education sciences

Assessment and Evaluation in Higher Education

Edited by
Sandra Raquel Gonçalves Fernandes,
Marta Abelha and Ana Teresa Ferreira-Oliveira

Printed Edition of the Special Issue Published in *Education Sciences*

www.mdpi.com/journal/education

MDPI

Assessment and Evaluation in Higher Education

Assessment and Evaluation in Higher Education

Editors

Sandra Raquel Gonçalves Fernandes
Marta Abelha
Ana Teresa Ferreira-Oliveira

MDPI • Basel • Beijing • Wuhan • Barcelona • Belgrade • Manchester • Tokyo • Cluj • Tianjin

Editors

Sandra Raquel Gonçalves Fernandes
Universidade Portucalense Infante D. Henrique
Portugal

Marta Abelha
Universidade Aberta (UAb)
Portugal

Ana Teresa Ferreira-Oliveira
CISAS, Instituto Politécnico de Viana do Castelo
Portugal

Editorial Office

MDPI
St. Alban-Anlage 66
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Education Sciences* (ISSN 2227-7102) (available at: https://www.mdpi.com/journal/education/special_issues/assess_educ).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

LastName, A.A.; LastName, B.B.; LastName, C.C. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-0365-6750-1 (Hbk)
ISBN 978-3-0365-6751-8 (PDF)

Contents

About the Editors

Sandra Raquel Gonçalves Fernandes

Sandra Raquel Gonçalves Fernandes, PhD., is Assistant Professor of the Department of Psychology and Education, at Portucalense University, Porto, Portugal. She holds a PhD in Education Sciences, specialization in Curriculum Development, at the University of Minho. Graduate degree in Education, by the same University. She is the Coordinator of the second cycle of studies in Education Administration and Management at Portucalense University. She coordinates the Pedagogic Innovation Office (Gabinete de Inovação Pedagógica—GIP) at Universidade Portucalense, an institutional structure aimed to promote the improvement of the quality of teaching through pedagogical training and the professional development of teachers. She is an integrated member of Portucalense Institute for Human Development (INPP) of Universidade Portucalense and an external collaborating member of the Center for Research in Child Studies (CIEC) of the University of Minho. She is member of several national and international research projects, with public funding. She is an Associate Editor of the Journal Teachers and Teaching: Theory and Practice and member of the Editorial Board of Frontiers in Education and Education Sciences (MDPI). She is a founding member and vice president of the Project Approaches in Engineering Education (PAEE) Association. Her research interests include Teacher Education, Higher Education, Curriculum Development, Education Administration and Management, Teacher Performance Assessment, Active Learning, Project-based Learning (PBL), Engineering Education, among others.

Marta Abelha

Marta Abelha, Ph.D., is Assistant Professor of Department of Education and Distance Learning (DEED), at Universidade Aberta (UAb), Portugal. She teaches in Higher Education, since 2011. She holds a PhD in Didactic (2011), from the University of Aveiro, Portugal. Concluded her Master in Curriculum Management (2005) and finished her Graduate Degree in Biology & Geology Teaching (2002) at the same University (University of Aveiro). She is an integrated researcher at UID-FCT n° 4372, Laboratory of Distance Education and eLearning (LE@D) of Universidade Aberta (Portugal). Additionally, she is a collaborating researcher at UID 00460/FCT, Centre for Interdisciplinary Studies (CEIS20), a research unit of the University of Coimbra, Portugal. She has been involved in national and international Educational Research Projects, with public funding. Her research interests focus on Teacher Education, Higher Education, Curriculum Studies, Teacher Collaboration, Assessment and Evaluation in Education, Active Learning, which are also the topics of her publications.

Ana Teresa Ferreira-Oliveira

Ana Teresa Ferreira-Oliveira, PhD., is an Assistant Professor at Escola Superior de Tecnologia e Gestão, Instituto Politécnico de Viana do Castelo, Portugal. She has a Pos Doc in Management and Economics from School of Economics and Management of the University of Porto and a PhD in Work and Organizational Psychology by University of Minho. She is a researcher at CISAS and a professor at Escola Superior de Tecnologia e Gestão, Instituto Politécnico de Viana do Castelo. Her main research topic includes trust, sustainability, agro-food systems, sustainable organizations, human resource management, behavior, equality, diversity and organizational change in the green deal society; social impact. She coordinates, at IPVC, an Horizon Project funded by the European Comission, FEAST—Food systems that support transitions to healthy and sustainable diets and the project" and a project coordinated by Centre for Educational Research and Innovation (CERI)

OECD Directorate for Education and funded by the Ministry of Higher Education "Fostering and assessing students' creative and critical thinking skills in higher education". She coordinated other previous international and national funded projects, the latest being an ERASMUS+ project funded by the European Comission: "BeyondScale: Developing the Organisational Capacity of Higher Education Institutions using the HEInnovate platform to facilitate peer learning and a pan-European community of practice". As a researcher she participates in several other funded projects, nationally and internationally. She represents IPVC in PRME, (Principles for Responsible Education) for United Nations and she's an ambassador of Alliance SDS Portugal - SDS 4 - Quality Education through the Global Compact Network Portugal. Previously to academia she was Human Resources Director in the industrial sector, entrepreneur and HR consultant in big companies and municipalities.

Preface to "Assessment and Evaluation in Higher Education"

This book compiles relevant research on students' learning processes in Higher Education. It presents different dimensions and forms of learning, crucial to understand and tackle diversity and integration regarding more sustainable and inclusive goals.

Major challenges have taken place recently considering assessment within higher education and its known impact on students' learning processes. The application of different assessment methods, considering a general public desire to focus on more skills development and assessment has brought higher education challenges regarding students learning processes and its assessment that are bringing severe changes towards academia. Higher education institutions should act as knowledge and research drivers, and the consideration about the learning processes and assessment methods are key.

Special attention is given to assessment methods and purposes, assessment rubrics and the assessment of learning outcomes (knowledge and skills). Research on the challenges, strengths, and opportunities of online and virtual assessment, as well as best practices and recommendations for assessment and evaluation in higher education, are also explored and discussed in this book. This book, which consists of 15 articles written by research experts in their topic of interest, reports the most recent research concerning assessment and evaluation in higher education. The book includes changes in education and higher education that can lead to a systematic modification of higher education.

Topic one—Pedagogical innovation

Jung Hee Park, Woo Sok Han, Jinkyung Kim and Hyunjung Lee—Strategies for Flipped Learning in the Health Professions Education in South Korea and Their Effects: A Systematic Review; https://doi.org/10.3390/educsci11010009.

William Swart and Ken MacLeod—Evaluating Learning Space Designs for Flipped and Collaborative Learning: A Transactional Distance Approach; https://doi.org/10.3390/educsci11060292.

Jiří Bečica and Roman Vavrek—A Qualitative Assessment of the Pedagogical Process at Czech Public Universities; https://doi.org/10.3390/educsci11080389.

Salleh Amat, Harizah Izyan Samsudin, Abu Yazid Abu Bakar, Mohd Izwan Mahmud and Mohd Hasrul Kamarulzaman—Needs Analysis of Psychosocial Module Development Based on Psychoeducation Approach for Public University Students in Malaysia; https://doi.org/10.3390/educsci11100604.

Humberto Arruda and Édison Renato Silva—Assessment and Evaluation in Active Learning Implementations: Introducing the Engineering Education Active Learning Maturity Model; https://doi.org/10.3390/educsci11110690.

We begin the book by introducing the reader to a very interesting work in which Park et al. develops a student-centred approach considering the real advantages in developing students learning abilities. The effect and strategies of flipped learning in the health professions education

were assessed through a Participant–intervention-comparator-outcome (PICO). Findings suggests that implementing flipped learning design based on class characteristics and appropriate post-class activities can enhance students' learning abilities.

The topic continues with the work from Swart et al. where they evaluate Learning Space Designs for Flipped and Collaborative Learning. A highly revealing study considering the relevance of the classroom´s design as a relevant mainframe for collaborative work. They developed a very interesting study where they present that from a student perspective the investment in high-tech classrooms is not necessarily directly associated with more student engagement.

Amat and colleagues presents a research which showed a highly positive response from the students toward the development of a psychosocial module in order to strengthen the assertiveness, social skills, and psychological well-being.

Assessment and Evaluation in Active Learning Implementations has been widely worked on engineering Education. Arruda and colleagues present a paper where they propose a framework that assesses the maturity of Active Learning implementation in a program or a course. A valuable tool to be implemented and replicated, on diagnosis and practical improvements in real settings.

Topic Two—Assessment

Fidel Salas Vicente, Ángel Vicente Escuder, Miguel Ángel Pérez Puig and Francisco Segovia López—Effect on Procrastination and Learning of Mistakes in the Design of the Formative and Summative Assessments: A Case Study; https://doi.org/10.3390/educsci11080428.

Eddy Sutadji, Herawati Susilo, Aji Prasetya Wibawa, Nidal A.M. Jabari and Syaiful Nur Rohmad—Authentic Assessment Implementation in Natural and Social Science; https://doi.org/10.3390/educsci11090534

Jurgita Lenkauskaitė, Remigijus Bubnys, Erika Masiliauskienė and Daiva Malinauskienė—Participation in the Assessment Processes in Problem-Based Learning: Experiences of the Students of Social Sciences in Lithuania; https://doi.org/10.3390/educsci11110678.

Tatiana Baranova, Dmitriy Mokhorov, Aleksandra Kobicheva and Elena Tokareva—Assessment of the Application of Content and Language Integrated Learning in a Multilingual Classroom; https://doi.org/10.3390/educsci11120808.

Another interesting research is the one presented by Vicent et al. where we are guided through students' procrastination processes, the formative assessments and the summative evaluation. This case study presents the unexpected and undesirable effects of both on the behavior of students and on both grading and learning.

Authentic assessment is another pedagogical approach based on real world impact. Sutadji and colleagues presented a research paper where they debate epistemological and scientific constructions to social and natural sciences.

Lenkauskaitė and colleagues explored the idea of change in the higher educational process using a problem-based learning strategy. The need we have in education institutions to empower students is reflected in this study that shows empowerment and enhancement of critical thinking from students when they assess and are involved in the assessment strategy.

To end this topic Baranova and colleagues present a research based on the introduction of content and language into classes with a multilingual approach, which will allow students to use several foreign languages in the process of professional communication. The purpose of the paper is to evaluate the efficiency of a newly introduced integrated learning model from the perspective of students and to understand its impact on students.

Topic three—Digitalization- and peer-support-technology-based

Rana Saeed Al-Maroof, Khadija Alhumaid and Said Salloum—The Continuous Intention to Use E-Learning, from Two Different Perspectives; https://doi.org/10.3390/educsci 11010006.

Isabel López-Tocón—Moodle Quizzes as a Continuous Assessment in Higher Education: An Exploratory Approach in Physical Chemistry; https://doi.org/10.3390/educsci11090500.

Mehdi Berriri, Sofiane Djema, Gaëtan Rey and Christel Dartigues-Pallez—Multi-Class Assessment Based on Random Forests.; https://doi.org/10.3390/educsci11030092.

In Topic 3, the book presents relevant pos pandemic strategies like e-learning methods. The paper by Al-Maroof et al. show that teachers' and students' perceived technology self-efficacy, ease of use and usefulness are the main factors directly affecting the continuous intention to use technology.

Their relevant findings also bring about intrinsic and extrinsic motivation associated and developed throughout the process of learning. Considered the key factors that support the use of technology continuously.

The chapter continues with López-Tocón and colleagues work on Moodle Quizzes as a Continuous Assessment. Moodle quizzes were explored in this study, and they acted as a reliable practice for teaching and learning.

This topic regards crucial aspects that intend to enhance the need for higher education institutions to support additional student-centred dynamics not based on the teacher and its role as the main actor in the classroom or even as main facilitator of knowledge. In this chapter it is explored by the work of Berriri, a novel study that tried to provide relevant information regarding counseling processes. Additionally, it is explored how teaching staff can propose training courses adapted to students by anticipating their possible difficulties using new technologies, a machine learning algorithm called Random Forest, allowing for the classification of the students depending on their results.

Topic four—Organizations and governance

Cristina Checa-Morales, Carmen De-Pablos-Heredero, Angela Lorena Carreño, Sajid Haider and Antón García—Organizational Differences among Universities in Three Socioeconomic Contexts: Finland, Spain and Ecuador. Relational Coordination Approach; https://doi.org/10.3390/educsci11080445.

Hue Thi Truong, Hung Manh Le, Duc Anh Do, Duc Anh Le, Huyen Thi Nguyen and Thanh Kim Nguyen—Impact of Governance Factors over Lecturers' Scientific Research Output: An Empirical Evidence; https://doi.org/10.3390/educsci11090553.

Cecília Guerra and Nilza Costa—Can Pedagogical Innovations Be Sustainable? One Evaluation Outlook for Research Developed in Portuguese Higher Education; https://doi.org/10.3390/educsci11110725.

The book also presents institutional research, conducted in a macro level approach that can be very helpful to the reader that is interested in more information regarding the assessment of pedagogical innovation at an institutional level. Bečica and colleagues present a relevant qualitative Assessment of the Pedagogical Process at Czech Public Universities. This work increments the debate towards more quality-based measures considering that the monitoring and quantification of quality in education is a very demanding and controverse topic.

Checa-Morales and colleagues presented an institutional study that focus on organizational differences Universities in Three Socioeconomic Contexts: Finland, Spain and Ecuador. Results show that shared goals with among faculty members and problem-solving communication among students were key. Organizational practices show relevant differences among the three universities.

Truong and colleagues analysed governance factors that influence the scientific research output of lecturers. The results reveal that resources for scientific research have the most impact on lecturers' scientific research output, followed by policies for lecturers, support for scientific research activities, scientific research objectives of HEIs, and finally, leadership.

Pedagogical innovation is key in order to develop society. Guerra and colleagues introduce a very interesting and critical approach to the topic, conducting research that focused on the extent to which funded national research-based education projects, developed in public Portuguese higher education institutions (universities and polytechnic institutes), have considered the sustainability of research results (e.g., pedagogical innovations), after funding ends.

We believe there is a need for this book to provide clear and relevant scientific research which takes into consideration pedagogical innovation, assessment processes and institutional models. The aim of this book is to provide the reader with key and updated research on the information necessary to understand students learning and teaching, innovation at assessment, the need to assess involving the students and also to understand global macro level research, national level and institutional level. Therefore, we hope to reach policy makers, educators, researchers, teachers, students and a global community of people that are interested and concerned with the research agenda, acting in student's empowerment, knowledge and updated skills.

Sandra Raquel Gonçalves Fernandes, Marta Abelha, and Ana Teresa Ferreira-Oliveira
Editors

education sciences

MDPI

Review

Strategies for Flipped Learning in the Health Professions Education in South Korea and Their Effects: A Systematic Review

Jung Hee Park [1,†], Woo Sok Han [2,†], Jinkyung Kim [2,*,‡] and Hyunjung Lee [3,*,‡]

[1] Department of Emergency Medical Service, Konyang University, Daejeon 158, Korea; jhpug@konyang.ac.kr
[2] Department of Hospital Management, Konyang University, Daejeon 158, Korea; wshan@konyang.ac.kr
[3] College of Nursing, Konyang University, Daejeon 158, Korea
* Correspondence: jkim@konyang.ac.kr (J.K.); leehj18@konyang.ac.kr (H.L.); Tel.: +82-42-600-8422 (J.K.); +82-42-600-8584 (H.L.)
† Two authors contributed equally to this work as co-first author.
‡ These two authors equally contributed to the present work as the corresponding author.

Abstract: This study aims to identify and synthesize recent literature on the effect and strategies of flipped learning in the health professions education. Participant–intervention-comparator-outcome (PICO) strategies were used to identify articles from published peer-reviewed papers from January 2017 to March 2020 in Korea Med, Korean Citation Index, National Digital Science Library, and Korean Studies Information Service System. Of the 83 screened articles, 10 published articles met all the inclusion criteria. Most of articles targeted nursing students and focused on practicum classes. The effects of flipped learning were measured based on satisfaction, self-motivated learning, information literacy, and critical thinking disposition. Further, pre-class, in-class, and post-class activities were analyzed. The findings revealed that flipped learning improved class performance, overall evaluation, self-motivated learning, self-efficacy, and problem-solving abilities. The study suggests implementing a tailored flipped learning design based on class characteristics and appropriate post-class activities for enhancing students' learning abilities.

Keywords: health professions education; flipped learning; learning method

Citation: Park, J.H.; Han, W.S.; Kim, J.; Lee, H. Strategies for Flipped Learning in the Health Professions Education in South Korea and Their Effects: A Systematic Review. *Educ. Sci.* **2021**, *11*, 9. https://doi.org/10.3390/educsci11010009

Received: 1 December 2020
Accepted: 29 December 2020
Published: 31 December 2020

1. Introduction

The European Union Joint Research Center argues that traditional lecture-based education is severely limited and incapable of adequately embodying the key visions of future education [1]. The traditional classroom, focusing on knowledge delivery via rote learning, has seen a shift toward a more learner-centered classroom with the introduction of various teaching and learning methods. Flipped learning is one such methodology that is focused on interaction with students as opposed to a unidirectional monologue; it is also known as reverse learning. Flipped learning, or a student-oriented learning method, consists of pre-class, in-class, and post-class activities. Pre-class activities guide students to actively learn the lesson beforehand using multimedia materials; in-class activities involve instructor-learner interaction based on knowledge acquired via pre-class activities. Post-class activities help reinforce and expand on the acquired learning [2].

Flipped learning promotes instructor-student interaction in contrast to unidirectional teaching. With the introduction and popularization of various educational media that help increase learners' understanding, the demand for high-quality lectures is also on the rise [3]. Many colleges use learner-centered classes to strengthen students' competence, in which the class contents are explored in depth [4]. There is also active ongoing research on flipped learning [5,6]. Since students learn the lesson in advance using media that can be accessed anywhere and at any time, their understanding of the lesson is improved during the actual class, which leads to greater satisfaction in class [7]. However, to maximize the

effectiveness of flipped learning, few conditions need to be met, such as active student participation, appropriate instructor intervention, and regulation of pre-learning. Courses must be designed specifically for each subject based on an analysis of numerous cases prior to applying flipped learning.

Learner-centered instruction is also being implemented currently in undergraduate health professions education to foster the competencies of health professionals [8]. Studies on learner-centered instruction, including flipped learning, have reported that such learning methods are effective in improving students' self-directed learning abilities, problem-solving, interpersonal skills, clinical performance, critical thinking, and academic achievements, thereby increasing academic performance [9–12]. A recent systematic literature review on flipped learning in South Korean nursing education found improvements in students' learning capability [5]. Hew and Lo conducted a meta-analysis on the flipped classroom approach in education for the health professions using journals in the Journal Citation Report [7]. However, no systematic literature review has been published that targeted health professions education in South Korea. It is necessary to conduct an in-depth analysis of learner-centered instruction cases in health and allied health majors to foster prospective health professionals who will be in charge of public health. Further, as flipped learning was introduced in colleges in Korea only in 2013, efforts to stabilize and enhance its efficiency of operation and develop instruments for quality control have been inadequate [13]. Therefore, this study proposes directions for research on flipped learning in health professions education and for an effective application of flipped learning in health professions education by systematically reviewing the latest pertinent studies. The specific objectives of this study are to analyze the general characteristics, the measured variables and outcomes, and the strategies used in studies on flipped learning in the health professions education.

2. Materials and Methods

2.1. Study Design

This paper is a systematic review of studies analyzing the effects of flipped learning in undergraduate health majors in Korea to identify the latest research trend and assess the teaching strategies and their effectiveness. The literature search strategy was based on the participant-intervention-comparators-outcome (PICO) framework [14].

2.2. Literature Search

The keywords for the participants (P) of the study were undergraduate allied health majors (excluding medicine, dentistry, Korean medicine, pharmacy). The keyword for the intervention (I) was flipped learning. The comparator (C) referred to conventional lecture-based classes. The keyword of the outcomes (O) referred to the major outcomes of flipped learning. To examine the latest trend in flipped learning research, the search was limited to Korean studies published between January 2017 and March 2020. The search was performed on Korea Med, National Digital Science Library (NDSL), Korean Studies Information Service System (KISS), and the Korean Citation Index (KCI). The database used for the search was based on the report of Kim et al. [15] and included the search engine of the Core of the Core, standard ideal (COSI) model, which is the protocol for literature review databases in Korea. Four researchers performed the search independently, and studies were selected after reviewing them against the inclusion and exclusion criteria. The search strategy generated a total of 2155 studies with 20 studies from Korea Med, 549 studies from KISS, 784 studies from NDSL, and 802 studies from KCI. After excluding 1539 studies on medicine and non-health majors, 533 duplicate searches, and 62 studies that meet the exclusion criteria (but were qualitative studies, literature reviews, or published before 2017), the manuscripts of 21 studies were selected and subsequently acquired.

The quality of the selected studies was appraised independently by four researchers using the methodology checklist for cohort studies published by the Scottish Intercollegiate Guidelines Network (SIGN) [16]. Studies that compared the outcomes of flipped learning

and conventional lecture-based classes were included, and after re-discussing studies on which the researchers had divided opinions regarding the comparison groups, a total of 10 studies were finally selected for the analysis (see Figure 1).

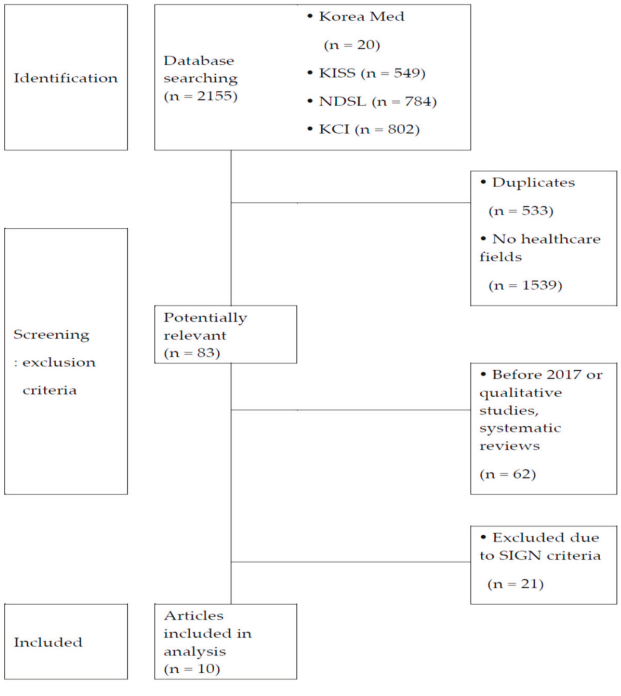

Figure 1. Precedent studies' selection flowchart.

2.3. Literature Analysis

All the 10 selected studies were quantitative studies. The features of each study were organized in a table by four researchers, and the assessments were reviewed among them. The following features of the studies were analyzed: author and publication year, institutional review board (IRB) review, course subject, pedagogical features (strategy, intervention period), participant features (sample size, inclusion criteria), measured variables (instruments, reliability and validity), and outcome variables (Table 1).

Table 1. Overview of studies included in the analysis.

