.

Regarding improving and optimising the course, the analysis of educational data through data mining reveals the importance of evaluative instances and teachers' commitment and active participation in course development. The need for ongoing monitoring of the course's progress by teachers is emphasised. Educational data mining enables the identification of the precise inflexion point that determines the success or failure of the course. In this regard, it is essential to identify these emerging patterns and focus on them during the redesign, analysis, and optimisation of future e-learning educational computational chemistry course implementations.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/educsci13080796/s1>, Table S1: Learning activity planning—Overview. Table S2: Learning activity planning—Didactic sequence. Table S3: Learning activity planning—Problem scenario. Table S4: Learning activity planning—Instructor Support Information. Table S5: Learning activity planning—Evaluation rubric, investigation report. Table S6: Characterisation questionnaire. Table S7: TPASK Survey. Table S8: Variable importance for all variables used in this study. Table S9: Error matrix for the generated models. References [32–34] are cited in the Supplementary Materials.

Author Contributions: Conceptualization, J.H.-R. and J.R.-B.; methodology, J.H.-R. and J.R.-B.; formal analysis, J.H.-R., L.C.-J. and J.R.-B.; writing of the first draft of the manuscript, J.H.-R.; writing-revising and editing of the manuscript, L.C.-J. and J.R.-B.; supervision, J.R.-B.; project administration, J.R.-B.; funding acquisition, J.R.-B. and L.C.-J. All authors have read and agreed to the published version of the manuscript.

Funding: The author J.H. thanks the Programa Extraordinario de Becas de Postgrado—Doctorado en Educación—UMCE and the Proyecto FONDECYT Regular-1221942 and FONDECYT Regular 1221634.

Institutional Review Board Statement: The study was approved by the Comité de Ética Institucional, Universidad Santiago de Chile, approval code 158/2022, approved on 19 April 2022.

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

Data Availability Statement: Not applicable.

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

References

1. Cannon, A.S.; Anderson, K.R.; Enright, M.C.; Kleinsasser, D.G.; Klotz, A.R.; O'Neil, N.J.; Tucker, L.J. Green Chemistry Teacher Professional Development in New York State High Schools: A Model for Advancing Green Chemistry. *J. Chem. Educ.* **2023**, *100*, 2224–2232. [\[CrossRef\]](#)
2. Hernández-Ramos, J.; Perna, J.; Cáceres-Jensen, L.; Rodríguez-Becerra, J. The Effects of Using Socio-Scientific Issues and Technology in Problem-Based Learning: A Systematic Review. *Educ. Sci.* **2021**, *11*, 640. [\[CrossRef\]](#)
3. Parrill, A.L.; Lipkowitz, K.B. *Reviews in Computational Chemistry*; Wiley Online Library: Hoboken, NJ, USA, 2022; Volume 32.
4. Jimoyiannis, A. Designing and implementing an integrated technological pedagogical science knowledge framework for science teachers professional development. *Comput. Educ.* **2010**, *55*, 1259–1269. [\[CrossRef\]](#)
5. Tuvi-Arad, I.; Blonder, R. Technology in the Service of Pedagogy: Teaching with Chemistry Databases. *Isr. J. Chem.* **2019**, *59*, 572–582. [\[CrossRef\]](#)
6. Lehtola, S.; Karttunen, A.J. Free and open source software for computational chemistry education. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2022**, *12*, 33. [\[CrossRef\]](#)
7. Rodríguez-Becerra, J.; Cáceres-Jensen, L.; Díaz, T.; Druker, S.; Bahamonde, V.; Perna, J.; Aksela, M. Developing technological pedagogical science knowledge through educational computational chemistry: A case study of pre-service chemistry teachers' perceptions. *Chem. Educ. Res. Pract.* **2020**, *21*, 638–654. [\[CrossRef\]](#)
8. Adedoyin, O.B.; Soykan, E. COVID-19 pandemic and online learning: The challenges and opportunities. *Interact. Learn. Environ.* **2020**, *31*, 863–875. [\[CrossRef\]](#)
9. Hernández-Ramos, J.; Rodríguez-Becerra, J.; Cáceres-Jensen, L.; Aksela, M. Constructing a Novel E-Learning Course, Educational Computational Chemistry through Instructional Design Approach in the TPASK Framework. *Educ. Sci.* **2023**, *13*, 648. [\[CrossRef\]](#)
10. Hachicha, W.; Ghorbel, L.; Champagnat, R.; Zayani, C.A.; Amous, I. Using Process Mining for Learning Resource Recommendation: A Moodle Case Study. *Procedia Comput. Sci.* **2021**, *192*, 853–862. [\[CrossRef\]](#)
11. Davies, R.; Allen, G.; Albrecht, C.; Bakir, N.; Ball, N. Using Educational Data Mining to Identify and Analyse Student Learning Strategies in an Online Flipped Classroom. *Educ. Sci.* **2021**, *11*, 668. [\[CrossRef\]](#)
12. Gaftandzhieva, S.; Talukder, A.; Gohain, N.; Hussain, S.; Theodorou, P.; Salal, Y.K.; Doneva, R. Exploring Online Activities to Predict the Final Grade of Student. *Mathematics* **2022**, *10*, 3758. [\[CrossRef\]](#)
13. Williams, G. *Data Mining with Rattle and R: The Art of Excavating Data for Knowledge Discovery*; Springer Science & Business Media: Berlin, Germany, 2011.
14. Al-Kindi, I.; Al-Khanjari, Z. A Comparative Study of Classification Algorithms of Moodle Course Logfile using Weka Tool. *Int. J. Comput. Their Appl.* **2022**, *29*, 202–211.
15. Esfijani, A.; Zamani, B.E. Factors influencing teachers' utilisation of ICT: The role of in-service training courses and access. *Res. Learn. Technol.* **2020**, *28*, 2313. [\[CrossRef\]](#)