No.	Author (Year)	IRB	Course	Pedagogical Feature		Participant Feature		Measured Variables		Results
				Strategy	Period	Sample Size	Inclusion Criteria	Instruments	Reliability Validity	
1	Kim and Park (2019) [17]	+	Comprehensive practicum II	Jigsaw model, flipped learning, cooperative learning	5 weeks	81	+	Nursing knowledge, critical thinking, self-directed learning, learning satisfaction	+	Improved learning satisfaction
2	Park et al. (2019) [18]	+	Geriatric nursing	Team-based learning	4 weeks	54	-	Class experience, level of communication anxiety, learning satisfaction, intent to continue learning	+	High class experience, learning satisfaction, intent to continue learning
3	Jung and Yang (2019) [19]	+	Fundamentals of nursing practicum	Create video, self-evaluation journal	11 weeks	74	+	Information literacy, critical thinking	+	Improved information literacy and critical thinking
4	H. Kim and E. Kim (2018) [20]	+	Health assessment	Video- and scenario-based learning, self-evaluation journal	10 weeks	178	+	Knowledge, performance, comprehensive knowledge	+	Improved academic achievement
5	Lee and Park (2018) [21]	+	Surgical nursing practicum	e-learning	1 week	102	-	Clinical self-efficacy, self-directed learning, problem-solving skills	+	Improved goal setting, self-directed learning, and problem-solving skills
6	Kim (2017) [22]	-	Health education	Team-based learning	7 weeks	48	-	Academic achievement, course satisfaction, learning attitude, health education self-efficacy	+	High academic achievement, learning attitude, health education self-efficacy

Table 1. *Cont.*

No.	Author (Year)	IRB	Course	Pedagogical Feature		Participant Feature		Measured Variables		Results
				Strategy	Period	Sample Size	Inclusion Criteria	Instruments	Reliability Validity	
7	Park and Woo (2017) [23]	+	Fundamentals of nursing	Question formation by students, team-based learning, MOOCs (Massive Open Online Courses)	5 weeks	102	+	Self-directed learning, problem-solving skills	+	Student question-centered teaching method is effective for self-directed learning and problem-solving skills
8	Lee and Han (2017) [24]	-	Health education and methodology	Cooperative learning	6 weeks	49	-	Learning motivation, class participation, course satisfaction	+	Learning motivation and class participation have positive impact on course satisfaction
9	Lee et al. (2017) [25]	+	Psychiatric nursing	Develop video module Write WSQ (Watch-Summarize-Question) journal	2 weeks	146	+	Undergraduate key competencies, academic performance, satisfaction with teaching method, and perceived usefulness	+	Improved academic performance and undergraduate key competencies, highly rated satisfaction with teaching method and perceived usefulness
10	You and Kim (2017) [26]	+	Fundamentals of nursing	Create video, utilize the time for self-training, self-evaluation journal	15 weeks	78	+	Self-directed learning, information literacy	+	High self-directed learning, information literacy

3. Results

3.1. General Features of the Selected Studies

The 10 selected studies were published between 2017 and 2020. The study design used to examine the effects of flipped learning was either pre-test–post-test or quasi-experimental design, and all studies were conducted on nursing students. The mean intervention period was six weeks and six days, and the mean number of participants in the flipped learning group was 91. Six studies presented evidence for determining the sample size for flipped learning (study no. 1, 3, 4, 7, 9, and 10). Eight studies were either approved by an IRB or mentioned obtaining consent from students undergoing flipped learning. Although flipped learning was generally applied to practicum courses, it was also sometimes used in theoretical courses such as fundamentals of nursing, health assessment, health education and methodology, and geriatric nursing.

3.2. Variables for Flipped Learning Outcomes and Results

The most common variable used to measure the outcomes of flipped learning was satisfaction, namely, satisfaction with learning, the course, or the teaching method (n = 5). This was followed by self-directed learning ability (n = 4), knowledge (n = 3), and critical thinking, problem-solving skills, information literacy, and academic achievement (n = 2 each). Other variables included class experience, level of communication anxiety, intent to continue learning, clinical performance, self-efficacy, learning accomplishment, learning attitude, learning motivation, class participation, undergraduate competencies, academic performance, and usefulness. The studies generally reported that flipped learning had a positive impact on the measured variables (Table 1).

3.3. Flipped Learning Teaching Strategies

To compare the teaching strategies used in studies that examined the effects of flipped learning, the activities were classified into pre-class, in-class, and post-class activities for analysis. Pre-class activities included team-based learning, video-based learning, e-learning courses, individual study, student question formation, watch-summary-question (WSQ) journal, skill practicing, and self-evaluation journal. In-class activities included team-based discussion and cooperative learning, team-based presentation, instructor feedback, hands-on training, pre-learning readiness assessment, scenario-based application, quiz on the e-learning materials, nursing diagnosis training, discussion on student-formed questions, instructor-led training (discussion, simulation, case study, conference), and 1:1 evaluation and guidance by the instructor. Post-class activities included review of key contents, peer evaluation, post-class evaluation and survey, Q&A, post-class self-study, self-evaluation journal, instructor feedback, application of case study, and team-based answer review and sharing (Table 2).

Table 2. Major features of flipped learning strategies.

No.	Author (Year)	Courses	Pre-Class	In-Class	Post-Class
1	Kim and Park (2019) [17]	Comprehensive practicum II	Assign team-based learning tasks	Team-based discussion and cooperative learning	Review key contents, peer review, post-class evaluation and survey
2	Park et al. (2019) [18]	Geriatric nursing	Team-based learning	Team-based presentation and discussion	Q&A, peer review
3	Jung and Yang (2019) [19]	Fundamentals of nursing practicum	Video-based learning	Team-based discussion, instructor feedback, hands-on training	Post-class self-training, self-evaluation journal
4	H. Kim and E. Kim (2018) [20]	Health assessment	Video-based learning	Evaluation of pre-learning readiness, scenario-based learning, hands-on training	Self-evaluation journal, instructor feedback
5	Lee and Park (2018) [21]	Surgical nursing practicum	e-learning	Quiz on the e-learning materials, nursing diagnosis training	Apply to case study, instructor feedback
6	Kim (2017) [22]	Health education	Self-study	Team-based discussion and presentation	Instructor feedback
7	Park and Woo (2017) [23]	Fundamentals of nursing	Video-based learning, question formation by student	Evaluation of pre-learning readiness, discussion of student-formed questions	Team-based answer checking and sharing
8	Lee and Han (2017) [24]	Health education and methodology	Team-based learning	Team-based discussion and presentation, instructor feedback	None
9	Lee et al. (2017) [25]	Psychiatric nursing	Video-based learning, write WSQ (Watch-Summarize-Question) journal	Instructor-guided training (e.g., discussion, simulation, case study, conference)	None
10	You and Kim (2017) [26]	Fundamentals of nursing	Video-based learning, Self-skills training, self-evaluation journal	Check on pre-class activities, 1:1 evaluation, and guidance by instructor	Self-training

4. Discussion

This study systematically reviewed Korean studies on flipped learning in allied health majors published within the past three years and based on the results, identified the latest trends in flipped learning research. Although we attempted to include studies on flipped learning in various health majors, all 10 selected studies were conducted on nursing students. As per the SIGN checklist for cohort studies, only studies that extracted both a control group and experimental group from the source population were selected.

This led to the exclusion of a number of studies that did not include a comparator (C). These excluded studies were conducted on students majoring in radiology, emergency medicine, occupational therapy (n = 2 each), and physical therapy, dental hygiene, and dental technology (n = 1 each). This suggests that flipped learning is less frequently applied in other health majors compared to nursing, necessitating an expanded application of flipped learning in more health professions courses. A pre-post analysis with a single group is limited in comparing the effects of flipped learning with that of the traditional

classroom approach. Thus, study designs should be chosen carefully to clearly assess the effects of flipped learning.

The courses in which flipped learning was applied were generally practicum courses, including comprehensive practicum II, geriatric nursing, fundamentals of nursing practicum, health assessment, surgical nursing practicum, health education, fundamentals of nursing, health education and methodology, and psychiatric nursing practicum, which was consistent with other studies that analyzed flipped learning methods in health professions education [27–31]. The educational effects of flipped learning approach in practicum courses included improved student attitude, fewer errors in practice [5], and improved student perception, self-directed learning, academic achievement, and satisfaction with the class [27,29,31].

In a meta-analysis of the effects of learner-centered class in nursing, Lee and Yang [9] reported that a learner-centered class is effective in enhancing clinical performance and learning of major-related knowledge; they demonstrated that learner-centered approaches, such as flipped learning, are more effective than the traditional lectures in the study of nursing. Prior to 2017, common measured variables were self-directed learning [27], academic achievement [27], self-efficacy [32,33], critical thinking and communication [32], and learning motivation [33]. On the other hand, in the past three years, a number of variables, such as knowledge and performance, self-directed learning [34,35], critical thinking [34,36,37], academic self-efficacy [36,37], satisfaction with major [36], and communication and problem-solving skills [36]. As flipped learning enables students to repeatedly learn the contents related to their lesson of the day through pre-class learning without restrictions of time and place, students are able to adequately familiarize themselves with the lesson in advance, which increases their understanding of the lesson during the actual class. Many studies on flipped learning confirmed positive changes after applying flipped learning, and they generally used variables that influence students' tendencies or academic competence and performance, that is, variables positively correlated with the flipped learning approach [4,5,7,9]. As shown in previous studies, the results of the literature review section in this study confirmed that flipped learning improved satisfaction with class (n = 4), academic achievement (n = 3), self-directed learning, problem-solving skills, and information literacy (n = 2), and critical thinking. Future studies should discover new significant variables to examine the effects of flipped learning. To expand the utilization of the flipped learning approach, active research and effort are needed to verify its effects.

Flipped learning strategies were analyzed by dividing the class into pre-class, in-class, and post-class activities. Video lectures were primarily used in pre-class activities for practicum courses, while team-based learning or individual study were performed in pre-class activities for theoretical courses. In-class strategies included pre-class readiness assessment, instructor feedback, and hands-on training during practicum courses, while team-based discussions or presentation were performed in theoretical education. Popular post-class strategies included peer evaluation, self-evaluation journal, and instructor feedback, and post-class activities were not used in two cases. Gan et al. [13] analyzed flipped learning cases and proposed a basic operational model for theory and practicum courses, and a similar flipped learning model was used in the ten studies reviewed in this article. Lee and Chang [25] had students write a WSQ journal, which required them to watch, summarize, and question activities as a form of pre-class activity. WSQ journal writing was proposed by Kirch [38], which involves instructor-student interaction in finding the solution during in-class activities and helps students take responsibility for completing the pre-class portion. Lee and Chang [25] reported that the WSQ strategy has a positive influence on self-directed learning and academic achievement. Friedman and Friedman [39] proposed the Analysis, Design, Development, Implementation, and Evaluation (ADDIE) model. Lee [40] applied the ADDIE model to flipped learning and proposed pre-learning, in-class learning activities, and evaluation and self-evaluation after class as the major teaching strategies. Gan et al. [13] argued that the post-class part of learning is the stage in which the learned materials are maintained and expanded. Among the studies analyzed

in this review, two studies did not use post-class activities, but appropriate post-class activities, such as reviewing, summarizing, complementing, and writing self-evaluation journals, must be implemented to facilitate continuity of learning. The results of previous studies and this study highlight the need to use flipped learning designs and strategies tailored for each subject to boost their learning effects on students.

5. Conclusions

The purpose of the study is to identify and synthesize recent literature on the effects and strategies of flipped learning in education for the health professions in South Korea. We used participant–intervention-comparators-outcome (PICO) strategies to identify the published peer-reviewed articles from January 2017 to March 2020 in Korea Med, KCI, NDSL, and KISS. Ten published articles from 83 screened articles met the inclusion criteria. All articles targeted nursing students and mostly focused on practicum classes. The effects of flipped learning were measured by satisfaction, self-motivated learning, information literacy, and critical thinking disposition. We further analyzed pre-class, in-class, and post-class activities. The limitation of this study is that it was not possible to evaluate the effect of different types of flipped learning approaches under the limited number of analyzed studies. Despite these limitation, we found that the flipped learning model improved class performance, overall evaluation, self-motivated learning, self-efficacy, and problem-solving abilities. The study results suggest implementing a tailored flipped learning design based on class characteristics and emphasizing appropriate post-class activities to enhance students' learning abilities.

Author Contributions: Conceptualization, H.L. and J.K.; methodology, H.L. and J.K.; formal analysis, J.H.P. and W.S.H.; investigation, J.H.P. and W.S.H.; data curation, J.H.P. and W.S.H.; writing—original draft preparation, J.H.P. and W.S.H.; writing—review and editing, H.L. and J.K.; visualization, J.H.P. and W.S.H.; supervision, H.L. and J.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

1. Redecker, C.; Leis, M.; Leendertse, M.; Punie, Y.; Gijsbers, G.; Kirschners, P.; Stoyanov, S.; Hoogveld, B. *The Future of Learning: Preparing for Change*; Institute for Prospective Technological Studies, European Commission Joint Research Center: Luxembourg, 2011.
2. Goodwin, B.; Miller, K. Evidence on flipped classrooms is still coming in. *Educ. Leadersh.* **2013**, *70*, 78–80.
3. Shim, J.G.; Kim, Y.M.; Park, S.J. Effect of smart learning applied on achievement goal, self-directed learning for students in health college. *Korean J. Radiol.* **2017**, *39*, 661–667. [CrossRef]
4. Estrada, A.C.M.; Vera, J.G.; Ruiz, G.R.; Arrebola, I.A. Flipped Classroom to improve university student centered learning and academic performance. *Soc. Sci.* **2019**, *8*, 315. [CrossRef]
5. Bae, S.H.; Shin, S.J. The effect and strategies of flipped learning in nursing education: A systematic review. *Health Nurs.* **2018**, *30*, 1–9. [CrossRef]
6. Shin, J.H. Analyzing the level of concerns about learner-centered pedagogical method of university faculty. *J. Lifelong Learn. Soc.* **2014**, *10*, 109–132. [CrossRef]
7. Hew, K.F.; Lo, C.K. Flipped classroom improves student learning in health professions education: A meta-analysis. *BMC Med. Educ.* **2018**, *18*, 38–50. [CrossRef]
8. Rezaee, R.; Moadeb, N.; Shokrpour, N. Team-based learning: A new approach toward improving education. *Acta Med. Iran.* **2016**, *54*, 679–683.
9. Lee, J.E.; Yang, S.H. The effects of learner-centered instruction in nursing: A meta-analysis. *Asian J. Educ.* **2018**, *19*, 1049–1077. [CrossRef]

10. Kim, E.J. A meta-analysis of team-based learning effects in university classes. *J. Learn. Cent. Curr. Instr.* **2017**, *17*, 721–743. [CrossRef]
11. Je, M.S.; Nam, C.W. Developing a teaching and learning model using flipped learning for the course of core fundamental nursing skills in nursing education. *J. Learn. Cent. Curr. Instr.* **2017**, *17*, 375–401. [CrossRef]
12. Woo, C.H.; Park, J.Y. The preference of instructional methods and Kolb's learning styles of nursing students. *J. Digit. Converg.* **2016**, *14*, 339–348. [CrossRef]
13. Gan, J.S.; Han, S.B.; Seo, Y.J.; Lee, E.J.; Bae, S.M.; Shin, M.S.; Kim, Y.H. *Case Analysis for Flipped Learning Course Operation and Development of Operational Model*; The 9th Research Project Contest; Korean Association of Center for Teaching and Learning: Bucheon, Korea, 2017; pp. 3–52.
14. Yensen, J. PICO search strategies. *Online J. Nurs. Inform.* **2013**, *17*. Available online: http://ojni.org/issues/?p=2860 (accessed on 10 June 2020).
15. Kim, S.Y.; Park, J.E.; Seo, H.J.; Lee, Y.J.; Son, H.J.; Jang, B.H.; Seo, H.S.; Shin, C.M. NECA's guidance for undertaking systematic reviews and meta-analyses for intervention. *Natl. Evid. Based Healthc. Collab. Agency* **2011**, *25*, 1–19.
16. Harbour, R.; Miller, J. A new system for grading recommendations in evidence-based guidelines. *BMJ* **2001**, *323*, 334–336. [CrossRef]
17. Kim, H.J.; Park, D. Effects of convergence education by jigsaw model and flipped learning in nursing students. *J. Converg. Inform. Technol.* **2019**, *9*, 36–43.
18. Park, B.N.; Shin, M.K.; Jeon, H.J. Analysis about the effect of flipped learning based team activity. *J. Converg. Inform. Technol.* **2019**, *9*, 44–51.
19. Jung, H.G.; Yang, Y.K. Effects of flipped learning-based fundamentals of nursing practice. *J. Learn. Cent. Curr. Instr.* **2019**, *19*, 555–572. [CrossRef]
20. Kim, H.; Kim, E. The effects of flipped learning based health assessment on academic achievement of nursing students. *J. Korea Acad. Ind. Coop. Soc.* **2018**, *19*, 201–210.
21. Lee, M.K.; Park, B.K. Effects of flipped learning using online materials in a surgical nursing practicum: A Pilot Stratified group-randomized trial. *Health Inform. Res.* **2018**, *24*, 69–78. [CrossRef]
22. Kim, H. A case study on flipped learning of health education class-comparison with the lesson that applied the reflective journal method. *J. Educ. Cult.* **2017**, *23*, 135–148. [CrossRef]
23. Park, J.Y.; Woo, C.H. The effects of lesson with student-generated questions: Based on flipped learning utilizing massive open online courses. *J. Learn. Cent. Curr. Instr.* **2017**, *17*, 283–306. [CrossRef]
24. Lee, N.Y.; Han, J.Y. Nursing students' learning motivation, class participation, and class satisfaction on flipped class and instructor-centered class. *J. Learn. Cent. Curr. Instr.* **2017**, *17*, 253–267. [CrossRef]
25. Lee, M.K.; Chang, S.J.; Jang, S.J. Effects of the flipped classroom approach on the psychiatric nursing practicum course. *J. Korean Acad. Psychiatr. Ment. Health Nurs.* **2017**, *26*, 196–203. [CrossRef]
26. You, H.S.; Kim, N.Y. The effects of flipped learning on the self-directed learning and information literacy of nursing students. *J. Learn. Cent. Curr. Instr.* **2017**, *17*, 491–510. [CrossRef]
27. Lee, Y.A.; Kim, H.S. Effect of flipped learning on self-directed learning and academic achievement in emergency medical technology students. *Clin. Exp. Emerg. Med.* **2016**, *20*, 99–111. [CrossRef]
28. Park, S.H.; Lee, M.Y. A study on the satisfaction of flipped learning with pre-class using video in basic medical subjects. *J. Korea Entertain. Ind. Assoc.* **2017**, *11*, 261–269. [CrossRef]
29. Chung, E.J. The effect of flipped learning on self-directed learning and class satisfaction in a class of college physical therapy students. *J. Korean Soc. Integr. Med.* **2018**, *6*, 63–73.
30. Lee, N.H.; Jung, H.R. The effects of flipped learning in sensory integration therapy class using online learning platform on learning participation. *J. Korean Soc. Integr. Med.* **2018**, *12*, 247–256. [CrossRef]
31. Choi, M.S. A case study on flipped learning convergence in dental hygiene major: Focusing on learning awareness and academic achievement. *J. Converg. Inform. Technol.* **2019**, *9*, 252–263.
32. Lee, Y.S.; Eun, Y. The effect of the flipped learning on self-efficacy, critical thinking disposition, and communication competence of nursing students. *J. Korea Acad. Soc. Nurs. Educ.* **2016**, *22*, 567–576. [CrossRef]
33. Gu, H.J. Effectiveness of flipped learning in fundamental nursing practice education. *J. Korean Stat. Soc.* **2016**, *18*, 2829–2841.
34. Kim, H.; Kim, Y.H. An action research on flipped learning for fundamental nursing practice courses. *J. Korean Acad. Fundam. Nurs.* **2017**, *24*, 265–276. [CrossRef]
35. Choi, Y.S.; Kim, J.Y. Effects of flipped learning in core competencies of nursing students: Based on communication, problem-solving, and self-directed learning. *J. Learn. Cent. Curr. Instr.* **2018**, *18*, 1163–1182. [CrossRef]
36. Jung, M.R.; Jeong, E. Effects of e-book-based flipped learning education on critical thinking disposition, academic self-efficacy, and major satisfaction of nursing students. *J. Korea Contents Assoc.* **2018**, *18*, 490–501.
37. Lee, O.S.; Noh, Y.G. The effect of flipped learning on learning motivation, academic self-efficacy, and critical thinking disposition of nursing students. *J. Digit. Converg.* **2019**, *17*, 253–260.
38. Kirch, C. *Flipping with Kirch: The Ups and Downs from Inside My Flipped Classroom*; Bretzmann Group: New Berlin, WI, USA, 2016.
39. Friedman, H.H.; Friedman, L.W. Crises in education: Online learning as a solution. *Creat. Educ.* **2011**, *2*, 156–163. [CrossRef]
40. Lee, D.Y. Research on developing instructional design models for flipped learning. *J. Digit. Converg.* **2013**, *11*, 83–89. [CrossRef]

education sciences

MDPI

Article

The Continuous Intention to Use E-Learning, from Two Different Perspectives

Rana Saeed Al-Maroof [1], Khadija Alhumaid [2] and Said Salloum [3,*]

[1] English Language & Linguistics, Al-Buraimi University College, Al-Buraimi 512, Oman; rana@buc.edu.om
[2] College of Education, Zayed University, Abu Dhabi, UAE; Khadija.Alhumaid@zu.ac.ae
[3] Research Institute of Sciences and Engineering (RISE), University of Sharjah, Sharjah, UAE
* Correspondence: ssalloum@sharjah.ac.ae

Abstract: During the recent vast growth of digitalization, e-learning methods have become the most influential phenomenon at higher educational institutions. E-learning adoption has proved able to shift educational circumstances from the traditional face-to-face teaching environment to a flexible and sharable type of education. An online survey was conducted, consisting of 30 teachers and 342 students in one of the universities in the United Arab Emirates. The results show that teachers' and students' perceived technology self-efficacy (TSE), ease of use (PEOU), and usefulness (PU) are the main factors directly affecting the continuous intention to use technology. Instructors' technological pedagogical content knowledge (TPACK) and perceived organizational support (POS) positively affect the intention to use the technology, whereas students' controlled motivation (CTRLM) has a greater influence on their intention to use the technology, due to the type of intrinsic and extrinsic motivation that they have and which they can develop throughout the process of learning. The findings support the given hypotheses. In addition, they provide empirical evidence of a relationship between perceived organizational support and perceived pedagogical content knowledge. In fact, they are considered the key factors that support the use of technology continuously.

Keywords: e-learning platform; PACK; perceived usefulness; perceived ease of use; perceived organizational support and technology self-efficacy

Citation: Saeed Al-Maroof, R.; Alhumaid, K.; Salloum, S. The Continuous Intention to Use E-Learning, from Two Different Perspectives. *Educ. Sci.* **2021**, *11*, 6. https://dx.doi.org/10.3390/educsci11010006

Received: 29 November 2020
Accepted: 21 December 2020
Published: 25 December 2020

1. Introduction

Teachers and students may perceive the importance of e-learning differently. Teachers usually focus on the importance of training and support that may enhance the effective use of e-learning platforms, whereas the perceived usefulness and ease of use are influential factors from the students' perspective. The differences in their perspectives stem from the fact that their roles are different. Students usually receive the product through the e-learning system and can get all the different advantages that the system may offer; thus, they act as the consumers of the product [1]. On the other hand, teachers are the providers of the educational product, as they provide learners with the content and synthesize the given information in simple and concise language [1,2].

Past studies have focused on the importance of e-learning and its implementation all over the world; some have focused on the continuous use of e-learning [3–19] and some on the effect of either teachers' or students' attitudes towards e-learning [20–26]. In other words, no studies have put forward the implementation of two models that focus on how the perceived interactivity of education technology influences teachers' and students' perceptions and urges them to continue using the technology. This study assumes that some factors affecting teachers' intentions to use e-learning platforms continually are different from those affecting students' intentions to use the same platforms continuously. Therefore, this study proposes two different models that tackle both teachers' and students' continuous intentions to use technology. The two models will focus on a certain predictive power that has a more direct relationship with the teachers' and students' perceptions regarding

the continued use of the technology. For instance, one of the factors that contributes to the teachers' continuous use of the e-learning platform is the support they get from the university to enhance the use of the e-learning environment. On the other hand, one of the crucial factors for the students' perception is the controlled motivation that embraces certain intrinsic and extrinsic factors. It is worth mentioning that all previous studies have focused on students' perspectives on any technology-based technique. The acceptance of technology from a student's perspective has been dealt with in many papers, such as [27–29]. The fact that this study focuses on the effect of the same e-learning tool from a teacher's perspective separately is what sets this paper apart from other previous studies.

The objective of this study was to propose a theoretical framework that could be validated later through a proposed model that predicates the continuous intention to use e-learning among students at public universities in Dubai. There are numerous examples of literature related to technology acceptance [25,30–34] and continuous intention [35–38] that have been reviewed to identify the most common factors affecting the continuous intention to use the e-learning platform. The main concentration has been on theories that have been proven to have great predictive power in understanding users' perceptions and on theories that help to explain the importance of continuous use from two different perspectives. Hence, the Technological Pedagogical Content Knowledge (TPACK) was initiated by [39], Technology Acceptance Model (TAM) acceptance theory by [40], Social Cognitive Theory by [41], Perceived Organizational Support (POS) by [42], and Motivational Theory (MT) by [43,44]. The main factors that have been derived from these theories are perceived use and perceived usefulness, computer self-efficacy, controlled motivation, and so forth. The table below (Table 1) summarizes the studies that have tackled the continuous use of e-learning platforms.

Table 1. Most relevant studies of e-learning platforms in different sectors.

Authors/ Reference	Target Population	Objective/Goal	Models Adopted
[45]	Students	To explain the f-variables that affect continued use of m-learning.	TAM, Theory of Planned Behavior (TPB), and Expectation Confirmation Model (ECM).
[46]	Students	To examine students' continuous use of blended learning, with reference to behavioral attitudes, motivations, and barriers.	TAM, TPB and self-determination theory (SDT).
[47]	Students	To make a connection between learners' adoption and satisfaction with LMS in blended learning in relation to certain learners' personal characteristics in terms of continuous use of the e-learning environment.	TAM and satisfaction factor (SAT).
[48]	Instructors	To examine the influential factors which may contribute to instructors' satisfaction with LMS use in a blended learning atmosphere.	LMS, system and instructors' characteristics that are derived from well-established factors.
[49]	Students	To investigate students' behavior of continuance intentions to use the double reinforcement interactive e-portfolio learning system.	TAM and IS continuance post-acceptance model (IS-TAM).

Table 1. *Cont.*

Authors/ Reference	Target Population	Objective/Goal	Models Adopted
[50]	Learners	To investigate the basic determinants behind the continuous intention to use e-learning.	TAM and Negative Critical Incident (NCI).
[51]	People chosen randomly through a high-traffic website	To investigate the motivational factors that affect the synthesized model that is composed of a combination of TAM, ECM, COGM and SDM.	TAM, ECM and cognitive model (COGM).
[52]	Technology users	To investigate and predict the main reason behind users' intentions to continue using e-learning.	ECM, TAM, and, TPB.

As seen in the previous table showing the studies in the existing literature, much research has been conducted focusing on students and/or teachers within one proposed model. Nevertheless, searching for the predictive power behind both teachers' and students' intentions by proposing different variables is still neglected. To our knowledge, no research has examined the continued-use intention (CU) of teachers and instructors using e-learning platforms in higher education. Without knowledge of teachers' and students' CU, it is impossible to enhance e-learning in the Gulf area or to support its programs, systems, or administrative policies in terms of helping to sustain the e-learning platform.

2. Theoretical Framework and Hypotheses

The proposed framework has certain factors that can make the intention to use e-learning more measurable from two different perspectives. TPACK and POS are crucial elements that usually guide the teaching and learning environments from the teachers' perspectives. On the other hand, controlled motivation (CTRLM) is a factor that combines students' intrinsic and extrinsic motivations. Nevertheless, certain factors are equally important to both teachers and students, such as technology self-efficacy (TSE), perceived usefulness (PU), and perceived ease of use (PEOU).

2.1. Technological Pedagogical Content Knowledge (TPACK)

Teachers' knowledge cannot be tackled easily, as it is a complex concept that has many embedded elements [39]. The most important element is pedagogical content knowledge (PCK), which has been the domain of study for many researchers and practitioners. Its importance stems from the fact that it comprises both the content and the pedagogy that can explain how a particular topic is organized and how it is represented to the learners [53].

Since its emergence, TPACK has become a must since all teachers want to have a full understanding of the relationship between pedagogy and technology. TPACK refers to the type of technological pedagogical knowledge that teachers need to organize and present the intended teaching material effectively [39]. TPACK is one of the influential factors affecting teachers' perspectives. It refers to technological pedagogical content knowledge which includes: TCK (technological content knowledge), TPK (technological pedagogical knowledge), and PCK (pedagogical content knowledge) [54]. Self-assessment surveys and performance-based assessments are the basic instruments for evaluating TPACK [55,56].

The framework that comprises TPACK can be explained as combing different elements, such as content knowledge (CK), that highlight the teacher's knowledge of the subject matter. It includes knowledge of different types, such as knowledge of the theory, discipline, psychological aspects, historical aspects, and so forth [57]. The other two closely related elements are the pedagogy knowledge (PK), which is closely related to the teacher's

knowledge concerning methodology, process, and practice, and the pedagogical content knowledge, which is concerned with how teachers interpret and tailor the teaching material to suit certain pedagogical aims and purposes. This implies that the difference between the former element and the latter is the fact that the latter is related to methodology, assessment, and teaching style knowledge which can be used differently based on students' prior knowledge. Other elements are related to technology, as it comprises technology knowledge, technological content knowledge, and technological pedagogical knowledge. They are related to the ability of users to use technology to accomplish different tasks. The technological content knowledge is related to how technology can affect teaching material and vice versa. The final element is technological pedagogical knowledge, which has to do with the constraints that technology may impose on teaching material. This stems from the fact that certain technology is not developed for educational purposes and should be accustomed to suit educational purposes [57,58]. Accordingly, the following hypothesis can be formed:

Hypothesis 1 (H1). *TPACK will positively affect teachers' CU in the e-learning environment.*

2.2. Technology Self-Efficacy (TSE)

Self-efficacy is an effective factor that can reflect how students' own belief in their abilities to use technology affects their acceptance of the learning environment. Therefore, self-efficacy and learning are two factors that can, interactively and dynamically, affect each other [59]. Self-efficacy in the e-learning environment is considered an intrinsic motivator as far as continuous intention is concerned. It usually refers to the degree of confidence that users have in making use of technology [60]. Technology self-efficacy is usually identified as the ability to use technology without facing any crucial problems. It embraces two subdivisions: the estimation of result (users' estimations about their own input) and estimation efficacy (users' estimations in achieving the final result) [61–63]. Within the environment of e-learning, self-efficacy is highly connected to users' own beliefs regarding technology. Some believe that using technology is tremendously easy and achievable, while others may share a contradictory belief, as they may face problems in their ability to learn the appropriate way of using technology [64]. This simply implies that whenever users have a high level of technological self-belief, they may perceive the whole system properly; hence, they will be able to continue using the technology in a positive way. Accordingly, the following hypothesis may be formulated:

Hypothesis 2 (H2). *Technology Self-efficacy will positively affect teachers' and students' CU in the e-learning environment.*

2.3. Technology Acceptance Model (TAM)

A review of recent studies has shown that certain variables are crucial to understanding the reasons behind the continuous intention to use e-learning. Regarding Davis's TAM [65], it has been proven that PU and PEOU are the most influential factors in users' continuous-use intentions. Interestingly, PU is more effective than PEOU when one wants to deal with the use of technology [60]. This study focuses only on two constructs within the TAM theory, which have proven effective in investigating the continuous use of technology; these are perceived usefulness and ease of use. [65] adopted the view that the perceived usefulness (PU) and the perceived ease of use (PEOU) of technology form the baseline for examining individuals' usage intentions. PU is defined as the degree to which a person believes that using a technological system supports user performance, whereas PEOU tends to refer to the degree to which a person believes that use will be free of effort. Due to the fact that PEOU has proved to be of great significance only during the early-acceptance stage of technology use [54,65], PEOU may not directly affect teachers' and students' CU in the e-learning environment. Hence, we hypothesize the following:

Hypothesis 3 (H3). *The level of PU will positively affect teachers' and students' CU.*

Hypothesis 4 (H4). *The level of PEOU will not affect teachers' and students' CU.*

2.4. Perceived Organizational Support (POS)

Organizational support theory has a close relation with how users perceive organizational feedback regarding the use of technology. Users' perceptions may vary in accordance with an organization's rewards, fairness, and supervisor support [42]. To put it differently, when users had a positive attitude towards their organization, they were more willing to pursue their intention to use its technology and vice versa. The organization has a crucial role in enhancing the use of technology by motivating users internally at the organizational level [44,66]. Related literature is guided by the fact that organizational support for use of technology has a high impact on teachers' and students' CU to adhere to computer technology, particularly in technical support [67–69]. In this respect, teachers and students may have different subjective perspectives of the role of educational institutions (colleges and universities) in creating a motivational atmosphere regarding the continuous use of technology. Hence, it is hypothesized that if teachers and students had a positive perception of organizational support (POS), they would support the continuous usage of the technology. Thus, the following hypothesis may be formed:

Hypothesis 5 (H5). *The level of POS support will positively affect teachers' CU.*

2.5. Controlled Motivation (CTRLM)

According to [43], students' intrinsic and extrinsic motivations can be dealt with in terms of a hieratical model in which three factors play an influential role: contextual, situational, and global factors. One of the motivations that have a massive effect on user perception is called Controlled Motivation (CTRLM), which refers to a source of negative perception that is illustrated by the pressure that students may be under, both internally and externally. This type of pressure may lead to maladaptive outcomes, which are, in turn, illustrated by a combination of negative effects, perceived incompetence, and dissatisfaction [44]. A more updated view regarding controlled motivation is given by [44], who proposed two different types within CTRLM: introjected and external regulation. Introjected regulation has a close relationship with an individual's behavioral engagement, such as obligation, avoidance of guilt, ego-enhancement, and internal rewards. On the other hand, external regulation affects behavior engagement, including compliance, external rewards, and avoidance of punishments.

Hypothesis 6 (H6). *Controlled Motivation negatively affects students' continuous intention to use e-learning platforms.*

2.6. The Proposed Research Models

Based on the review of previous studies, it has been noticed that most of the studies on continuous intention have focused on one model that has factors that may be crucial to teachers but not students [60,70]. Therefore, this study attempted to build two models that can meet both teachers' and students' continuous intentions to meet e-learning demands. The proposed research models rely on these hypotheses, as illustrated in Figures 1 and 2.

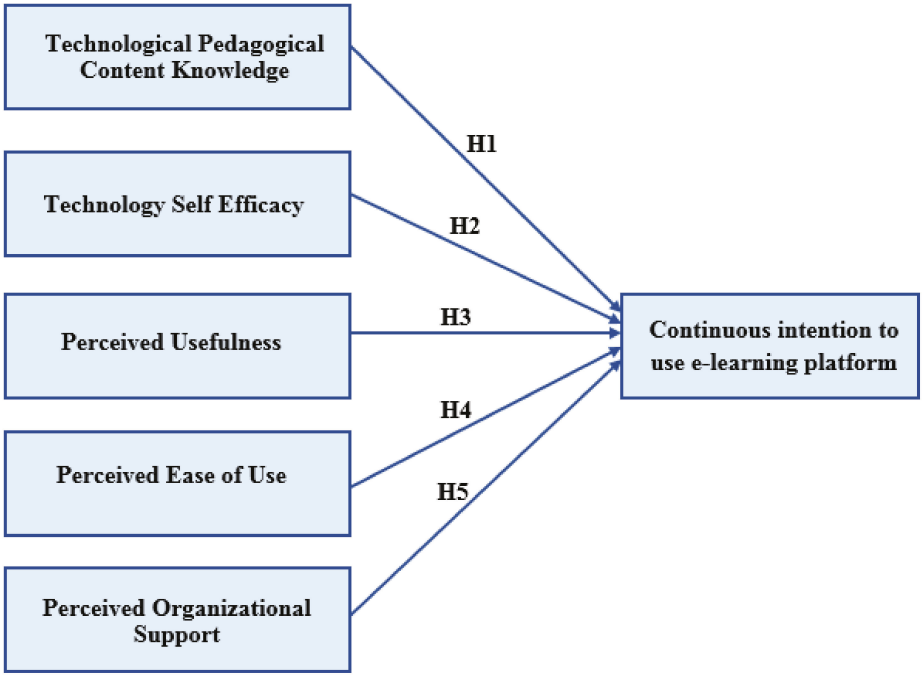

Figure 1. The e-learning technology model adopted for teachers. Note: H1–5 = Hypotheses 1–5.

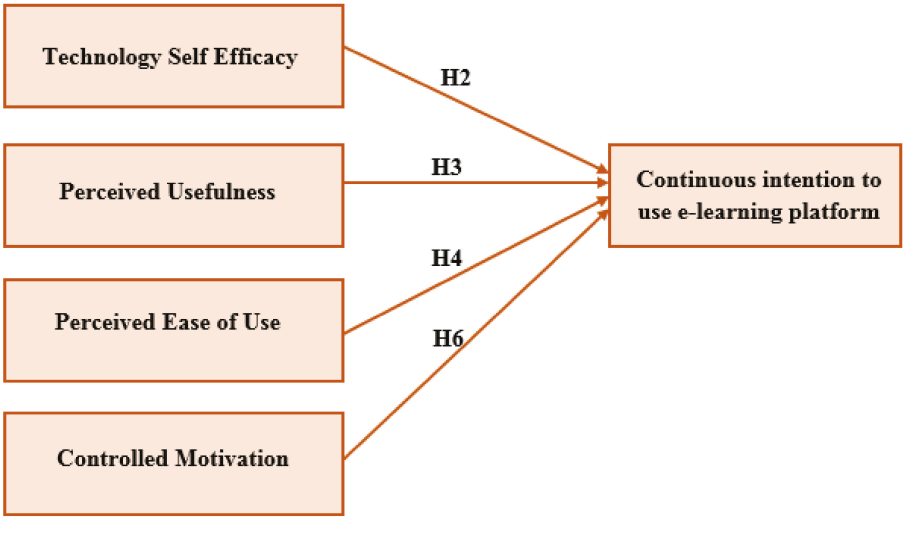

Figure 2. The e-learning technology model adopted for students.

3. Methodology

3.1. Participants

The participants (n = 372) were classified into two categories—teachers and students—in accordance with the two proposed models. This two-level selection was based on certain

influential factors. First, participants were identified as having sufficient experience in using e-learning platforms. Second, two different surveys were prepared and distributed to the two targeted groups of participants. The two surveys may have had similar and shared questions, but the teachers' survey may have had additional questions that the students' survey may have lacked. Responses were received from 372 participants. Therefore, the total number of teacher participants in this study was 30. These were college instructors with a nearly equal female-to-male gender ratio. On the other hand, the total number of student participants was 342. These were students at The British University in Dubai (BUiD).

3.2. Data Collection

During the winter semester of 2019/2020, the data was collected through online surveys from individuals studying at the British University in Dubai (BUiD) from 15 January to 20 February 2020. The aggregated response rate was 93%; 400 questionnaires were circulated, out of which 372 were answered by respondents. This means that 372 questionnaires were filled out correctly and found to be useful, while 28 were rejected because of missing values. The prospective sample size was 306 respondents with respect to a population of 1500. Thus, the sample size of 372 correct responses was suitable, according to [71], because—bearing in mind the required sample size—the sample size of 372 is a higher figure. Thus, this sample size could be reviewed using structural equation modeling [72] to verify the hypotheses. It must be noted that hypotheses were based on the current theories and were adjusted to the e-learning context. In order to assess the measurement model, the researchers used structural equation modeling (SEM) [72]. Further treatment was performed using a final path model.

3.3. Students' Personal Information/Demographic Data

The assessment of personal/demographic data is covered in Table 2. The percentage of males was 53%, while for females it was 47%. A total of 33% of students had ages ranging from 18 to 29 years, while 67% of respondents were aged over 29. In terms of academic background, 39% were students from the Faculty of Engineering and IT, 35% were from the Faculty of Education and 26% belonged to the Faculty of Business and Law. The majority of respondents came from sophisticated families and held university degrees; 49% of participants had bachelor's degrees, 42% had master's degrees, and 9% had a doctoral degree. When the respondents were ready to volunteer and were easily approachable, the purposive sampling approach was used as per [3]. This sample was created by students coming from different faculties, with different ages, enrolling in diverse programs at different levels. Moreover, with the aid of IBM SPSS Statistics ver. 23, the demographic data was evaluated. Table 2 depicts the complete demographic data of the respondents.

Table 2. Demographic data of the respondents.

Criterion	Factor	Frequency	Percentage
Gender	Female	175	47%
	Male	197	53%
Age	Between 18 and 29	122	33%
	Between 30 and 39	98	26%
	Between 40 and 49	88	24%
	Between 50 and 59	64	17%
Faculties	Faculty of Engineering and IT	145	39%
	Faculty of Education	129	35%
	Faculty of Business and Law	98	26%
Education qualification	Bachelor	182	49%
	Master	157	42%
	Doctorate	33	9%

3.4. Study Instrument

The survey instrument used to validate the hypothesis was determined in this research. The survey, consisting of 30 items, was used for the measurement of seven constructs in the questionnaire. Table 3 depicts the sources of the constructs. The questions from prior studies were modified in order to enhance the appropriateness of the research.

Table 3. Constructs and their sources.

Constructs	Number of Items	Source
CU	2	[73–75]
CTRLM	5	[43]
TPACK	4	[39,76]
TSE	7	[41,77]
PEOU	3	[40]
PU	4	[40]
POS	5	[42]

Note: TPACK = Technological pedagogical content knowledge; TSE = Technology self-efficacy; PEOU = Perceived ease of use; PU = Perceived usefulness; POS = Perceived organizational support; CTRLM = Controlled motivation; CU = Continuous intention to use e-learning platform.

3.5. Pilot Study for the Questionnaire

A pilot study was conducted to check the reliability of the questionnaire items. Approximately 40 students and teachers were chosen on a random basis from the given population to establish the pilot study. The sample size was set based on 10% of the aggregated sample size of this study (400 students and teachers) and thus adhered strictly to the research criteria. Cronbach's alpha test was utilized for the computation of internal reliability through IBM SPSS Statistics ver. 23, in order to judge the outcomes of the pilot study. Thus, the appropriate findings were shown for the measurement items. A value of 0.7 was taken to be an acceptable value for the reliability coefficient, considering the model for social science research [14]. Tables 4 and 5 show the Cronbach's alpha values for the seven measurement scales for teachers and students.

Table 4. Cronbach's alpha values for the pilot study (Cronbach's alpha ≥ 0.70) for teachers (Model A).

Constructs	Cronbach's Alpha
CU	0.756
TPACK	0.779
TSE	0.864
PEOU	0.889
PU	0.734
POS	0.852

Note: TPACK = Technological pedagogical content knowledge; TSE = Technology self-efficacy; PEOU = Perceived ease of use; PU = Perceived usefulness; POS = Perceived organizational support; CU = Continuous intention to use e-learning platform.

3.6. Survey Structure

The questionnaire survey given to students and teachers had two sections. Within the first part, personal data was given to gather information about students and teachers. The second section had a group of questions related to the main factors of the proposed models. The teachers' questionnaire had six sub-sections coinciding with the six factors proposed in the model. Similarly, the students' questionnaire had five sub-sections related to the five factors proposed in the model. With the help of the five-point Likert Scale, the 42

items were evaluated. The scales included the following: (1) strongly disagree, (2) disagree, (3) neutral, (4) agree, and (5) strongly agree.

Table 5. Cronbach's alpha values for the pilot study (Cronbach's alpha ≥ 0.70) for students (Model B).

Constructs	Cronbach's Alpha
CU	0.872
CTRLM	0.881
TSE	0.798
PEOU	0.736
PU	0.797

Note: TSE = Technology self-efficacy; PEOU = Perceived ease of use; PU = Perceived usefulness; CTRLM = Controlled motivation; CU = Continuous intention to use e-learning platform.

4. Findings and Discussion

4.1. Data Analysis

Along with the help of SmartPLS V.3.2.7 software, the partial least squares-structural equation modeling (PLS-SEM) was utilized to conduct the data analysis in this research [15]. The assessment approach had two steps of a structural model and a measurement model allowed to study the collected data [16]. There were various reasons for choosing PLS-SEM in the study. First, as the research is an extension of a current theory, PLS-SEM was considered the best option [17]. Second, the complex models within exploratory research can be effectively tackled with the help of PLS-SEM [18]. Third, PLS-SEM analyzes a complete model as a single unit, so there is no need to divide it [19]. Lastly, PLS-SEM provides concurrent analysis for measurement, as well as a structural model, leading to more accurate calculations [20].

4.2. Convergent Validity

In order to review the measurement model, it was suggested by [16] that construct reliability—including composite reliability (CR), Dijkstra–Henseler's rho (pA), and Cronbach's alpha (CA) and validity (including convergent and discriminant validity)—must be considered. Cronbach's alpha (CA) has values between 0.782 and 0.895, as Tables 6 and 7 show, in order to determine construct reliability. These statistics are higher than the threshold value of 0.7 [78]. According to Tables 6 and 7, the outcomes also show that the composite reliability (CR) has values from 0.796 to 0.882; these values are evidently bigger than the recommended value of 0.7 [79]. As an alternative, the construct reliability must be appraised by researchers by means of the Dijkstra–Henseler's rho (pA) reliability coefficient [80]. Like CA and CR, the reliability coefficient ρA must show 0.70 or higher in exploratory studies and values of more than 0.80 or 0.90 for further stages of research [78,81,82]. The reliability coefficient ρA of each measurement construct is above 0.70 according to Tables 6 and 7. According to these outcomes, the construct reliability is verified and all the constructs were considered to be accurate.

Convergent validity can be measured by testing the average variance extracted (AVE) as well as the factor loading [16]. Tables 6 and 7 suggest that all values of factor loadings exceeded the threshold value of 0.7. Moreover, Tables 6 and 7 show that the values obtained for the AVE were higher than the threshold value of 0.5, ranging from 0.509 to 0.718. Depending on the expected results, the convergent reliability can be obtained for all the constructs.

Table 6. Convergent validity results that ensure acceptable values (Factor loading, Cronbach's alpha, composite reliability (CR), Dijkstra–Henseler's rho (pA) $\geq$ 0.70 and average variance extracted (AVE) > 0.5) (Model A).