16. Jebb, A.T.; Ng, V.; Tay, L. A Review of Key Likert Scale Development Advances: 1995–2019. *Front. Psychol.* **2021**, *12*, 637547. [CrossRef]
17. Habibi, A.; Yusop, F.D.; Razak, R.A. The dataset for validation of factors affecting pre-service teachers' use of ICT during teaching practices: Indonesian context. *Data Brief* **2020**, *28*, 104875. [CrossRef]
18. Lin, T.; Tsai, C.; Chai, C.; Lee, M. Identifying science teachers' perceptions of technological pedagogical and content knowledge (TPACK). *J. Sci. Educ. Technol.* **2013**, *22*, 325–336. [CrossRef]
19. Schmidt, D.A.; Baran, E.; Thompson, A.D.; Mishra, P.; Koehler, M.J.; Shin, T.S. Technological pedagogical content knowledge (TPACK) the development and validation of an assessment instrument for pre-service teachers. *J. Res. Technol. Educ.* **2009**, *42*, 123–149. [CrossRef]
20. Forman, J.; Damschroder, L. Qualitative content analysis. In *Empirical Methods for Bioethics: A Primer*; Emerald Group Publishing Limited: Bingley, UK, 2007; Volume 11, pp. 39–62.
21. McHugh, M.L. Interrater reliability: The kappa statistic. *Biochem. Medica* **2012**, *22*, 276–282.
22. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021.
23. Williams, G. Rattle: A data mining GUI for R. *R J.* **2009**, *1*, 45–55.
24. Lawrence, M.; Lang, D.T. RGtk2: A graphical user interface toolkit for R. *J. Stat. Softw.* **2010**, *37*, 1–52. [CrossRef]
25. Krepf, M.; Plöger, W.; Scholl, D.; Seifert, A. Pedagogical content knowledge of experts and novices—What knowledge do they activate when analysing science lessons? *J. Res. Sci. Teach.* **2018**, *55*, 44–67. [CrossRef]
26. Duc, T.; Hop, N.; Dung, T.; Ha, V. The Effectiveness of Chemistry e-Teaching and e-Learning during the COVID-19 Pandemic in Northern Viet Nam. *Int. J. Inf. Educ. Technol.* **2022**, *12*, 240–247. [CrossRef]
27. Cáceres-Jensen, L.; Rodríguez-Becerra, J.; Jorquera-Moreno, B.; Escudey, M.; Druker-Ibañez, S.; Hernández-Ramos, J.; Díaz-Arce, T.; Perna, J.; Aksela, M. Learning Reaction Kinetics through Sustainable Chemistry of Herbicides: A Case Study of Preservice Chemistry Teachers' Perceptions of Problem-Based Technology Enhanced Learning. *J. Chem. Educ.* **2021**, *98*, 1571–1582. [CrossRef]
28. Bulut, O.; Yavuz, H.C. Educational data mining: A tutorial for the rattle package in R. *Int. J. Assess. Tools Educ.* **2019**, *6*, 20–36. [CrossRef]
29. Shea, P.; Bidjerano, T. Learning presence: Towards a theory of self-efficacy, self-regulation, and the development of a communities of inquiry in online and blended learning environments. *Comput. Educ.* **2010**, *55*, 1721–1731. [CrossRef]
30. Garrison, D.R.; Vaughan, N.D. *Blended Learning in Higher Education: Framework, Principles, and Guidelines*; John Wiley & Sons: Hoboken, NJ, USA, 2008.
31. Kizilcec, R.F.; Piech, C.; Schneider, E. Deconstructing disengagement: Analysing learner subpopulations in massive open online courses. In Proceedings of the Third International Conference on Learning Analytics and Knowledge, Arlington, TX, USA, 13–17 March 2013; pp. 170–179.
32. Rao, P.K. CCR5 inhibitors: Emerging promising HIV therapeutic strategy. *Indian J. Sex. Transm. Ted Dis. AIDS* **2009**, *30*, 1–9. [CrossRef]
33. Tan, Q.; Zhu, Y.; Li, J.; Chen, Z.; Han, G.W.; Kufareva, I.; Li, T.; Ma, L.; Fenalti, G.; Li, J.; et al. Structure of the CCR5 chemokine receptor-HIV entry inhibitor maraviroc complex. *Science* **2013**, *341*, 1387–1390. [CrossRef]
34. Motivational Video. Available online: <https://www.youtube.com/watch?v=v2PcQ-449p4> (accessed on 1 January 2023).

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.