Constructs	Items	Factor Loading	Cronbach's Alpha	CR	pA	AVE
Technology Self-Efficacy	TSE1	0.775	0.874	0.799	0.832	0.536
	TSE2	0.736				
	TSE3	0.820				
	TSE4	0.901				
	TSE5	0.756				
	TSE6	0.723				
	TSE7	0.797				
Technological Pedagogical Content Knowledge	TPACK 1	0.711	0.829	0.882	0.791	0.552
	TPACK 2	0.869				
	TPACK 3	0.909				
	TPACK 4	0.790				
Perceived Ease of Use	PEOU1	0.829	0.844	0.812	0.817	0.661
	PEOU2	0.847				
	PEOU3	0.746				
Perceived Usefulness	PU1	0.734	0.816	0.828	0.825	0.623
	PU2	0.766				
	PU3	0.889				
	PU4	0.850				
Perceived Organizational Support	POS1	0.729	0.863	0.814	0.883	0.718
	POS2	0.848				
	POS3	0.758				
	POS4	0.819				
	POS5	0.878				
Continuous intention to use e-learning platform	CU1	0.796	0.815	0.876	0.898	0.673
	CU2	0.801				

Table 7. Convergent validity results that ensure acceptable values (Factor loading, Cronbach's alpha, composite reliability, Dijkstra–Henseler's rho $\geq$ 0.70 & AVE > 0.5) (Model B).

Constructs	Items	Factor Loading	Cronbach's Alpha	CR	PA	AVE
Technology Self-Efficacy	TSE1	0.726	0.782	0.833	0.823	0.705
	TSE2	0.826				
	TSE3	0.710				
	TSE4	0.868				
	TSE5	0.746				
	TSE6	0.733				

Table 7. *Cont.*

Constructs	Items	Factor Loading	Cronbach's Alpha	CR	PA	AVE
Perceived Ease of Use	PEOU1	0.763	0.895	0.800	0.836	0.559
	PEOU2	0.890				
	PEOU3	0.849				
Perceived Usefulness	PU1	0.793	0.856	0.879	0.808	0.696
	PU2	0.709				
	PU3	0.873				
	PU4	0.821				
Controlled Motivation	CTRL1	0.832	0.805	0.796	0.807	0.700
	CTRL2	0.802				
	CTRL3	0.875				
	CTRL4	0.810				
	CTRL5	0.796				
Continuous intention to use e-learning platform	CU1	0.725	0.878	0.818	0.816	0.509
	CU2	0.878				

4.3. Discriminant Validity

The two criteria that were suggested should be measured to obtain the measurement of discriminant validity were the Fornell–Larcker measure and the Heterotrait–Monotrait ratio (HTMT) [16]. As per the findings of Tables 8 and 9, these needs have been verified by the Fornell–Larker criterion as each AVE value, together with its square root, exceeds the value of the correlation of AVE with other constructs [83].

Table 8. Fornell–Larcker Scale (Model A).

	TSE	TPACK	PEOU	PU	POS	CU
TSE	**0.876**					
TPACK	0.165	**0.845**				
PEOU	0.125	0.253	**0.802**			
PU	0.569	0.487	0.558	**0.790**		
POS	0.187	0.202	0.291	0.115	**0.787**	
CU	0.369	0.198	0.378	0.383	0.178	**0.803**

Note: TPACK = Technological pedagogical content knowledge; TSE = Technology self-efficacy; PEOU = Perceived ease of use; PU = Perceived usefulness; POS = Perceived organizational support; CU = Continuous intention to use e-learning platform.

Table 9. Fornell–Larcker Scale (Model B).

	TSE	PEOU	PU	CTRL	CU
TSE	**0.768**				
PEOU	0.368	**0.801**			
PU	0.267	0.229	**0.887**		
CTRL	0.649	0.492	0.399	**0.844**	
CU	0.422	0.327	0.302	0.188	**0.870**

Note: TSE = Technology self-efficacy; PEOU = Perceived ease of use; PU = Perceived usefulness; CTRLM = Controlled motivation; CU = Continuous intention to use e-learning platform.

Tables 10 and 11 show the outcomes of the HTMT ratio, indicating that the threshold value of 0.85 is bigger than the values of other constructs [27], and hence confirming the HTMT ratio. These outcomes play a role in the evaluation of the discriminant validity. The outcomes of the analysis indicated a smooth and simple assessment of the measurement model in terms of the model's validity and reliability. To conclude, it can be said that the collected data was appropriate for additionally evaluating the structural model.

Table 10. Heterotrait–Monotrait Ratio (HTMT) (Model A).

	TSE	TPACK	PEOU	PU	POS	CU
TSE						
TPACK	0.560					
PEOU	0.136	0.487				
PU	0.266	0.363	0.556			
POS	0.296	0.200	0.270	0.544		
CU	0.389	0.635	0.378	0.638	0.555	

Note: TPACK = Technological pedagogical content knowledge; TSE = Technology self-efficacy; PEOU = Perceived ease of use; PU = Perceived usefulness; POS = Perceived organizational support; CU = Continuous intention to use e-learning platform.

Table 11. Heterotrait–Monotrait Ratio (HTMT) (Model B).

	TSE	PEOU	PU	CTRL	CU
TSE					
PEOU	0.232				
PU	0.506	0.436			
CTRL	0.392	0.457	0.503		
CU	0.697	0.609	0.210	0.264	

Note TSE = Technology self-efficacy; PEOU = Perceived ease of use; PU = Perceived usefulness; CTRLM = Controlled motivation; CU = Continuous intention to use e-learning platform.

4.4. Model Fit

The subsequently mentioned fit measures are ensured by SmartPLS: the standard root mean square residual (SRMR), exact fit criteria, Euclidean distance (d_ULS), geodesic distance (d_G), Chi-square, Normed Fit Index (NFI), and RMS Theta show the model fit in PLS-SEM [84]. The difference between experimental correlations and the correlation matrix inferred from model [85] are indicated by SRMR, and values smaller than 0.08 are assumed to serve as good model-fit measures [86]. The NFI values that are higher than 0.90 point out a good model fit [87]. The NFI is a ratio of the Chi-square value of the proposed model to the null model (also known as the benchmark model) [88]. The NFI increases with larger parameters and therefore, the NPI is not suggested as a model-fit pointer [85]. The discrepancy between the empirical covariance matrix and the covariance matrix, inferred from the composite factor model, is indicated by the metrics of squared Euclidean distance (d_ULS) and the geodesic distance (d_G) [80,85]. The RMS Theta helps in the measurement of the degree of outer model residuals correlation and is appropriate for reflective models only [88]. The nearer the RMS Theta value is to zero, the more superior the PLS-SEM model, and their values of less than 0.12, are assumed to be a good fit, with anything other than this suggesting an absence of fit [89]. The saturated model evaluates the correlation between all constructs, as recommended by [29], while the approximate model takes all the effects and model structure into consideration. The RMS Theta value was 0.073 in Model A and 0.073 in Model B, as given in Tables 12 and 13, which gives an idea that the specific goodness-of-fit for the PLS-SEM model was big enough to prove global PLS model validity.

Table 12. Model fit indicators (Model A).

	Complete Model	
	Saturated Model	**Estimated Mod**
SRMR	0.031	0.041
d_ULS	0.786	3.216
d_G	0.565	0.535
Chi-Square	466.736	473.348
NFI	0.624	0.627
RMS Theta	0.073	

Table 13. Model fit indicators (Model B).

	Complete Model	
	Saturated Model	**Estimated Mod**
SRMR	0.012	0.031
d_ULS	0.605	2.317
d_G	0.516	0.506
Chi-Square	461.646	472.347
NFI	0.633	0.642
RMS Theta	0.061	

4.5. Hypotheses Testing Using PLS-SEM

The interdependence between different theoretical constructs related to the structural model was studied by using a combination of the structural equation model with maximum-likelihood estimation and SmartPLS [38,39]. This indicates the procedure of analysis of the proposed hypothesis. About 83% and 71% variance were found within the continuous intention to use the e-learning platform, as shown in Tables 14 and 15, which indicates a high predictive power of Models A and B [37]. For all the proposed hypotheses, outcomes of the PLS-SEM technique provided the beta (β) values, *t*-values, and *p*-values, which have been stated in Tables 16 and 17. It is evident that each and every hypothesis is supported by all the researchers. The empirical data backs hypotheses H1, H2, H3, H4, H5, and H6 on the basis of the analyzed data. The standardized path coefficients and path significances are demonstrated in Figures 3 and 4.

Table 14. R^2 of the endogenous latent variables (Model A).

Constructs	**R^2**	**Results**
Continuous intention to use e-learning platform	0.832	High

Table 15. R^2 of the endogenous latent variables (Model B).

Constructs	**R^2**	**Results**
Continuous intention to use e-learning platform	0.709	High

In Model A, technological pedagogical content knowledge (TPACK), technology self-efficacy (TSE), perceived ease of use (PEOU), perceived usefulness (PU), and perceived organizational support (POS) have significant effects on continuous intention to use the e-learning platform (CU) ((β = 0.336, $p < 0.001$), (β = 0.426, $p < 0.05$), β = 0.589, $p < 0.05$), (β = 0.625, $p < 0.05$) and (β = 0.553, $p < 0.001$), respectively); hence, H1, H2, H3, H4, and H5 are supported.

In Model B, technology self-efficacy (TSE), perceived ease of use (PEOU), perceived usefulness (PU), and controlled motivation (CTRLM) have significant effects on continuous intention to use the e-learning platform (CU) (($\beta = 0.290$, $p < 0.001$), ($\beta = 0.357$, $p < 0.05$), $\beta = 0.465$, $p < 0.05$) and ($\beta = 0.243$, $p < 0.05$), respectively); hence, H2, H3, H4, and H6 are supported.

Table 16. Hypotheses testing of the research model (significant at ** $p <= 0.01$, * $p < 0.05$) (Model A).

H	Relationship	Path	*t*-Value	*p*-Value	Direction	Decision
H1	TPACK -> CU	0.336	12.223	0.001	Positive	Supported **
H2	TSE -> CU	0.426	5.269	0.026	Positive	Supported *
H3	PU -> CU	0.589	6.716	0.018	Positive	Supported *
H4	PEOU -> CU	0.625	5.584	0.023	Positive	Supported *
H5	POS -> CU	0.553	16.108	0.000	Positive	Supported **

Note: TPACK = Technological pedagogical content knowledge; TSE = Technology self-efficacy; PEOU = Perceived ease of use; PU = Perceived usefulness; POS = Perceived organizational support; CU = Continuous intention to use e-learning platform.

Table 17. Hypotheses testing of the research model (significant at ** $p <= 0.01$, * $p < 0.05$) (Model B).

H	Relationship	Path	*t*-Value	*p*-Value	Direction	Decision
H2	TSE -> CU	0.290	14.578	0.000	Positive	Supported **
H3	PEOU -> CU	0.357	3.116	0.043	Positive	Supported *
H4	PU -> CU	0.465	2.646	0.035	Positive	Supported *
H6	CTRLM -> CU	0.243	4.361	0.033	Positive	Supported *

Note: TSE = Technology self-efficacy; PEOU = Perceived ease of use; PU = Perceived usefulness; CTRLM = Controlled motivation; CU = Continuous intention to use e-learning platform.

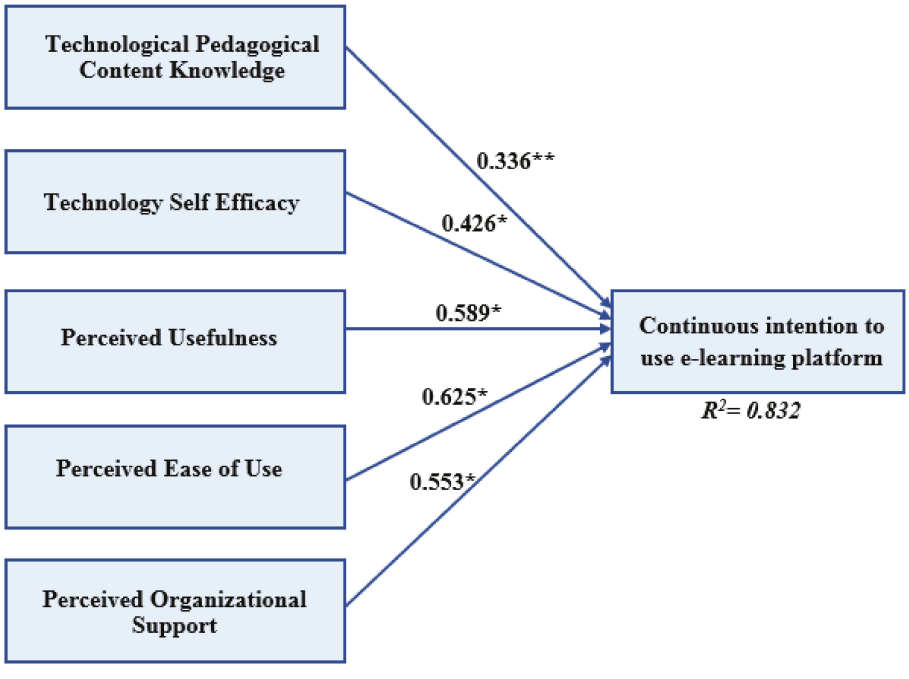

Figure 3. Path coefficient of the model (significant at ** $p <= 0.01$, * $p < 0.05$) (Model A).

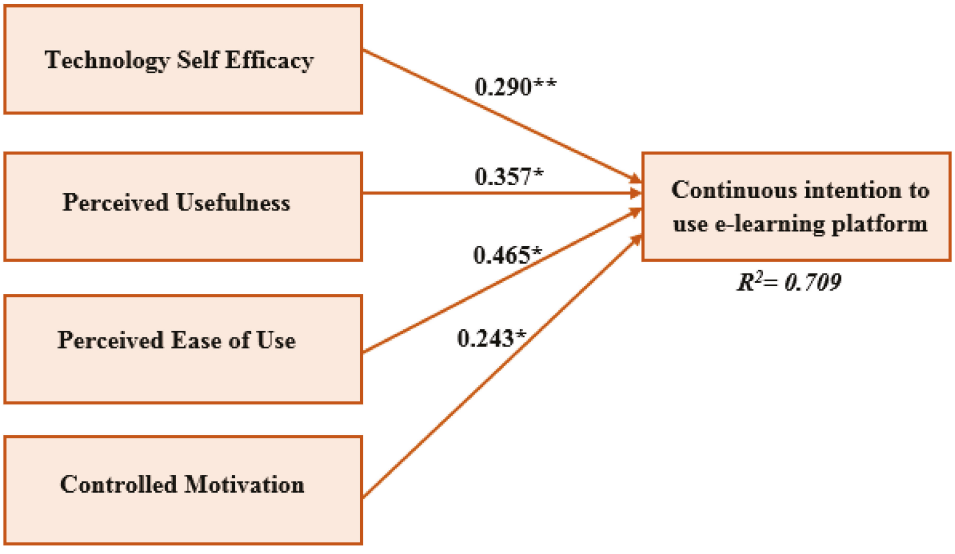

Figure 4. Path coefficient of the model (significant at ** $p < = 0.01$, * $p < 0.05$) (Model B).

5. Discussion and Conclusions

This study proposed two unique CU models that took into consideration factors that affect both instructors' and students' attitudes. The two models can be theoretically extended to enhance other technology-supported educational environments and instructional processes. The first research model of instructors' CU was proposed, taking into consideration certain social cognitive theory along with personal, behavioral, and environmental elements that are closely related to instructors' CU. In general, the results of SEM analysis has supported all the proposed hypothesis. From a practical perspective, this study has proven that POS is the most influential factor that affects instructors' CU of e-learning platforms.

Ref. [90–92] seem to agree with the current conclusion in stating that POS could motivate its staff members, leading to an upgrade of the organization. However, a study by [93] placed an emphasis on staff members' personalities and readiness to change. This implies that the lack of organizational support may have negative consequences. One of the studies by [60] has proposed that when instructors feel that there is no adequate organizational support, they are less likely to continuously use the technology, especially in an educational atmosphere where instructors are supposed to implement various in-class pedagogical changes to facilitate a better learning environment for the students.

In fact, POS is not the only factor that affects instructors' CU, but rather instructors' TPACK is another key factor that affects CU. Most of the previous research has proven that the organizational support may affect users' motivation to use the technology. A study by [94] put emphasis on the effect of PACK in facilitating the e-learning process by both teachers and adult learners. This seems to be in line with the results obtained from this study where TPACK affected, to great extent, the teachers' performance. It is assumed that whenever teachers' content, technological, and pedagogical knowledge is high, it implies that his or her ability to change the teaching material to suit the newly used technology will be more practical and effective. Obviously, teachers are more motivated when there is reliable technical support and IT staff that can facilitate the process of establishing new computer-supported knowledge [66,68].

Similarly, CTRLM has an effect on students' CU. CTRLM, along with technology self-efficacy (TSE), are the key factors that have a great impact on students' CU. The study has proven that CTRLM is connected deeply with the willingness to use the e-learning teaching platform. The higher the motivation is, the more effective results are obtained. Previous studies have tackled the effect of CTRLM on students' performance and have proven that there are both pedagogical and non-pedagogical elements that affect motivation [95–97]. These studies have indicated that technology development has placed a positive effect on motivation, as it urges students to get engaged in the new learning platform. This makes motivation very high and involvement in classes even higher, increasing the students' willingness to learn, and thus, using the technology continuously [98,99].

5.1. Practical Implications

According to the study outcomes, salient factors were noticed in determining users' acceptance of the e-learning system and technological pedagogical content knowledge, technology self-efficacy, perceived ease of use, perceived usefulness, perceived organizational support, and controlled motivation were significant. Hence, upon implementing a new e-learning system, faculty members must be informed about the system's features and its technical issues, as well as its usefulness, so that they feel self-assured and can gain insight into the system. To increase faculty use of the e-learning system, faculty members are of the view that universities should deliver workshops, extensive training, and awareness programs on the system's features, benefits, and usage [100]. In addition, a national survey was conducted, in which 57% of faculty members said that they could become more productive [100] if the use of e-learning technology was communicated thoroughly to them in their courses. Faculty members in this study also reported that, if they were aware of the positive impact of such technology on student learning, they would be inspired to learn and use the e-learning system. Hence, extended online help and periodic training programs for the e-learning system should be offered by the universities to ensure increased use of the system and to increase the faculty's self-efficacy. As a result, this overall phenomenon would help faculty members to obtain practical exposure, acquire better skills and become more proficient in using the e-learning system; as a result of this, their use of the system would be enhanced. In accordance with our findings, there was a weak influence on faculty attitudes toward the e-learning system because of facilitating conditions. Consequently, to ensure the smooth running of the e-learning system, attention must be given by universities to ensuring technical support and reliable network access. Moreover, online and face-to-face support and guidance should be provided by the universities for faculty members to ensure that members have positive attitudes toward the e-learning system and, consequently, that they become capable of extended use of the system [65,101].

5.2. Limitations and Further Research

This study has the following limitations. Only one university in the United Arab Emirates was considered for studying the impact of factors on the adoption of e-learning systems; this is the key limitation. The applicability and pertinence of this study would have been enhanced if more institutes and universities in the United Arab Emirates had been taken into consideration. With further analysis and insight into the e-learning system, the researchers could have better understood the factors that influence such a system. In addition, the participation of a limited number of students (372) is another limitation. According to [45], a survey questionnaire method was used for data collection. For appropriate and improved outcomes, an enhanced instrument is being sought and more institutes will be taken into account from other regions, such as the Arab Gulf, including Kuwait, Saudi Arabia, and Bahrain. Furthermore, invitations to join the study will be sent to more students, and researchers will conduct focus groups and interviews for suitable results. Moreover, we are seeking to implement the e-learning system in specific Arab universities that have contributed to the research.

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

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

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

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

References

1. Wagner, N.; Hassanein, K.; Head, M. Who is responsible for e-learning success in higher education? A stakeholders' analysis. *J. Educ. Technol. Soc.* **2008**, *11*, 26–36.
2. Kituyi, G.; Tusubira, I. A Framework for the Integration of E-learning in Higher Education Institutions in Developing Countries. *Int. J. Educ. Dev. Using ICT* **2013**, *9*, 19–36.
3. Luo, N.; Zhang, M.; Qi, D. Effects of Different Interactions on Students' Sense of Community in E-learning Environment. *Comput. Educ.* **2017**, *115*, 153–160. [CrossRef]
4. Hong, J.-C.; Tai, K.-H.; Hwang, M.-Y.; Kuo, Y.-C.; Chen, J.-S. Internet Cognitive Failure Relevant to Users' Satisfaction with Content and Interface Design to Reflect Continuance Intention to Use a Government E-learning System. *Comput. Human Behav.* **2017**, *66*, 353–362. [CrossRef]
5. Alghizzawi, M.; Habes, M.; Salloum, S.A.; Ghani, M.A.; Mhamdi, C.; Shaalan, K. The Effect of Social Media Usage on Students'E-learning Acceptance in Higher Education: A case study from the United Arab Emirates. *Int. J. Inf. Technol. Lang. Stud.* **2019**, *3*, 13–26.
6. Alshurideh, M.; al Kurdi, B.; Salloum, S.A. Examining the Main Mobile Learning System Drivers' Effects: A Mix Empirical Examination of Both the Expectation-Confirmation Model (ECM) and the Technology Acceptance Model (TAM). In Proceedings of the 5th International Conference on Advanced Intelligent Systems and Informatics (AISI'19), Cairo, Egypt, 26–28 October 2019; Volume 1058.
7. Salloum, S.A.; Al-Emran, M.; Habes, M.; Alghizzawi, M.; Ghani, M.A.; Shaalan, K. Understanding the Impact of Social Media Practices on E-Learning Systems Acceptance. In Proceedings of the 5th International Conference on Advanced Intelligent Systems and Informatics (AISI'19), Cairo, Egypt, 26–28 October 2019; Volume 1058.
8. Alshurideh, M.; Salloum, S.A.; al Kurdi, B.; Monem, A.A.; Shaalan, K. Understanding the quality determinants that influence the intention to use the mobile learning platforms: A practical study. *Int. J. Interact. Mob. Technol.* **2019**, *13*, 157–183. [CrossRef]
9. Al-Emran, M.; Arpaci, I.; Salloum, S.A. An empirical examination of continuous intention to use m-learning: An integrated model. *Educ. Inf. Technol.* **2020**, *25*, 2899–2918. [CrossRef]
10. Salloum, S.A.; Shaalan, K. *Investigating Students' Acceptance of E-Learning System in Higher Educational Environments in the UAE: Applying the Extended Technology Acceptance Model (TAM)*; The British University in Dubai: Dubai, UAE, 2018.
11. Salloum, S.A.; Alhamad, A.Q.M.; Al-Emran, M.; Monem, A.A.; Shaalan, K. Exploring Students' Acceptance of E-Learning Through the Development of a Comprehensive Technology Acceptance Model. *IEEE Access* **2019**, *7*, 128445–128462. [CrossRef]
12. Muqtadiroh, F.A.; Nisafani, A.S.; Saraswati, R.M.; Herdiyanti, A. Analysis of User Resistance Towards Adopting E-Learning. *Procedia Comput. Sci.* **2019**, *161*, 123–132. [CrossRef]
13. Tawafak, R.M.; Romli, A.B.T.; Arshah, R.b.A.; Malik, S.I. Framework design of university communication model (UCOM) to enhance continuous intentions in teaching and e-learning process. *Educ. Inf. Technol.* **2019**, *25*, 817–843. [CrossRef]
14. al Kurdi, B.; Alshurideh, M.; Salloum, S.A. Investigating a theoretical framework for e-learning technology acceptance. *Int. J. Electr. Comput. Eng.* **2020**, *10*, 6484–6496. [CrossRef]
15. Al-Maroof, R.S.; Salloum, S.A. An Integrated Model of Continuous Intention to Use of Google Classroom. In *Recent Advances in Intelligent Systems and Smart Applications*; Springer: Cham, Switzerland, 2020; Volume 295.
16. Al-Maroof, R.A.; Arpaci, I.; Al-Emran, M.; Salloum, S.A.; Shaalan, K. Examining the Acceptance of WhatsApp Stickers Through Machine Learning Algorithms. In *Recent Advances in Intelligent Systems and Smart Applications*; Springer: Cham, Switzerland, 2020; Volume 295.
17. Al-Maroof, R.S.; Salloum, S.A.; AlHamadand, A.Q.M.; Shaalan, K. A Unified Model for the Use and Acceptance of Stickers in Social Media Messaging. In *International Conference on Advanced Intelligent Systems and Informatics*; Springer: Cham, Switzerland, 2019; pp. 370–381.
18. Al Kurdi, B.; Alshurideh, M.; Salloum, S.A.; Obeidat, Z.M.; Al-dweeri, R.M. An Empirical Investigation into Examination of Factors Influencing University Students' Behavior towards Elearning Acceptance Using SEM Approach. *Int. J. Interact. Mob. Technol.* **2020**, *14*, 19–41. [CrossRef]

19. Al-Maroof, R.S.; Salloum, S.A.; Al Hamadand, A.Q.; Shaalan, K. Understanding an Extension Technology Acceptance Model of Google Translation: A Multi-Cultural Study in United Arab Emirates. *Int. J. Interact. Mob. Technol.* **2020**, *14*, 157–178. [CrossRef]
20. Demetriadis, S.; Barbas, A.; Molohides, A.; Palaigeorgiou, G.; Psillos, D.; Vlahavas, I.; Tsoukalas, I.; Pombortsis, A. 'Cultures in negotiation': Teachers' acceptance/resistance attitudes considering the infusion of technology into schools. *Comput. Educ.* **2003**, *41*, 19–37. [CrossRef]
21. Buckenmeyer, J. Revisiting teacher adoption of technology: Research implications and recommendations for successful full technology integration. *Coll. Teach. Methods Styles J.* **2008**, *4*, 7–10. [CrossRef]
22. Aldunate, R.; Nussbaum, M. Teacher adoption of technology. *Comput. Human Behav.* **2013**, *29*, 519–524. [CrossRef]
23. Eyyam, R.; Yaratan, H.S. Impact of use of technology in mathematics lessons on student achievement and attitudes. *Soc. Behav. Personal. Int. J.* **2014**, *42*, 31S–42S. [CrossRef]
24. Nelson, M.J.; Hawk, N.A. The impact of field experiences on prospective preservice teachers' technology integration beliefs and intentions. *Teach. Teach. Educ.* **2020**, *89*, 103006. [CrossRef]
25. Scherer, R.; Siddiq, F.; Tondeur, J. All the same or different? Revisiting measures of teachers' technology acceptance. *Comput. Educ.* **2020**, *143*, 103656. [CrossRef]
26. Raes, A.; Depaepe, F. A longitudinal study to understand students' acceptance of technological reform. When experiences exceed expectations. *Educ. Inf. Technol.* **2020**, *25*, 533–552. [CrossRef]
27. Salloum, S.A.; Al-Emran, M.; Shaalan, K.; Tarhini, A. Factors affecting the E-learning acceptance: A case study from UAE. *Educ. Inf. Technol.* **2019**, *24*, 509–530. [CrossRef]
28. Cheng, Y. Antecedents and consequences of e-learning acceptance. *Inf. Syst. J.* **2011**, *21*, 269–299. [CrossRef]
29. Al-Gahtani, S.S. Empirical investigation of e-learning acceptance and assimilation: A structural equation model. *Appl. Comput. Inform.* **2016**, *12*, 27–50. [CrossRef]
30. Sunny, S.; Patrick, L.; Rob, L. Impact of cultural values on technology acceptance and technology readiness. *Int. J. Hosp. Manag.* **2019**, *77*, 89–96. [CrossRef]
31. Jacobs, J.V.; Hettinger, L.J.; Huang, Y.H.; Jeffries, S.; Lesch, M.F.; Simmons, L.A.; Verma, S.K.; Willetts, J.L. Employee acceptance of wearable technology in the workplace. *Appl. Ergon.* **2019**, *78*, 148–156. [CrossRef]
32. Scherer, R.; Siddiq, F.; Tondeur, J. The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Comput. Educ.* **2019**, *128*, 13–35. [CrossRef]
33. Estriegana, R.; Medina-Merodio, J.-A.; Barchino, R. Student acceptance of virtual laboratory and practical work: An extension of the technology acceptance model. *Comput. Educ.* **2019**, *135*, 1–14. [CrossRef]
34. Taherdoost, H. A review of technology acceptance and adoption models and theories. *Procedia Manuf.* **2018**, *22*, 960–967. [CrossRef]
35. Dehghani, M.; Kim, K.J.; Dangelico, R.M. Will smartwatches last? Factors contributing to intention to keep using smart wearable technology. *Telemat. Inform.* **2018**, *35*, 480–490. [CrossRef]
36. Obal, M. What drives post-adoption usage? Investigating the negative and positive antecedents of disruptive technology continuous adoption intentions. *Ind. Mark. Manag.* **2017**, *63*, 42–52. [CrossRef]
37. Choi, B.K.; Yeo, W.-D.; Won, D. The implication of ANT (Actor-Network-Theory) methodology for R&D policy in open innovation paradigm. *Knowl. Manag. Res. Pract.* **2018**, *16*, 315–326.
38. Ifinedo, P. Examining students' intention to continue using blogs for learning: Perspectives from technology acceptance, motivational, and social-cognitive frameworks. *Comput. Human Behav.* **2017**, *72*, 189–199. [CrossRef]
39. Mishra, P.; Koehler, M.J. Technological pedagogical content knowledge: A framework for teacher knowledge. *Teach. Coll. Rec.* **2006**, *108*, 1017–1054. [CrossRef]
40. Davis, F.D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* **1989**, *13*, 319–340. [CrossRef]
41. Bandura, A. Social Foundations of Thought and Action. In *The Health Psychology Reader*; SAGE: Englewood Cliffs, NJ, USA, 1986.
42. Eisenberger, R.; Huntington, R.; Hutchison, S.; Sowa, D. Perceived organizational support. *J. Appl. Psychol.* **1986**, *71*, 500. [CrossRef]
43. Vallerand, R.J. Toward a hierarchical model of intrinsic and extrinsic motivation. In *Advances in Experimental Social Psychology*; Elsevier: Amsterdam, The Netherlands, 1997; Volume 29, pp. 271–360.
44. Ryan, R.M.; Deci, E.L. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* **2000**, *55*, 68. [CrossRef]
45. Al-Emran, M.; Salloum, S.A. Students' Attitudes Towards the Use of Mobile Technologies in e-Evaluation. *Int. J. Interact. Mob. Technol.* **2017**, *11*, 195–202. [CrossRef]
46. Sabah, N.M. Motivation factors and barriers to the continuous use of blended learning approach using Moodle: Students' perceptions and individual differences. *Behav. Inf. Technol.* **2020**, *39*, 875–898. [CrossRef]
47. Al-Busaidi, K.A. An empirical investigation linking learners' adoption of blended learning to their intention of full e-learning. *Behav. Inf. Technol.* **2013**, *32*, 1168–1176. [CrossRef]
48. Al-Busaidi, K.A.; Al-Shihi, H. Key factors to instructors' satisfaction of learning management systems in blended learning. *J. Comput. High. Educ.* **2012**, *24*, 18–39. [CrossRef]

49. Hwang, I.-H.; Tsai, S.-J.; Yu, C.-C.; Lin, C.-H. An empirical study on the factors affecting continuous usage intention of double reinforcement interactive e-portfolio learning system. In Proceedings of the 2011 6th IEEE Joint International Information Technology and Artificial Intelligence Conference, Chongqing, China, 20–22 August 2011; Volume 1, pp. 246–249.
50. Lin, K.-M. e-Learning continuance intention: Moderating effects of user e-learning experience. *Comput. Educ.* **2011**, *56*, 515–526. [CrossRef]
51. Ho, C.-H. Continuance intention of e-learning platform: Toward an integrated model. *Int. J. Electron. Bus. Manag.* **2010**, *8*, 206.
52. Lee, M.-C. Explaining and predicting users' continuance intention toward e-learning: An extension of the expectation–confirmation model. *Comput. Educ.* **2010**, *54*, 506–516. [CrossRef]
53. Shulman, L. Knowledge and teaching: Foundations of the new reform. *Harv. Educ. Rev.* **1987**, *57*, 1–23. [CrossRef]
54. 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 preservice teachers. *J. Res. Technol. Educ.* **2009**, *42*, 123–149. [CrossRef]
55. Chai, C.S.; Koh, J.H.L.; Tsai, C.-C.; Tan, L.L.W. Modeling primary school pre-service teachers' Technological Pedagogical Content Knowledge (TPACK) for meaningful learning with information and communication technology (ICT). *Comput. Educ.* **2011**, *57*, 1184–1193. [CrossRef]
56. Oner, D. A virtual internship for developing technological pedagogical content knowledge. *Australas. J. Educ. Technol.* **2020**, *36*, 27–42. [CrossRef]
57. Koehler, M.; Mishra, P. What is technological pedagogical content knowledge (TPACK)? *Contemp. Issues Technol. Teach. Educ.* **2009**, *9*, 60–70. [CrossRef]
58. Koehler, M.J.; Shin, T.S.; Mishra, P. How do we measure TPACK? Let me count the ways. In *Educational Technology, Teacher Knowledge, and Classroom Impact: A Research Handbook on Frameworks and Approaches*; IGI Global: Hershey, PA, USA, 2012; pp. 16–31.
59. Lu, Y.-L.; Lien, C.-J. Are they learning or playing? Students' perception traits and their learning self-efficacy in a game-based learning environment. *J. Educ. Comput. Res.* **2020**, *57*, 1879–1909. [CrossRef]
60. Cai, J.; Yang, H.H.; Gong, D.; MacLeod, J.; Zhu, S. Understanding the continued use of flipped classroom instruction: A personal beliefs model in Chinese higher education. *J. Comput. High. Educ.* **2019**, *31*, 137–155. [CrossRef]
61. Venkatesh, V. Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the technology acceptance model. *Inf. Syst. Res.* **2000**, *11*, 342–365. [CrossRef]
62. Chen, K.; Chen, J.V.; Yen, D.C. Dimensions of self-efficacy in the study of smart phone acceptance. *Comput. Stand. Interfaces* **2011**, *33*, 422–431. [CrossRef]
63. Cebeci, U.; Ertug, A.; Turkcan, H. Exploring the determinants of intention to use self-checkout systems in super market chain and its application. *Manag. Sci. Lett.* **2020**, *10*, 1027–1036. [CrossRef]
64. Bailey, A.A.; Pentina, I.; Mishra, A.S.; Mimoun, M.S.B. Mobile payments adoption by US consumers: An extended TAM. *Int. J. Retail Distrib. Manag.* **2017**, *45*, 626–640. [CrossRef]
65. Fathema, N.; Shannon, D.; Ross, M. Expanding the Technology Acceptance Model (TAM) to Examine Faculty Use of Learning Management Systems (LMSs) In Higher Education Institutions. *J. Online Learn. Teach.* **2015**, *11*, 210–232.
66. Lam, S.; Cheng, R.W.; Choy, H.C. School support and teacher motivation to implement project-based learning. *Learn. Instr.* **2010**, *20*, 487–497. [CrossRef]
67. Kurup, V.; Hersey, D. The changing landscape of anesthesia education: Is Flipped Classroom the answer? *Curr. Opin. Anesthesiol.* **2013**, *26*, 726–731. [CrossRef]
68. Missildine, K.; Fountain, R.; Summers, L.; Gosselin, K. Flipping the classroom to improve student performance and satisfaction. *J. Nurs. Educ.* **2013**, *52*, 597–599. [CrossRef]
69. Bøe, T.; Gulbrandsen, B.; Sørebø, Ø. How to stimulate the continued use of ICT in higher education: Integrating information systems continuance theory and agency theory. *Comput. Human Behav.* **2015**, *50*, 375–384. [CrossRef]
70. Bölen, M.C. Exploring the determinants of users' continuance intention in smartwatches. *Technol. Soc.* **2020**, *60*, 101209. [CrossRef]
71. Krejcie, R.V.; Morgan, D.W. Determining sample size for research activities. *Educ. Psychol. Meas.* **1970**, *30*, 607–610. [CrossRef]
72. Chuan, C.L.; Penyelidikan, J. Sample size estimation using Krejcie and Morgan and Cohen statistical power analysis: A comparison. *J. Penyelid. IPBL* **2006**, *7*, 78–86.
73. Bhattacherjee, A. Understanding information systems continuance: An expectation-confirmation model. *MIS Q.* **2001**, *25*, 351–370. [CrossRef]
74. Almahamid, S.; Rub, F.A. Factors that determine continuance intention to use e-learning system: An empirical investigation. *Int. Conf. Telecommun. Tech. Appl. Proc. CSIT* **2011**, *5*, 242–246.
75. Thiruchelvi, A.; Koteeswari, S. A conceptual framework of employees' continuance intention to use e-learning system. *Asian J. Res. Bus. Econ. Manag.* **2013**, *3*, 14–20.
76. Grandgenett, N.F. Perhaps a matter of imagination: TPCK in mathematics education. In *Handbook of Technological Pedagogical Content Knowledge (TPCK) for Educators*; Routledge: London, UK, 2008; p. 145166.
77. Compeau, D.R.; Higgins, C.A. Computer self-efficacy: Development of a measure and initial test. *MIS Q.* **1995**, *19*, 189–211. [CrossRef]
78. Nunnally, J.C.; Bernstein, I.H. *Psychometric Theory*; McGraw-Hill: New York, NY, USA, 1994.
79. Kline, R.B. *Principles and Practice of Structural Equation Modeling*; Guilford Publications: New York, NY, USA, 2015.

80. Dijkstra, T.K.; Henseler, J. Consistent and asymptotically normal PLS estimators for linear structural equations. *Comput. Stat. Data Anal.* **2015**, *81*, 10–23. [CrossRef]
81. Hair, J.F.; Ringle, C.M.; Sarstedt, M. PLS-SEM: Indeed a silver bullet. *J. Mark. Theory Pract.* **2011**, *19*, 139–152. [CrossRef]
82. Henseler, J.; Ringle, C.M.; Sinkovics, R.R. The use of partial least squares path modeling in international marketing. In *New Challenges to International Marketing*; Emerald Group Publishing Limited: Bingley, UK, 2009; pp. 277–319.
83. Fornell, C.; Larcker, D.F. Evaluating Structural Equation Models With Unobservable Variables and Measurement Error. *J. Mark. Res.* **1981**, *18*, 39–50. [CrossRef]
84. Trial, D. Model Fit. Available online: https://www.modelfit.com/ (accessed on 20 December 2020).
85. Hair, J.F., Jr.; Hult, G.T.M.; Ringle, C.; Sarstedt, M. *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*; Sage Publications: New York, NY, USA, 2016.
86. Hu, L.; Bentler, P.M. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychol. Methods* **1998**, *3*, 424. [CrossRef]
87. Bentler, P.M.; Bonett, D.G. Significance tests and goodness of fit in the analysis of covariance structures. *Psychol. Bull.* **1980**, *88*, 588. [CrossRef]
88. Lohmöller, J.B. *Latent Variable Path Modeling with Partial Least Squares*; Physica: Heidelberg, Germany, 1989.
89. Henseler, J.; Dijkstra, T.K.; Sarstedt, M.; Ringle, C.M.; Diamantopoulos, A.; Straub, D.W.; Ketchen, D.J., Jr.; Hair, J.F.; Hult, G.T.M.; Calantone, R.J. Common beliefs and reality about PLS: Comments on Rönkkö and Evermann (2013). *Organ. Res. Methods* **2014**, *17*, 182–209. [CrossRef]
90. Baranik, L.E.; Roling, E.A.; Eby, L.T. Why does mentoring work? The role of perceived organizational support. *J. Vocat. Behav.* **2010**, *76*, 366–373. [CrossRef] [PubMed]
91. Aselage, J.; Eisenberger, R. Perceived organizational support and psychological contracts: A theoretical integration. *J. Organ. Behav. Int. J. Ind. Occup. Organ. Psychol. Behav.* **2003**, *24*, 491–509. [CrossRef]
92. Stamper, C.L.; Johlke, M.C. The impact of perceived organizational support on the relationship between boundary spanner role stress and work outcomes. *J. Manag.* **2003**, *29*, 569–588.
93. Mitchell, J.I.; Gagné, M.; Beaudry, A.; Dyer, L. The role of perceived organizational support, distributive justice and motivation in reactions to new information technology. *Comput. Human Behav.* **2012**, *28*, 729–738. [CrossRef]
94. Alsofyani, M.M.; Aris, B.b.; Eynon, R.; Majid, N.A. A preliminary evaluation of short blended online training workshop for TPACK development using technology acceptance model. *Turkish Online J. Educ. Technol.* **2012**, *11*, 20–32.
95. Siahaan, A.U.; Aji, S.B.; Antoni, C.; Handayani, Y. Online Social Learning Platform vs E-Learning for Higher Vocational Education in Purpose to English Test Preparation. In Proceedings of the 7th International Conference on English Language and Teaching (ICOELT 2019), Padang, Indonesia, 4–5 November 2019; pp. 76–82.
96. Heidig, S.; Clarebout, G. Do pedagogical agents make a difference to student motivation and learning? *Educ. Res. Rev.* **2011**, *6*, 27–54. [CrossRef]
97. Corbalan, G.; Kester, L.; van Merriënboer, J.J.G. Towards a personalized task selection model with shared instructional control. *Instr. Sci.* **2006**, *34*, 399–422. [CrossRef]
98. Zaric, N.; Lukarov, V.; Schroder, U. A Fundamental Study for Gamification Design: Exploring Learning Tendencies' Effects. *Int. J. Serious Games* **2020**, *7*, 3–25. [CrossRef]
99. Hanif, M. Students' Self-Regulated Learning in Iconic Mobile Learning System in English Cross-Disciplined Program. *Anatol. J. Educ.* **2020**, *5*, 121–130. [CrossRef]
100. Fathema, N.; Sutton, K.L. Factors influencing faculty members' Learning Management Systems adoption behavior: An analysis using the Technology Acceptance Model. *Int. J. Trends Econ. Manag. Technol.* **2013**, *2*, 20–28.
101. Panda, S.; Mishra, S. E-Learning in a Mega Open University: Faculty attitude, barriers and motivators. *EMI Educ. Media Int.* **2007**, *44*, 323–338. [CrossRef]

education sciences

MDPI

Article

Multi-Class Assessment Based on Random Forests

Mehdi Berriri, Sofiane Djema, Gaëtan Rey and Christel Dartigues-Pallez *

Université Côte d'Azur, CNRS, I3S, CEDEX 2, 06103 Nice, France; mehdi.berriri@etu.univ-cotedazur.fr (M.B.); sofiane.djema@etu.univ-cotedazur.fr (S.D.); Gaetan.REY@univ-cotedazur.fr (G.R.)
* Correspondence: Christel.DARTIGUES-PALLEZ@univ-cotedazur.fr

Abstract: Today, many students are moving towards higher education courses that do not suit them and end up failing. The purpose of this study is to help provide counselors with better knowledge so that they can offer future students courses corresponding to their profile. The second objective is to allow the teaching staff to propose training courses adapted to students by anticipating their possible difficulties. This is possible thanks to a machine learning algorithm called Random Forest, allowing for the classification of the students depending on their results. We had to process data, generate models using our algorithm, and cross the results obtained to have a better final prediction. We tested our method on different use cases, from two classes to five classes. These sets of classes represent the different intervals with an average ranging from 0 to 20. Thus, an accuracy of 75% was achieved with a set of five classes and up to 85% for sets of two and three classes.

Keywords: machine learning; Random Forest; selection feature; orientation

Citation: Berriri, M.; Djema, S.; Rey, G.; Dartigues-Pallez, C. Multi-Class Assessment Based on Random Forests. *Educ. Sci.* **2021**, *11*, 92. https://doi.org/10.3390/educsci11030092

Academic Editor: Sandra Raquel Gonçalves Fernandes

Received: 14 January 2021
Accepted: 18 February 2021
Published: 26 February 2021

1. Introduction

Many universities face the same problem: the first-year failure rate is far too high. For example, the average failure rate observed in the first year in France is more than 60%. This is a problem common to all areas, albeit to varying degrees [1]. It is particularly high in France in science, psychology, law, and so on. This phenomenon concerns both small and very large cohorts [2,3]. This failure rate is found both in training courses that do not select their students and in so-called selective training. Faced with this recurring problem, universities are struggling to find a valid explanation as the origins of students is varied. In fact, the studies taken before entering the University are often different from one student to another, the institutions of origin are often of very heterogeneous levels, the students have followed optional courses which vary from others, etc. All these points make it difficult at first glance to identify student profiles that may fail. This is where techniques based on machine learning can be of great help.

The field of educational data mining (EDM) is facing this kind of problem. More precisely, supervised learning techniques have been used more and more in recent years in very different fields, and each time they provide significant support. For example, they are used for image recognition, activity recognition, text classification, the medical field, and emotions. The idea of our project is to use the power of these algorithms to detect the profiles of students entering university who are likely to fail. The algorithms will use learning methods based on the profiles of students from previous years and their success or failure at university. Once such students are identified, it is then possible to set up an adapted pedagogy (tutoring, adapted course, hours of support, etc.) in order to break the spiral of failure. Thus, instead of waiting for students to fail, it will then be possible to set up processes to limit the number of failures.

In this article, we propose a new approach to learn the profiles of failed students. In the following we will present the existing work in terms of prediction of success and profiles, we will present the data from which we work as well as the pre-treatments that we apply to them (in particular in terms of anonymization), we will present then our approach

based on Random Forests and we will end with the presentation of our experiments and the promising results that we obtained.

For several years now, research around the world has been interested in this problem of too high failure rates, especially during the first year of university. The vast majority of approaches used are based on supervised learning techniques, but some innovate by testing semi-supervised approaches. In [4], for example, the authors decided to test semi-supervised approaches in order to learn the profiles of failed students. They use an algorithm called the Tri-Learning algorithm. This algorithm is particularly used in cases where it is difficult to label the data [5]. Three classifiers from the original set example are set. These classifiers are then refined using unlabeled examples in the tri-training process.

In his thesis [6], Pojon focuses on the comparison of three significant algorithms of supervised learning to know how much they improve prediction performance. The algorithms he chose are linear regression, decision trees, and Naive Bayes. These algorithms were tested on two datasets, and the decision trees and Naive Bayes gave the best performances.

In [7], the authors compare five algorithms. Faced with the large number of features that make up their dataset, they use feature selection algorithms to reduce it. Among the algorithms used are decision trees and Bayesian networks, Multi-Layer Perceptron (part of the Artificial Neural Network), Sequential Minimal Optimisation (part of the SVM class), and 1-NN (1-Nearest Neighborhood). In this study, the dataset used concerns students admitted to Computer Science in some undergraduate colleges in Kolkata from 2006 to 2012. The years 2006 to 2010 are used to create the model (309 students) and the years 2011 and 2012 are used to test it (104 students). In this study, decision trees gave the best results.

As we have just seen, the studies seem to focus on the same algorithms: the SVM, the Naïve Bayes, the decision trees, as well as the logistic regression [1,4–6,8]. The study of existing works has led, on many occasions, to note the points which make this task difficult [8]. These difficulties are intrinsically linked to the nature of the data to be considered. Indeed, these are generally few (it is not uncommon to have to consider student promotions of the order of several tens). These small samples correlate with a large number of characteristics, which can penalize the performances of certain algorithms. Another point that can greatly reduce the performance of predictions is the fact that the datasets considered are generally poorly balanced. It is only very recently that some studies are interested in Random Forests (RF) to solve this problem [2,8–10].

The RF algorithm consists of building a set of decision trees (Figure 1). The final decision is taken following a majority vote. The construction of these decision trees is based on two very important concepts: Bagging and Random Feature Selection. Bagging consists of not considering the entire dataset when creating one of the trees of the forest but a subset of randomly selected elements. The Random Selection Feature is a process that occurs at the time of the creation of the nodes of a tree: unlike a conventional decision tree where the creation of a new node is done by choosing the best argument among the set possible arguments, the choice of an argument for the creation of a new node in a forest is done by choosing the best argument among a subset of randomly chosen arguments.

One of the very important results in [11] is the theoretical proof of convergence of RF. That is, the generalization error of RF converges to a limit value with the increase of the number of trees. As demonstrated by Breiman, an upper bound for the generalization error (denoted PE in the original paper) is given by:

$$PE^* \leq \frac{\overline{\rho}(1 - s^2)}{s^2}$$

where $\overline{\rho}$ is the mean value of the correlation between two RF models and is the strength of the set of classifiers (i.e., decision trees). As stated by Breiman [11], although the bound is likely to be loose, it fulfills the same suggestive function for RF as Vapnik-Chervonekis-type bounds do for other types of classifiers. It shows that the two ingredients involved in the generalization error for RF are the strength of the individual classifiers in the forest and the correlation between them. The s^2 ratio is the correlation divided by the square of the

strength. In understanding the functioning of RF, this ratio will be a helpful guide–the smaller it is the better, which means that the error of generalization of the forest will be smaller if the decision from each individual tree is both more reliable and less correlated to the decisions from the other trees.

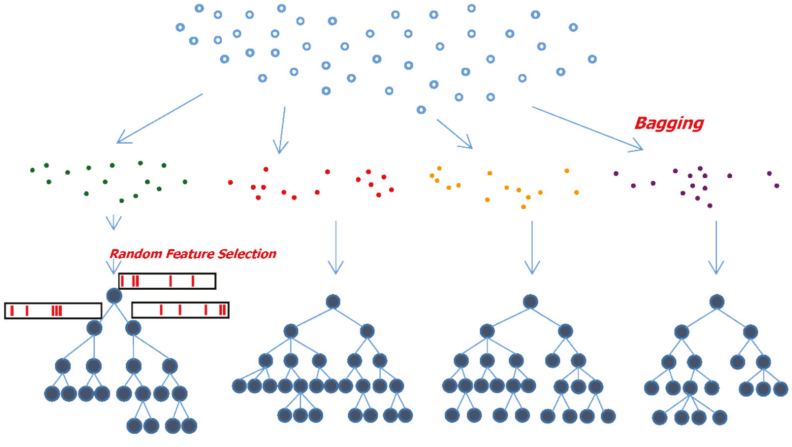

Figure 1. Principle of Random Forests (RF).

We have already been able to test the effectiveness of this algorithm in very different domains: the recognition of human activity from videos from a Kinect [12] or worn sensors [13], the classification of short texts [14], the prediction of marine currents, etc. We thus decided to apply this algorithm to try to find relevant answers to the problem of improving student success in their studies.

With this statement, we will show the interest in this technique to answer our problem in this article. We are also aware that the nature of the data we consider requires us to make a significant effort of pre-processing. This is because the dataset contains almost empty columns because all the subjects existing in the school are represented and not all the pupils follow them. Moreover, only the grades obtained in the subjects by the students, the best and worst averages, and the class average are available. A lot of information must be removed from the initial dataset in order to ensure strict anonymization (surname, first name, date of birth, address, telephone, e-mail, etc.). It is also important to be able to obtain indications such as the evolution of grades in a subject over time in order to have a better profile. In the following, we will show the strategies we have implemented to improve our prediction.

2. Material and Methods

2.1. Methodology

We started our research work on a promotion of students which well-illustrated our problem of failure: the first year of studies in computer science of the IUT (Technical University Institute) of Université Côte d'Azur. Indeed, this training is in great demand (more than 2000 requests for 80 to 90 places) and moreover, a selection of the best candidates is made each year. This selection is made by looking at scores in mathematics, French, and English, and also by looking at behavioral remarks. Despite this strict selection process, the failure rate in the first year is still far too high (around 30% in the first year). This type of training is therefore very symptomatic of our problem. We, therefore, worked our study from the list of students admitted in this formation for the school years 2017–2018 and 2018–2019. Over these two years, we have combined, for each student, all the information available concerning their past education with the marks obtained in the first semester of our reference training.

There were many data on past education. They contained the specialties followed by the students, the marks obtained in each subject by the candidate themself and also the average of the class, as well as the highest and lowest averages of the class. We also had information that characterized the high schools of the candidates (number of students having passed the bac, number of honors, etc.). We had thus obtained a fairly large set of data if we considered the information that characterized each candidate for our training. Thus, our initial data set contained more than 400 characteristics. After removing characteristics that were either unnecessary or related to the identity of the students and adding additional information, the dataset contained more than 1000 characteristics. On the other hand, the volume of candidates cumulated over the two years taken into account was around 150, which was obviously not much. Techniques associated with the processing of small data sets should therefore be used in the context of this study [15].

2.2. Data Processing

We first had to encode the data used by the RF algorithm [16]. Some information contained in the dataset can negatively impact predictions, for example, features that mainly have no values. The dataset had been designed to be compatible with all the university's courses, whatever their field of study. Thus, we considered all the subjects that can be followed before arriving at the university. Our dataset, therefore, contained all existing subjects. However, a high school student cannot follow all these subjects. For example, a high school student who had specialized in a literary field will not have any marks in the final year before university in science subjects such as mathematics or physics and chemistry. Thus, when we consider for our study a university department of studies specializing in computer science, all the columns in our dataset that correspond to non-scientific subjects (literature, Greek, Latin, etc.) will therefore be almost empty.

Thus, we decided not to consider those data. Finally, we added some new data correlated with our dataset. As an example, information about the school of the future students was initially the name of this school and its localization. We replace that information—which was useless—with numerical information related to the number of high school students who obtain their bac, as well as the number of honors.

Finally, we had in our dataset many grades obtained by high school students during their schooling. We calculated other information from those grades, such as the evolution of their results in mathematics during the year, the difference between their best and worst results in a semester, etc. More precisely, we calculated information for each of the subjects followed by the candidates:

- Averages: calculation of averages in all subjects for each term of each school year two by two (première and terminale).
- Weighted averages: weighted averages taking the subjects' scientific and non-scientific matters. A higher coefficient is given depending on whether the calculated average corresponds to a scientific or non-scientific subject.
- Candidate average delta: candidate average—class average.
- Low average delta: candidate average—lowest average.
- High average delta: highest average—candidate average.
- Difference: difference between the highest and the lowest mark in each subject.
- Range: highest average—lowest average.
- Bonus: (candidate average—average of the class)/(highest average—lowest average).

It emerged that the characteristics that have the greatest influence on prediction obviously depend on one training course to another. Nevertheless, low and high delta values stood out as important characteristics in several formations, so they seemed to be relevant metrics. As seen in article [17], the selection of features improved prediction rates, and the method which gives the best results was the "wrapper" technique [18]. Finally, in order to select a feature on the columns, we chose the "wrapper" technique.

2.3. Class Balancing

The pupil averages to be expected were separated into different intervals called classes. We separated the pupils by their final results. We decided to test different configurations in order to see which one was the most significant. The class numbered 0 was also the most obvious: success, ie [10, 20] in Table 1 means that students obtained an average between 10 and 20, and failure, ie [0, 10[in Table 1 means that students obtained an average between 0 and 10 (10 is excluded in this range). We also tried to be more precise in our classification. For class number 1, we, therefore, considered pupils with great difficulties (result below 7.5), those who succeed (result above 10), and pupils who have results below 10 but who have not completely failed. As this study was exploratory, we have divided this class into two other similar classes (classes 2 and 3) by slightly modifying the boundaries of each group, as shown in Table 1. Finally, we also planned to separate the pupils into two other even more precise configurations (classes 4 and 5) by dividing the notes into 4 and then 5 intervals.

Table 1. Set of classes.

Class Set Label	Class Intervals				
0	[0, 10[	[10, 20]			
1	[0, 7.5[	[7.5, 10[	[10, 20]		
2	[0, 8[	[8, 10[	[10, 20]		
3	[0, 8.5[	[8.5, 10.5[	[10.5, 20]		
4	[0, 7.5[	[7.5, 9[	[9, 10.5[	[10.5, 20]	
5	[0, 7.5[	[7.5,8.5[	[8.5, 9.5[	[9.5, [	[10.5, 20]

The students were not equally distributed in all the classes, as the majority of them had global results between 9 and 14, as shown in Figure 2. Thus, the number of classes and the limits of those classes had a direct incidence on the number of students in each class.

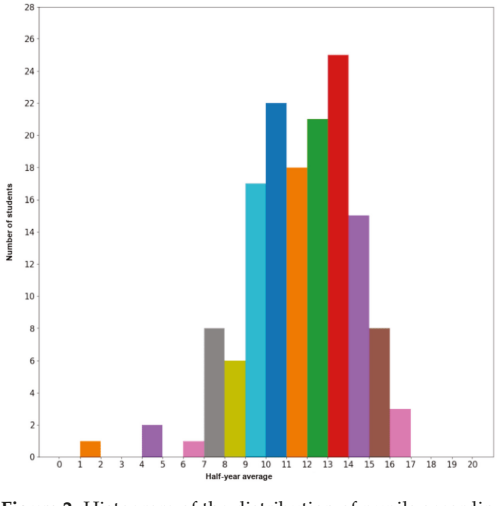

Figure 2. Histogram of the distribution of pupils according to their average.

An imbalance in the distribution of students according to class sets was notable. As shown in the following histogram on the left, the class of the pupils having average in range [10, 20] contained 119 pupils out of a total of 154, which represents 77% of the

total population. One of the first treatments we had to do on our dataset was to do class balancing in order to minimize this problem.

Class balancing was done in two ways. First of all, the classes with a large population have been divided into subclasses as shown in Figure 3. The disadvantage here is that for some groups, the class population has become small, less than 10 students by class.

The second way was to manage this in the RF algorithm itself, which allows the data of a class with a small population to have greater weight in decision-making. It can be therefore useful to add the hyperparameter "class_weight = balanced". This automatically weighed classes inversely proportional to the frequency of their appearance in the data.

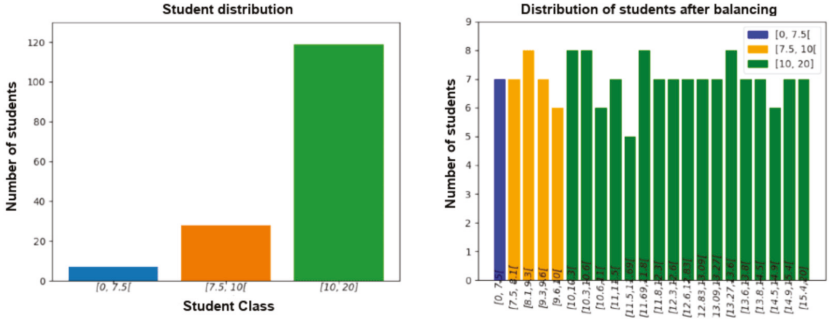

Figure 3. Histograms before and after class balancing.

2.4. Generation of Models

To build the different RF models, we first separate the dataset into two subsets: training dataset, which will be used to build the model, and test dataset, which will be used to test the efficiency of the model. To build the training dataset, we randomly selected 80% of the initial dataset, and the remaining 20% was affected by the test dataset.

As shown in Figure 4, the training dataset was used to build 6 RF corresponding to our six sets of classes, hence creating six models. We then tested those models with the test dataset.

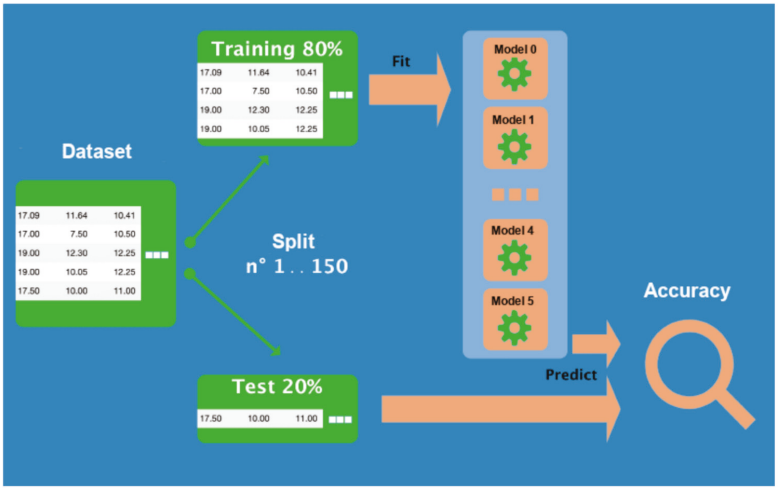

Figure 4. Model generation process.

Since the number of students was limited to 154, the test dataset consists of 31 students. This means that each student had a very important weight for the accuracy of the model. Making a mistake on a single student immediately dropped the precision. To face this problem, using several "seeds" made it possible to generate models whose supplied data and definition classes were identical in order to be able to combine their results. The number of seeds used was 150. Thus, the precision was then no longer calculated on 31 pupils but on 4650. This reduces the individual impact of 90% of a prediction and thus had a result closer to reality, to which more confidence can be allowed. In addition, the choice to use 150 "seeds" was taken in order to have results in a reasonable time while having a sufficiently large sample. So, for each set of classes, 150 "seeds" were used.

2.5. Crossover of Models

The objective of cross-checking the models was to use the results of the six models trained with the same training set and with different target class sets so that they correct their errors with each other (Figure 5). The goal was to merge their prediction on the same student to try to determine which predictions were wrong.

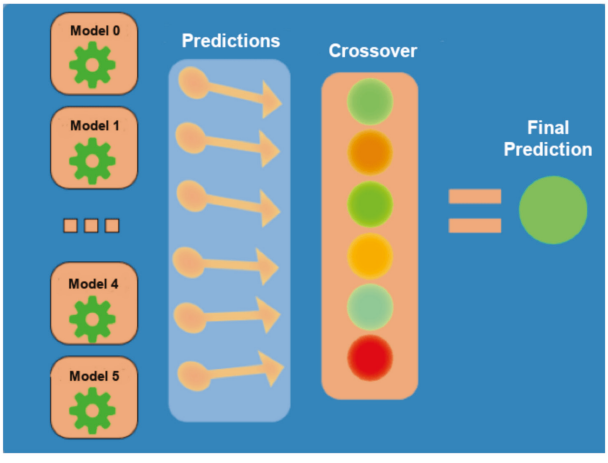

Figure 5. Models crossover.

The crossing of the predictions of the six models can be represented as shown in Figure 6. It represents a tree where the last level contains 1081 sheets, where each final leaf represents a possible combination of the predictions of the six models (2 * 3 * 3 * 3 * 4 * 5) for each set of target classes. Each level of the tree represents a model. Each node has m children and m represents the number of classes of the following model. As shown in Figure 6, the class bubble in range [10, 20] of the class 0 set is connected to three child nodes because the class 1 set is made up of three classes. Finally, a tree is made for each set of definition classes.

To know which class to predict after each combination, a training phase was carried out beforehand by making predictions using the models on the students of the test sets. Students whose actual class was different may refer to the same combination of predictions. If so, the class with the highest proportion of students for that combination becomes the combination prediction.

With this process, some combinations may never be encountered. If applicable, the prediction of the model whose definition classes are the same as that of the tree becomes the prediction of the combination of models. For example, if for a combination 'abcdef' whose target classes are in range [0, 10[with 10 cases encountered and range [10, 20] with 30 cases

encountered then for this combination the final decision will be a success: the prediction associated with this sheet will therefore be in range [10, 20].

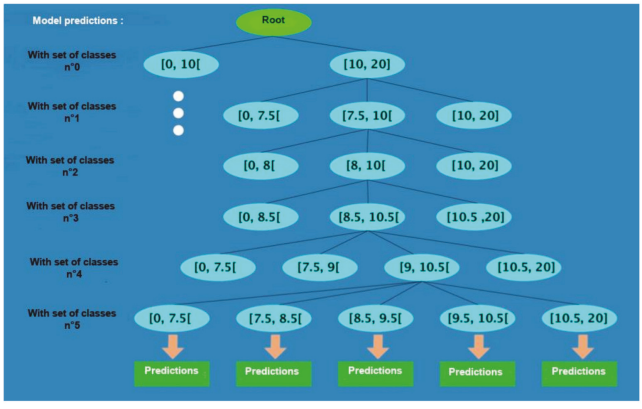

Figure 6. Tree representing the model crossing method.

3. Results

Following the increase in data, the balancing of classes, and the crossing of models; the results obtained are shown in Table 2. The "Virtual" column represents the results of the models with the balancing of the classes performed by cutting them up and the "Balanced" column represents the results of the models with the balancing classes performed by weighting the data in the algorithm.

Table 2. Accuracy for two balancings.

	Before Crossing		*After Crossing*	
Class Set	**Virtual**	**Balanced**	**Virtual**	**Balanced**
0	0.79	0.81	0.85	0.85
1	0.71	0.755	0.848	0.84
2	0.68	0.76	0.84	0.83
3	0.63	0.68	0.77	0.76
4	0.59	0.63	0.76	0.75
5	0.56	0.626	0.75	0.74

Crossing the models allowed a clear improvement of the predictions, in particular for the models with a larger number of classes such as the models with four and five classes which gained approximately 15% accuracy. After crossing only, the results seem similar with a slight advantage for the models with the "Virtual" class sets. The detailed results after crossing are presented in the form of a confusion matrix (Figure 7) with the prediction of the model on the x-axis and the student's real class on the y-axis. The diagonal therefore represents the good predictions. First, the results of the models with "Virtual" class sets.

For the classes of students who fail multiclass models, the worst precision is that of class in range [9.5, 10.5[of model 5 with 27% and the best is that of class in range [7.5, 10[of model 1 with 57%. Overall, the accuracy for classes of students who fail is about 40% and always over 92% for students who pass. Next, the results of the models with sets of "Balanced" classes (Figure 8).

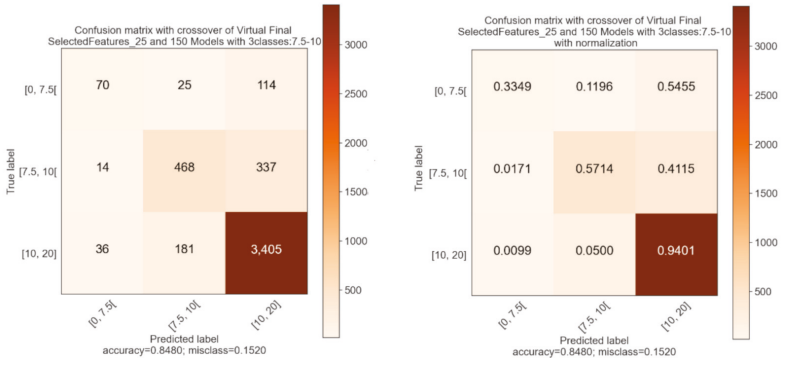

Figure 7. Tree "Virtual" confusion matrix model number 1 to 3 after crossing.

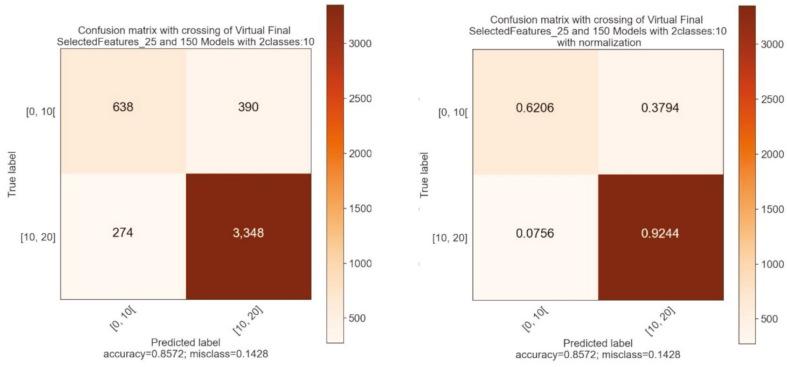

Figure 8. "Balanced" confusion matrix for two classes after crossing.

For the two-class model, we observe good precision for the pupils of class in range [10, 20]—more than 95% (Figure 9). However, only 52% of the pupils of class in range [0, 10[were correctly predicted, which is worse than the model with the previous balancing.

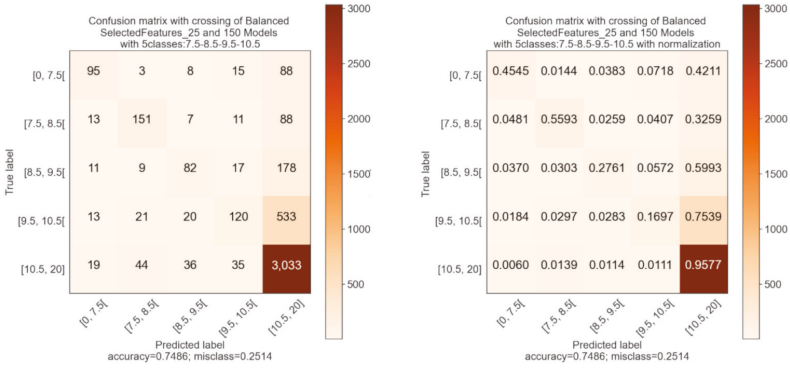

Figure 9. "Balanced" confusion matrix for five classes after crossing.

For the classes of students who fail multiclass models, the worst precision is that of class in range [9.5, 10.5[of model 5 with 17% and the best being that of class in range [7.5, 8.5[of the same model with 56% (Figure 9). Overall, the accuracy for classes of students who fail is around 40%.

4. Discussion

The confusion matrices show that increasing the number of classes results in a slight loss of precision. The decrease in precision between models with the sets of classes 0,1,2 and 3,4,5 can be explained by the fact that using class [10, 20] offers better precision compared to a model using the class in range [10.5, 20] (Figure 10).

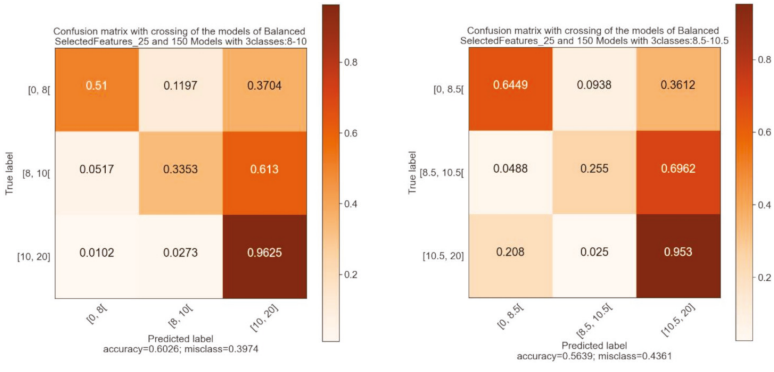

Figure 10. Results of a model with the class in range [10, 20] and of a model with the class in range [10.5, 20].

The quality of the prediction of the classes of students who fail is still poor—less than one pupil in two belonging to these classes is correctly identified.

The balancing which seems the most convincing is balancing with the weight of the data because even if there is low precision for the classes contiguous to the classes in ranges [10, 20] and [10.5, 20], the precision of the classes to students with the lowest averages is better. This model is therefore more useful for predicting students who will largely fail than those with an average of around 10/20. The first results were not conclusive. However, thanks to improvements in the data, the increase in features followed by a selection of these led to a gain in model accuracy.

Secondly, balancing the classes of the models allowed for smoothing the precision. That is to say, a gain of precision on the classes with a small amount of data representing them.

Finally, the intersection of the model predictions made it possible to obtain more satisfactory results. However, they are still insufficient because the quality of the prediction of pupils having difficulty is not important enough. However, the models can still be improved.

In order to improve the results, we can imagine a second level of crossing predictions between the two kinds of balancing. Improving the RF algorithm, we use is also possible, for example, trying to limit the training so as not to be overfitting as suggested in Figure 11.

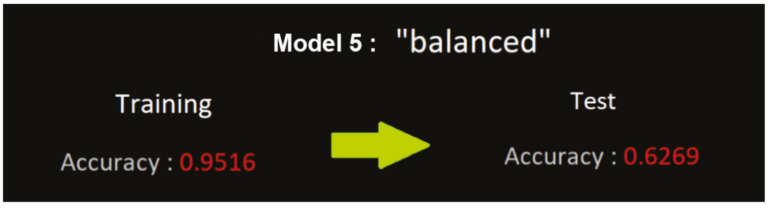

Figure 11. Accuracy in training and testing.

5. Conclusions

In this article, we addressed the problem of success in the first year of university by constructing a solution to estimate whether a candidate entering a training course will

succeed or not. This additional information can then be used by guidance counselors to guide the future student in their choices and by pedagogical teams to set up specific support measures for each candidate, even before the first evaluations. In order to obtain this information, we used a set of real data that we had previously cleaned and then processed using the RF algorithm. To improve the results, we also implemented a "model crossing" mechanism.

In addition, it should not be forgotten that the results are based on data from first-year students at a Technical University Institute in Computer Science. These have been pre-selected and the size of each promotion can be considered as average, around one hundred. Therefore, wanting to reproduce the process presented to other training is quite possible and relevant. However, the results will probably not be the same since the number of students differs according to the type of training. This can be considered very important like biology or medical training where it can exceed a thousand each year, or low like graphic design schools, for example, Itecom Art Design with around twenty a year. With enrollment differences, the pre-selection of students, as well as the different modules from one training to another, can have an impact on the results.

We are currently working on different solutions to consolidate and improve the results obtained. In addition to the evaluation of different data sources (promotions), we are investigating the possibility of using other algorithms instead of RF and even combining the different algorithms. Finally, we are also starting work to take into consideration not only numerical data characterizing the candidates but also textual data found in class reports.

Author Contributions: Conceptualization: G.R. and C.D.-P.; methodology: G.R. and C.D.-P.; software: M.B. and S.D.; supervision, G.R. and C.D.-P.; project administration, C.D.-P.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institut Universitaire Technologique of Université Côte d'Azur.

Informed Consent Statement: The patient's consent was revoked because all names and identifying information were removed by persons outside the project. The people who worked on the data were therefore never aware of the source of the data.

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

References

1. Hussain, M.; Zhu, W.; Zhang, W.; Abidi, S.M.R.; Ali, S. Using machine learning to predict student difficulties from learning session data. *Artif. Intell. Rev.* **2018**, *52*, 381–407. [CrossRef]
2. Wakelam, E.; Jefferies, A.; Davey, N.; Sun, Y. The potential for student performance prediction in small cohorts with minimal available attributes. *Br. J. Educ. Technol.* **2020**, *51*, 347–370. [CrossRef]
3. Hasan, R.; Palaniappan, S.; Raziff, A.R.A.; Mahmood, S.; Sarker, K.U. Student Academic Performance Prediction by using Decision Tree Algorithm. In Proceedings of the 4th International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, Malaysia, 13–14 August 2018; pp. 1–5.
4. Kostopoulos, G.; Kotsiantis, S.B.; Pintelas, P.E. Estimating student dropout in distance higher education using semi-supervised techniques. In Proceedings of the 19th Panhellenic Conference on Informatics, Athens, Greece, 1–3 October 2015; pp. 38–43.
5. Zhou, Z.-H.; Li, M. Tri-training: Exploiting unlabeled data using three classifiers. *IEEE Trans. Knowl. Data Eng.* **2005**, *17*, 1529–1541. [CrossRef]
6. Pojon, M. Using Machine Learning to Predict Student Performance. Master's Thesis, Faculty of Natural Sciences, University of Tampere, Tampere, Finland, June 2017.
7. Acharya, A.; Sinha, D. Early Prediction of Students Performance using Machine Learning Techniques. *Int. J. Comput. Appl.* **2014**, *107*, 37–43. [CrossRef]
8. Imran, M.; Latif, S.; Mehmood, D.; Shah, M.S. Student Academic Performance Prediction using Supervised Learning Techniques. *Int. J. Emerg. Technol. Learn.* **2019**, *14*, 92–104. [CrossRef]
9. Sultana, J.; Usha Rani, M.; Farquad, H. Student's Performance Prediction using Deep Learning and Data Mining Methods. *Int. J. Recent Technol. Eng.* **2019**, *8*, 1S4.

10. Shetty, I.D.; Shetty, D.; Roundhal, S. Student Performance Prediction. *Int. J. Comput. Appl. Technol. Res.* **2019**, *8*, 157–160. [CrossRef]
11. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
12. Halim, A.A.; Dartigues-Pallez, C.; Precioso, F.; Riveill, M.; Benslimane, A.; Ghoneim, S.A. Human action recognition based on 3d skeleton part-based pose estimation and temporal multi-resolution analysis. In Proceedings of the IEEE International Conference on Image Processing (ICIP 2016), Phoenix, AZ, USA, 25–28 September 2016; pp. 3041–3045.
13. Gioanni, L.; Dartigues-Pallez, C.; Lavirotte, S.; Tigli, J.-Y. Using random forest for opportunistic human activity recognition: A complete study on opportunity dataset. In Proceedings of the 11èmes Journées Francophones Mobilité et Ubiquité (Ubimob 2016), Lorient, France, 5 July 2016.
14. Bouaziz, A.; da Costa Pereira, C.; Dartigues Pallez, C.; Precioso, F. Interactive generic learning method (IGLM): A new approach to interactive short text classification. In Proceedings of the SAC '16 31st Annual ACM Symposium on Applied Computing, Pisa, Italy, 4–8 April 2016; pp. 847–852.
15. Fernandez-Delgado, M.; Cernadas, E.; Barro, S.; Amorim, D. Do we need hundreds of classifiers to solve real world classification problems? *J. Mach. Learn. Res.* **2014**, *15*, 3133–3181.
16. Amine, S.; Dartigues-Pallez, C.; Gaetan, R. *Classification Supervisée de Données Pédagogiques pour la Réussite dans L'enseignement Supérieur*; Master's Thesis, hal-02486729; I3S Université Côte d'Azur: Sophia-Antipolis, France, February 2020.
17. Nguyen, C.; Wang, Y.; Nguyen, H.N. Random forest classifier combined with feature selection for breast cancer diagnosis and prognostic. *J. Biomed. Sci. Eng.* **2013**, *6*, 551–560. [CrossRef]
18. Saeys, Y.; Abeel, T.; Van De Peer, Y. Robust Feature Selection Using Ensemble Feature Selection Techniques. In *Machine Learning and Knowledge Discovery in Databases*; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2008; Volume 5212, pp. 313–325.

education sciences

MDPI

Article

Evaluating Learning Space Designs for Flipped and Collaborative Learning: A Transactional Distance Approach

William Swart * and Ken MacLeod

Department of Marketing and Supply Chain Management, College of Business, East Carolina University, Greenville, NC 27858, USA; MACLEODK@ecu.edu
* Correspondence: SWARTW@ecu.edu

Abstract: Problem-based learning is the latest name for a teaching philosophy that is as old as Ancient Greece. Whether you call it Socratic Inquiry, case-based teaching, problem-based learning, interactive group learning, or "flipped" learning, the essential concept is to encourage the student to collaborate in applying their gained knowledge to solve a problem. As traditional lecture-based teaching has been challenged, the design of classrooms has been called into question. A flat or tiered room is not seen as an ideal setting for collaborative work. In our own College of Business, several traditional classrooms were converted to problem-based learning classrooms at considerable expense. This paper evaluates, using measures based on Michael G. Moore's theory of transactional distance, whether moving flipped classes into these high-tech classrooms improves the collaborative learning experience. Transactional distance can be defined as the barriers that exist to a student's engagement with their learning experience. These barriers arise due to the interaction between students and the teacher, other students, the subject matter content, and instructional technology being used. Our results suggest that, from a student engagement and outcome standpoint, the investment in costly high-tech classrooms is not warranted—a welcome result in times when university budgets are stretched to the limit.

Keywords: learning space design; transactional distance; student engagement; learning outcomes; collaborative learning; blended learning; problem-based learning; interactive group learning; flipped classroom

Citation: Swart, W.; MacLeod, K. Evaluating Learning Space Designs for Flipped and Collaborative Learning: A Transactional Distance Approach. *Educ. Sci.* **2021**, *11*, 292. https://doi.org/10.3390/educsci11060292

Academic Editors: Sandra Raquel Gonçalves Fernandes, Marta Abelha and Ana Teresa Ferreira-Oliveira

Received: 7 May 2021
Accepted: 9 June 2021
Published: 12 June 2021

1. Introduction

As instructors, we are assigned classrooms. We can make requests if we have enough seniority, but even then, there are no guarantees. We all want to get through to our students, and we all want the best environment for doing so. At one time, that environment was the porch of a Greek market where the Stoics met. Later, a one-room schoolhouse gave most students all the education they would receive, while higher education evolved from students gathering in their professor's office to listen and talk with them. Eventually, when the office could no longer accommodate the increased number of students, classrooms were built, giving students a place to sit, listen, and take notes. Figure 1 depicts such a traditional classroom.

Before the days of the printing press, books were generally unavailable. Those that possessed them would read them to those that did not have them (or could not read). As these "lecturers" achieved a reputation, it became necessary (and profitable) to extend their reach as far as possible. This was achieved by following the model of the Greek theatron, the stadium-like seating around the stage. Figure 2 illustrates how this model of conveying a lecture to an audience has persisted over a millennium despite books and a wide range of instructional technology being widely available and affordable.

Figure 1. A traditional classroom.

Figure 2. Tiered classrooms then and now.

Stadium-like classrooms are great for lecturing, but lecturing is not necessarily good for learning. An article in BBC NEWS [1] asked whether lectures should not be obsolete by now, quoting research showing that students remember as little as 10% of their lecture just days afterward and referencing a Harvard study that found, on average, attendance at lectures falls from 79% at the start of the term to 43% at the end.

Prince [2] found that problem-based learning (PBL) can be superior to traditional lecture-based learning. Bishop and Verleger [3] report similar results for the flipped classroom, a particular type of PBL, which they define as an educational technique that consists of two parts: interactive group learning activities mediated by the instructor in the classroom and direct computer-based individual instruction outside the classroom.

Selingo [4] proclaimed flipped learning as the future of higher education. This proclamation appears to have been a prophesy considering the exponential growth of research about the flipped classroom, shown in Figure 3 from Talbot [5]. This explosion in knowledge about flipped learning has encouraged its adoption in the classroom. According to the Flipped Learning Global Initiative, by 2017, approximately 16% of U.S. teachers were flipping their classes, 35% wanted training on how to flip their classes, and 46% of principals wanted new teachers who knew how to flip a class [6]. As can be expected, adoption in the classroom has created a market for flipped products (software, hardware and services) which was valued at $971 in 2018 and is forecasted to grow to $1.9 Billion by 2024—a compound annual growth rate of 15.5% [7].

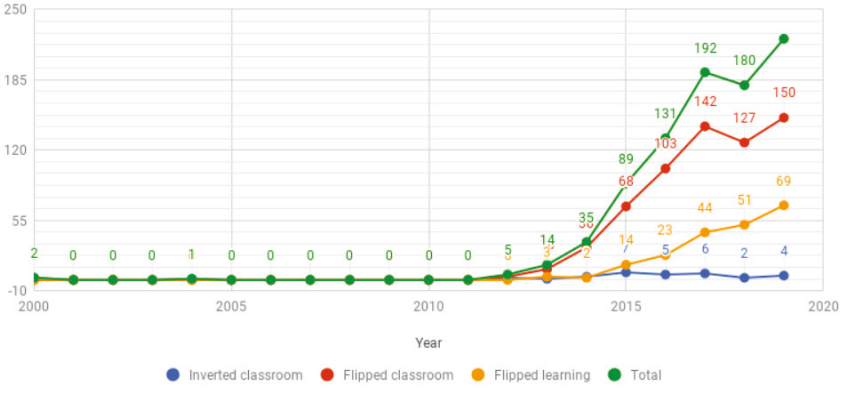

Figure 3. Published research on flipped learning (# of peer-reviewed articles).

Based on the theory that a new type of learning will require a new setting, in 2013, the then Dean of our College of Business initiated a remodeling program to convert existing classrooms to specifically support PBL, at a cost of approximately US $190,000 per classroom. The rooms are showpieces, with a very large (120″) screen at the front and from 5 to 8 group stations around the room, each consisting of tables (on wheels) with seating for six, a computer with internet access, and a large (90″) screen, where the students can project their work in process. Simultaneously, to encourage faculty to consider PBL, the Dean solicited proposals from the faculty for implementing it in their classes.

The authors proposal to implement flipped learning in a required undergraduate course in the Bachelor of Science in Business Administration (BSBA) program in the PBL classroom was accepted. With their course scheduled in the new PBL room, when they asked the Dean what he expected to see in the class, the response was "I do not want to see lecturing. I want to see student working in teams collaborating with each other, sweating to solve a complex problem." His vision aligned directly with the flipped classroom. Figure 4 shows a flipped class in action in one of the PBL classrooms.

Figure 4. A flipped class in action in a PBL room.

The authors experienced success teaching their course as a flipped class in the PBL classroom. The average final grades increased by approximately 15 points and the range between the highest and lowest average final grade in the class decreased by over 45 points [8]. Consequently, they adopted flipped learning in all their classes. Due to the limited number of PBL classrooms, however, their classes could not always be scheduled in PBL rooms, so they have taught their classes in the PBL rooms when fortunate enough to have been assigned to one and in flat or stadium classrooms when less fortunate. When not in a PBL room, they improvised to adapt the classroom to help, not interfere, with the collaborative work required of their students. For the flat classroom, this was simply to have the students drag desks together in groups of four to six and to share their work on their laptops. For the stadium classroom, the desktops and chairs are fastened to the

floor, so the groups sat together, either strung out along one row or in two rows, with the students in front turning around to talk to the ones behind, again sharing work on the laptops.

Regardless of the classroom type assigned, all classes were flipped in the manner described in Swart [9]. The school's Learning Management System contained the lecture material (notes and videos) the students read and watched outside of class. During class, the lecture material is applied to an assigned problem or task, upon completion of which an individual quiz assesses how well the material was mastered. The instructor, instead of lecturing, moves from group to group, answering questions and providing guidance. Swart [9] has shown that student outcomes and satisfaction are better with the flipped paradigm. We had not observed any change in grades, however, regardless of the type of classroom assigned. This led us to the overarching research question of what benefits in student learning are obtained from the monies that are being spent on creating PBL rooms?

In the next subsections, we will review pertinent literature to our research covering both learning spaces and transactional distances. The latter provides the theoretical foundations for the scale of transactional distance, which will be our measurement tool for the research.

1.1. Learning Spaces

Problem-based, or active learning, classrooms designed to facilitate collaborative work are becoming more common [10–12]. Some research [13,14] has shown that students do not like collaborative learning, while others [8,15] have found no difference or improved student performance in collaborative settings. Student preference for entertaining and interactive learning environments was noted by Leverett et al. [16]. Enhancements to learning spaces has been a topic for numerous authors [17–20]. In [13], the authors noted the need for inquiry into the relationship between active learning classrooms and student attitudes toward collaborative learning.

Discussion of learning spaces goes back before any talk of collaborative learning. By 1979, Weinstein [21] found enough research to provide a review concerning the physical environment of schools. The 2006 book *Learning Spaces* [22] provides a series of articles discussing how to "reconceptualize learning spaces to facilitate active, social, and experiential learning." Clinton and Wilson [23] note that classrooms were designed to facilitate lectures, but Peberdy [24] notes many institutions are redesigning classrooms to facilitate active learning, despite the substantial costs. While students indicate that collaborative learning classrooms help their concentration [25], Cleveland and Fisher [26] call for better understanding of student opinions of their learning spaces. Clinton and Wilson [23] provide a small study using their own survey of student perceptions of their course taught in both a traditional setting and a collaborative setting, concluding that "students perceived active-learning classrooms as better suited for collaborative learning than traditional classrooms . . . ". Swart [9] found that student performance and satisfaction improved in flipped classes in the PBL classrooms compared to lecture-based classes in traditional classrooms.

1.2. Transactional Distance

The theory of transactional distance was developed by Michael G. Moore [27] and is one of the seminal theories governing distance education. It constituted a paradigm shift in that it viewed "distance" as a social and communications gap, a space of potential misunderstandings between the instructor and learners. He proposed that transactional distance is a function of two sets of variables, dialogue (interaction) and structure (course design). High structure and low dialogue yield high transactional distance and vice versa.

Zhang [28] updated Moore's theory to include online learning. She posited that transactional distance (TD) represented the barriers existing between students and their full engagement with the learning environment. She defined it as a multidimensional construct consisting of: transactional distance between student and student (TDSS), transactional distance between student and teacher (TDST), transactional distance between student

and content (TDSC), and transactional distance between student and the instructional TECHnology used for teaching/learning (TDSTECH). She then postulated that these constructs resulted in three learning outcomes: student SATISFACTION with learning, PROGRESS toward achieving learning goals, and whether LEARNING has occurred in the class.

Zhang operationalized the above concepts by breaking down each construct into multiple elements, resulting in 31 questions plus the three outcomes, shown in Appendix A. Each element is associated with a five-point Likert scale (1 = Untrue, 5 = True). The resulting questionnaire, referred to as the scale of transactional distance, was subjected to structural equation modelling techniques including exploratory and confirmatory factor analysis, confirming goodness of fit and validation of the scale [28].

Figure 5 summarizes the content of the scale of transactional distance, detailed in Appendix A. Student engagement is obtained from the responses to the element questions under TDSI, TDSC, TDSS, and TDSTECH, while outcomes are measured from the responses for LEARNING, GOALS, and SATISFACTION. All the responses are recorded on the same survey.

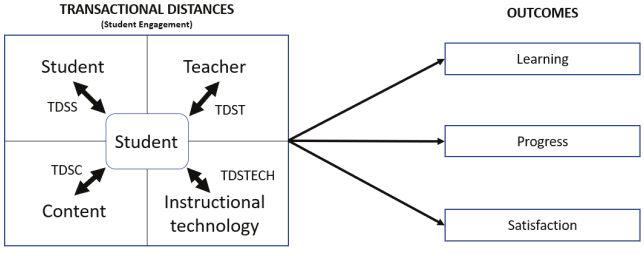

Figure 5. Content of the scale of transactional distance.

1.3. Research Hypotheses

Swart [9] achieved a 15 point (out of 100) increase in final average grade when changing from a lecture to a flipped format. However, the authors did not find a difference in grades when the flipped course was taught in either a stadium, flat of PBL classroom. Grades are important to students. However, as instructors, Schwab et al. [29] note that we are more interested in student learning. While we already knew that student grades in our flipped classes were not impacted by classroom type, we did want to find out whether classroom type impacted student engagement, which is "highly correlated with many desirable learning and personal developmental outcomes of college", according to Axelson and Flick [30]. Thus, from the data obtained from administering the scale of transactional distance survey, we tested the following four research hypotheses:

Research Hypothesis 1

Hypothesis $H1_0$. *Student engagement and outcomes in a flat classroom = student engagement and outcomes in a PBL classroom.*

Hypothesis $H1_a$. *Student engagement and outcomes in a flat classroom ≠ student engagement and outcomes in a PBL classroom.*

Independent-samples t-tests will be used to test these hypotheses. SPSS 27 will provide the results for both engagement and outcome, hence the reason for including both in the hypothesis.

Research Hypothesis 2

Hypothesis $H2_0$. *Student engagement and outcomes in a PBL classroom = student engagement and outcomes in a stadium classroom.*

Hypothesis H2$_a$. *Student engagement and outcomes in a PBL classroom ≠ student engagement and outcomes in a stadium classroom.*

Independent-samples t*-tests will be used to test these hypotheses. SPSS 27 will provide the results for both engagement and outcome, hence the reason for including both in the hypothesis.*

Research Hypothesis 3

Hypothesis H3$_0$. *Student engagement and outcomes in a stadium classroom = student engagement and outcomes in a flat classroom.*

Hypothesis H3$_a$. *Student engagement and outcomes in a stadium classroom ≠ student engagement and outcomes in a flat classroom.*

Independent-samples t*-tests will be used to test these hypotheses. SPSS 27 will provide the results for both engagement and outcome, hence the reason for including both in the hypothesis.*

Research Hypothesis 4

Hypothesis H4$_0$. *Student engagement and classroom type are unique significant predictors of outcomes.*

Hypothesis H4$_a$. *Student engagement and classroom type are NOT unique significant predictors of outcomes.*

Stepwise multiple regression will be used to test Research Hypothesis 4.

Hypotheses 1–3 refer to results obtained after the fact, when PBL rooms were already built. We envision the possibility of using the scale of transactional distance to evaluate alternate PBL room design using user focus groups. Focus groups could be presented with renderings of actual design alternatives, including virtual reality representations of the physical PBL room design alternatives. They could then fill out the scale of transactional distance survey indicating how they would respond to the questions for each design alternative. The results could then be entered into a robust regression model to predict student outcomes for each design alternative.

2. Materials and Methods

The data was gathered over a five-semester period (Fall 2017–Spring 2019) in a required undergraduate class in the Bachelor of Business Administration (BSBA) program. The instructor was the same for all classes and taught the class in a flipped format as described in [9]. The classes were scheduled in either a PBL classroom, a flat classroom, or a stadium classroom, based on availability and administration priorities.

The scale of transactional distance survey was used to collect student engagement and outcome data. Student engagement data was obtained from the transactional distance questions, while outcomes data was obtained from Zhang's three outcome questions. Table 1 indicates the total enrollments in each of the classroom types during the five semesters as well as the response rates obtained.

Table 1. Enrollment and response information by classroom type.

	PBL	FLAT	STADIUM
Responses	33	76	48
Enrollment	37	97	58
Response %	89%	78%	83%

3. Results

Several researchers have shown that the three outcome measures are significantly correlated [24,31,32]. We obtained the bivariate correlations to determine whether this was the case for our data as well. The results are presented in Table 2.

Table 2. Bivariate correlations for Zhang's learning outcomes.

		LEARNING	PROGRESS	SATISFACTION
LEARING	Pearson Correlation	1	0.734 **	0.816 **
	Sig. (2-tailed)		0.000	0.000
	N	157	157	157
PROGRESS	Pearson Correlation	0.734 **	1	0.697 **
	Sig. (2-tailed)	0.000		0.000
	N	157	157	157
SATISFACTION	Pearson Correlation	0.816 **	0.697 **	1
	Sig. (2-tailed)	0.000	0.000	
	N	157	157	157

**: Correlation is significant at the 0.01 level (2-tailed).

The bivariate correlations shown in Table 2 are statistically significant ($p < 0.001$) and their Pearson Correlation Coefficients are 0.695 or higher. As done in several other studies [23,24,30], we have elected SATISFACTION as the representative outcome in the rest of our analyses.

Table 3 exhibits the group statistics for the data obtained from the scale of transactional distance surveys.

Table 3. Group statistics for the scale of transactional distance data.

	PBL		FLAT		STADIUM	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
TDSS	4.313	0.627	4.134	0.690	3.688	1.137
TDSC	3.914	0.581	3.958	0.720	3.778	0.896
TDST	4.350	0.674	4.406	0.583	4.172	0.836
TDSTECH	3.841	0.579	3.814	0.638	3.703	0.738
SATISFACTION	4.303	1.045	4.316	1.086	3.813	1.315
N	33		76		48	

Figure 6 shows a radar graph comparing the means of the engagement factors and the outcome variable SATISFACTION for the three types of classrooms tested. The vertical axis shows the outcome variable SATISFACTION and each of the other four axes represent an engagement factor. The Likert scale in the scale of transactional distance was defined so that the greater the transactional distance, the better the engagement and outcomes. Although opposite to the traditional definition that a "big" distance is bad, we found that respondents were less confused filling out a survey in which big was associated with better. Thus, it appears as though the data indicates that the stadium classroom is not as good as either of the other two on all student engagement factors or outcome. To draw any conclusions from this data, we must test the statistical significance of the differences shown on the radar diagram.

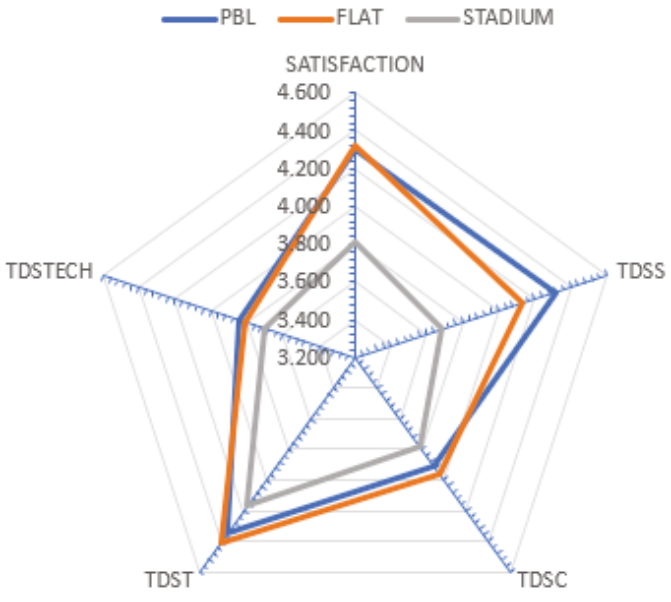

Figure 6. Radar diagram of mean SATISFACTION and engagement factors by classroom type.

3.1. Testing Research Hypothesis 1

Figure 7 shows pictures of the actual PBL and flat classroom used by our students when they completed the scale of transactional distance survey.

Figure 7. Pictures of Flat and PBL classrooms used in this research.

The results of analyzing that data with an independent-samples *t*-test are given in Table 4. Levene's Test for Equality of Variances was not rejected for any of the engagement factors or SATISFACTION, so the significance levels for the tests are drawn from the "Equal variances assumed" rows of the table. The results of the *t*-test for equality of means support $H1_0$, indicating that there was no reason to believe that there were any statistically significant differences in student engagement (transactional distances) or SATISFACTION between the PBL and the flat classroom.

Table 4. Independent-samples *t*-test for PBL vs. flat classroom.

		Levene's Test for Equality of Variances		*t*-Test for Equality of Means				
		F	**Sig.**	***t***	**df**	**Sig. (2-Tailed)**	**Mean Difference**	**Std. Error Difference**
TDSS	Equal variances assumed	0.034	0.854	1.281	107	0.203	0.179	0.140
TDSC	Equal variances assumed	0.762	0.385	–0.311	107	0.756	–0.044	0.142
TDST	Equal variances assumed	0.321	0.572	–0.444	107	0.658	–0.057	0.127
TDSTECH	Equal variances assumed	0.005	0.942	0.207	107	0.837	0.027	0.129
SATISFACTION	Equal variances assumed	0.022	0.883	–0.057	107	0.955	–0.013	0.224

3.2. Testing Research Hypothesis 2

Figure 8 shows pictures of the actual stadium and PBL classrooms used by our students when they completed the scale of transactional distance survey.

Figure 8. Pictures of stadium and PBL classrooms used in this research.

The results of analyzing that data with an independent-samples *t*-test are given in Table 5. The result of Levene's Test for Equality of variances indicated that equality of variances could not be assumed for TDSS, TDSC, and TDSTECH. Thus, the significance levels of the tests for these factors had to be obtained from the "Equal variances not assumed" rows of the table.

The results of the *t*-tests indicated that there was a significant difference between the transactional distance between student and student in the two rooms ($p < 0.002$). There was also a marginal statistical difference in SATISFACTION between the two rooms ($p < 0.078$). Thus, $H2_0$ is not supported for TDST or for SATISFACTION at a level of significance of 0.078.

Table 5. Independent-samples *t*-test for PBL vs. stadium classrooms.

		Levene's Test for Equality of Variances		*t*-Test for Equality of Means				
		F	**Sig.**	***t***	**df**	**Sig. (2-tailed)**	**Mean Difference**	**Std. Error Difference**
TDSS	Equal variances assumed	12.685	0.001	2.871	79	0.005	0.626	0.218
	Equal variances not assumed			3.174	75.967	0.002	0.626	0.197
TDSC	Equal variances assumed	7.371	0.008	0.766	79	0.446	0.136	0.177
	Equal variances not assumed			0.827	78.780	0.411	0.136	0.164
TDST	Equal variances assumed	2.535	0.115	1.013	79	0.314	0.178	0.175
TDSTECH	Equal variances assumed	5.555	0.021	0.898	79	0.372	0.138	0.153
	Equal variances not assumed			0.939	77.549	0.350	0.138	0.147
SATISFACTION	Equal variances assumed	3.213	0.077	1.788	79	0.078	0.491	0.274

3.3. Testing Research Hypotheses 3

Figure 9 shows pictures of the actual stadium and flat classrooms used by our students when they completed the scale of transactional distance survey.

Figure 9. Pictures of flat and stadium classrooms used in this research.

The results of analyzing that data with an independent-samples *t*-test are given in Table 6. The result of Levene's Test for Equality of Variances indicated that equality of variances could not be assumed for any of the factors or the outcome data. Thus, the significance levels of the tests for these all these had to be obtained from the "Equal variances not assumed" rows of the table.

The results of the *t*-tests indicated that there was a significant difference between the transactional distance between student and student in the two rooms ($p < 0.017$) as well as between SATISFACTION between the two rooms ($p < 0.029$). There also was a marginal statistical difference in TDST between the two rooms. Thus, $H2_0$ is not supported for TDSS and SATISFACTION nor for TDST at a significance level of 0.094.

Table 6. Independent-samples *t*-test for flat vs. stadium classrooms.

		Levene's Test for Equality of Variances		*t*-Test for Equality of Means				
		F	**Sig.**	***t***	**df**	**Sig. (2-tailed)**	**Mean Difference**	**Std. Error Difference**
TDSS	Equal variances assumed	18.323	0.000	2.723	122	0.007	0.446	0.164
	Equal variances not assumed			2.450	69.043	0.017	0.446	0.182
TDSC	Equal variances assumed	4.673	0.033	1.231	122	0.221	0.180	0.146
	Equal variances not assumed			1.172	84.378	0.244	0.180	0.153
TDST	Equal variances assumed	7.933	0.006	1.837	122	0.069	0.234	0.127
	Equal variances not assumed			1.697	75.795	0.094	0.234	0.138
TDSTECH	Equal variances assumed	5.469	0.021	0.888	122	0.376	0.111	0.125
	Equal variances not assumed			0.859	89.328	0.393	0.111	0.129
SATISFACTION	Equal variances assumed	4.001	0.048	2.314	122	0.022	0.503	0.217
	Equal variances not assumed			2.217	86.161	0.029	0.503	0.227

3.4. Research Hypothesis 4

Stepwise multiple regression was used to test Research Hypothesis 4. The dependent variable was SATISFACTION, and the independent variables were TDSS, TDSC, TDST, TDSTECH, STADIUM and FLAT. The last two are indicator variables to account for classroom type. Note that PBL is the base case so that when the two indicator variables are zero, the data correspond to teaching in a PBL room.

Table 7 exhibits the results of the stepwise multiple regression. The robust regression equation is given by:

$$\text{SATISFACTION} = 0.601 * \text{TDSC} + 0.452 * \text{TDSS} + 0.285 * \text{TDST}$$

Table 7. Stepwise multiple regression results with SATISFACTION as the dependent variable.

	Step 1				Step 2				Step 3				
	b	***Std. error***	***B***	***t***	***b***	***Std. error***	***B***	***t***	***b***	***Std. error***	***B***	***t***	**VIF**
TDSC	1.095	0.088	0.705	12.388 *	0.653	0.118	0.420	5.539 *	0.601	0.117	0.387	5.153 *	2.129
TDSS					0.531	0.102	0.396	5.214 *	0.452	0.104	0.336	4.357 *	2.245
TDST									0.285	0.101	0.168	2.809 **	1.355
R	0.705				0.757				0.771				
R^2	0.498				0.573				0.594				
F	153.485 *				103.305 *				74.180 *				
df-regression	1				2				3				
df-residual	155				154				153				

Note: * $p < 0.000$; ** $p < 0.006$.

The predictor variables are significant as indicated by the note at the bottom of the table. The regression equation accounts for 59.4% of the variance and is significant. The VIF values give no indication of multicollinearity. Classroom type is not a significant predictor of SATISFACTION and, because of their significant bivariate correlations with each other, neither are LEARNING or PROGRESS. Thus, $H4_0$ is not supported.

4. Discussion

Table 8 summarizes the significant results from testing research hypotheses 1–3 and indicates how the three classroom types that were used for flipped learning compared in terms of student engagement (measured by the four transactional distance factors) and outcomes.

Table 8. Summary significance levels for hypotheses tests.

		PBL and Flat	Stadium and PBL	Stadium and Flat
		RH 1	**RH 2**	**RH 3**
Engagement	TDSS	-	0.002	0.017
	TDSC	-	-	-
	TDST	-	-	0.094
	TDSTECH	-	-	-
Outcomes	SATISFACTION	-	0.078	0.029

Note: $p > 0.1000$ is indicated by "-".

To make sense of this data, we have copied the measurements on the TDSS, TDST, and satisfaction axes, showing the actual transactional distance scores, from the radar graph (Figure 5) and shown them in Figure 10. TDSC (student-to-content) and TDSTECH (student-to-technology) were not included in Figure 10 since they had no significant differences for any of the comparisons. For TDSC, this was not surprising since the content is delivered identically, via the LMS, to all students who must learn it outside of class, as required for flipped classes. We were surprised that the additional technology (TDSTECH) of the PBL room seems not to have mattered to the students: file-sharing is just easy on their portable computing devices, no matter what the classroom type.

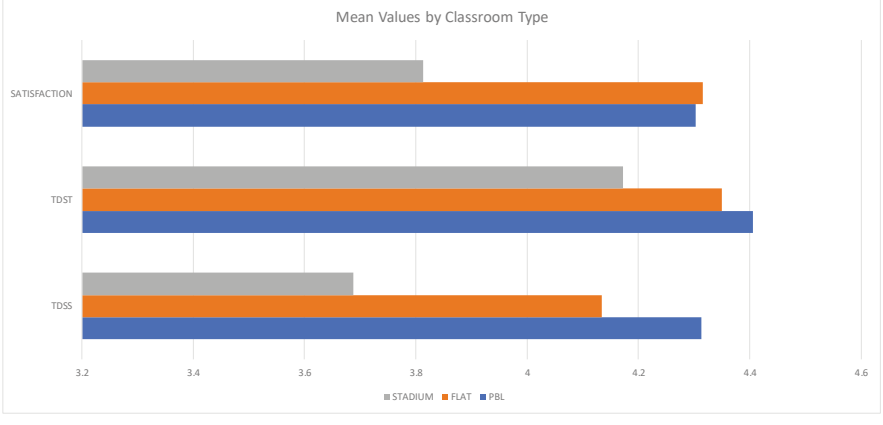

Figure 10. Comparison of classroom layouts.

Regarding Figure 10, beginning with SATISFACTION, both the flat and the PBL rooms were statistically different from the stadium room. On the graph, the PBL and flat bars both extend to the right, which is why they have no significant differences between them, while both show significant differences from the stadium classroom. If this were a lecture-based class, this result would have been surprising, since a stadium (or tiered) classroom is designed for lecturing. As previously noted, a stadium classroom is not likely the best design for a "flipped" class, and this appears to be reflected in the SATIFACTION data. The fixed desktops and chairs were simply awkward for the students working in

groups and equally awkward for the instructor when trying to squeeze between the rows to work with the groups, explaining details of the material or helping them get back on track to solve the problem. That kind of interaction is exactly what the PBL classroom was designed to facilitate, so it is rather surprising that the flat classroom garnered just as good SATISFACTION numbers. Provided the students were able to work together effectively, they appear to be equally satisfied whether in the high-tech PBL room or simply rearranging the desks in a flat classroom. As noted earlier, the scale of transactional distance has three auto-correlated outcomes represented by SATISFACTION, so similar result would have been obtained if LEARNING or PROGRESS had been used instead.

Considering student engagement with other students, both PBL and flat classrooms had statistically better (0.002 for PBL and 0.017 for flat) results for student-to-student (TDSS) interactions than the stadium classroom. Again, this can be attributed to the classes being flipped. In flipped classes, class time is devoted to interactive group learning, requiring real time, face-to-face communication. The tiered classroom simply does not meet the needs of the students, presenting a barrier to their ability to interact with one another, as shown in the transactional distance of the PBL and flat TDSS (4.313 and 4.134, respectively) scores from the stadium TDSS score (3.688). More surprising was the lack of significant difference between the PBL and flat classrooms. It would seem that the technology and pod arrangement were not as important as simply the ability to see one another (when the desks were rearranged into groups) and the ability to share files on their laptops. With the focus of the classroom time on solving a problem, the special advantages of the PBL room may have been lost in the student's concentration on the work to be done.

The final statistically different result was for the engagement of students with the instructor (TDST), where the flat classroom is significantly greater ($p < 0.094$) than the stadium classroom. Flipped classes require that the instructor be a learning coach and consultant to teams and individuals. This requires the instructor to move from one group to another. Such movement is easier in the flat room because the instructor can walk from one team to another without physical barriers. In a stadium room student teams are at various levels and it is awkward for the instructor to move from one level to the next because it requires exiting one row, moving to a different tier, and then moving across the row where the team requiring attention is located. Further, the instructor is either standing before a group of six students spread out in a row or standing in the middle of the group split between two rows. In either setting, approximately half of the group would have a difficult time seeing what the instructor was demonstrating.

Given that flat classrooms were statistically superior to stadium classrooms on the TDST score, it was almost unbelievable that the PBL classrooms were not. In both cases, the instructor stands at one edge of the group, so it would seem, at first glance, that the two were identical, but the students perceived a difference. Somehow, the physical barriers of the stadium layout equate to other barriers in the PBL layout, perhaps that the students face a wall (screen) in the PBL room rather than the instructor. In the PBL rooms, the groups are located along the walls, with an open area in the center of the room. In a flat classroom, the students invariably grouped the desks toward the center of the room rather than along the walls, so the instructor could literally turn around and be talking to another group. For that matter, it was not unusual for one group to "listen in" while the instructor was talking to a near-by group, thus gaining information before the instructor addressed that group. Perhaps it was simply the informality of the hastily rearranged flat classroom that made the students more comfortable, but clearly the PBL rooms gained no advantage over the flat classrooms in terms of students interacting to solve a problem.

The most unexpected result of our research is that there was no significant difference in student engagement or outcomes between the PBL and flat rooms. An entire book was devoted to reconceptualizing learning spaces to facilitate active, social, and experiential learning [16] and many institutions are redesigning classrooms to facilitate active learning despite substantial costs [14]. Yet, we found no evidence beyond this research that anyone has made such a decision based on data that proves that the investment in such classrooms

can be justified in terms of improving student engagement or outcomes. Our results indicate that providing students with the ability to rearrange a flat classroom to face each other provided similar engagement and satisfaction as an expensive classroom reconfiguration. In short, while the technology (and luxury) of PBL classrooms is impressive to donors, visitors, and prospective students and their parents, students in this study find that they can do just as well without it.

Another indirect measure of student satisfaction is their willingness to recommend the course to other students. That question was included in the questionnaire and Table 9 shows both the number of students and percentage for the three classrooms. These results reflected our statistical results. Similar percentage of students would recommend the class taught in the PBL and flat classrooms to their friends while a much lower numerical percentage would recommend the same course taught in the stadium room to their friends.

Table 9. Responses to "Would you recommend this class to another student?"

	PBL		**FLAT**		**STADIUM**	
YES	28	85%	63	83%	30	63%
NO	5	15%	13	17%	18	38%
TOTAL	33		76		48	

With the decline in on-campus enrollments [33], campuses are under pressure to use every possible marketing strategy to attract students [34]. Undoubtedly, therefore, impressive physical facilities, including classrooms, will continue to be built. Ideally, such classrooms will not only involve form, but also function. Thus, the question becomes one of how to predict student satisfaction for a classroom *before* it is built, so that the investment decision can include costs as well as benefits/drawbacks to students. Research Hypothesis 4 was developed to answer this question and provided the robust prediction equation developed earlier:

$$\text{SATISFACTION} = 0.601 * \text{TDSC} + 0.452 * \text{TDSS} + 0.285 * \text{TDST}$$

It indicates that the classroom type is not a predictor of student satisfaction. What determines satisfaction is dictated by how well the classroom facilitates students learning the material (TDSC), by how well students can engage with each other in the classroom (TDSS), and by how well students can engage with the instructor in the classroom (TDSC). The fact that classroom type is not a predictor in the equation indicates that our data shows that neither classroom design has accomplished this any better than the others Thus, for this study, no classroom type is inherently better than another to produce student satisfaction (and the related outcomes).

We propose that the robust prediction model be used to evaluate alternate designs for future PBL room. Alternate PBL classroom design plans can be presented, one at a time, to a focus group consisting of classroom users. These alternate plans can be traditional plans and architectural renderings, but preferably they would consist of virtual reality models [25,35] that would allow users to experience the classroom as though it were already built. Based upon their experience and interaction with the PBL classroom design team, the focus groups members would be asked to complete the scale of transactional distance survey for each of the PBL design alternatives from which the TDSC, TDSS, and TDST could be calculated and used as input to the robust prediction model. This would yield a predicted measure of SATISFACTION for each design.

We recognize that many considerations must go into the selection of a best PBL classroom design. We posit that the impact of such a design on student engagement and satisfaction must be a primary consideration. Heretofore, it appears as though PBL rooms have been built on the presumption that they will improve student outcomes. This research presents a methodology that will allow the presumption to be tested before scarce resources are invested.

5. Conclusions

Our university, like several others, has invested hundreds of thousands of dollars in transforming traditional flat classrooms to PBL rooms and is currently gathering inputs on new technology and furniture requirements for consideration in the planning for the next generation of PBL rooms. Our results indicating that these expenditures may not have been necessary to achieve student engagement and outcome improvement are enlightening. However, as often is the case, aesthetics does not guarantee functionality.

We attribute some of our findings to the increased portability and capability of computing devices. Today's smart phones and tablets make it extremely easy for students to pass around their devices to classmates. When the first generation of PBL rooms were constructed, great attention was given to facilitating the sharing of information within and between groups. At the time, sharing of devices was cumbersome at best, hence relying on projecting information on large screens seemed the best alternative. Our results suggest that it may no longer be necessary to invest in multiple workstations, each having a networked PC and large screen for problem-based learning.

There is no doubt that the PBL rooms are a great marketing and advancement tool. Every open house conducted by our College and University conducts tours of campus and the PBL rooms have a great WOW! factor associated them. They impress students and parents and do influence and steer students to consider our university as their choice. These rooms are also a favorite stopping point for the Dean to extoll our commitment to student learning to potential donors. To that end, every door to the PBL room serves as a window and it is common to see the Dean showing a potential donor a PBL room in action. The challenge is to find problem-based learning rooms that will serve the dual purpose of showing our commitment to student learning while simultaneously delivering on that promise.

6. Further Research

This research was motivated by the growing number of institutions, including ours, that have or are contemplating investing in special purpose classrooms dedicated to collaborative learning. We have dubbed these as problem-based learning (PBL) rooms.

The investments required are substantial—our institution has spent US $190,000 each to convert several traditional classrooms to PBL rooms. While the costs were clear, the benefits of doing so were, to the best of our knowledge, intuitive. It was assumed that outfitting classrooms with tables and chairs on wheels would facilitate collaboration and that networked computers with large display screens would facilitate student interaction with each other as they collaborated on solving problems.

While the above reasoning appealed to almost everyone's intuition, there was no data to support the intuition. Did PBL rooms do a better job of engaging students and did they lead to better outcomes than other types of classrooms that might already be available?

We measured student engagement and outcomes in three different types of classrooms: stadium, flat and PBL. The same course was taught by the same instructor using flipped learning. Flipped learning required collaboration when the course met in a classroom. We used the scale of transactional distance to measure student engagement and outcomes in each of the classroom. Our results indicated that there was no significant difference in student engagement and outcome between the PBL rooms and the flat classrooms. Hence, from a student perspective, the investment in PBL rooms was not justifiable since students taking the course in traditional flat rooms were just as engaged and had similar outcomes.

Our results were for a specific type of pedagogy (flipped) and a specific business course, Business Analytics, a quantitative course requiring students to use computer models to solve problems. Thus, the results may not be transferrable to other types of collaborative learning or other types of disciplines. Research is required for alternative disciplines and collaboration modes to determine whether our results can be generalized.

However, our methodology is not dependent on subject matter or collaboration mode. It can be used to determine student engagement and outcomes in any subject, collaboration

mode and/or classroom type. We tested the methodology on a specific flipped course in existent classrooms. Our results did make sense in identifying the stadium room as yielding the least student satisfaction and outcomes amongst the three types of classroom tested. This result that was confirmed by the much smaller number of students that would recommend the course to their friend when taught in a stadium room than if were taught in a PBL or flat classroom.

The purpose of our research was to provide a quantitative measure of student engagement and outcomes to decision makers *before* a decision was made to invest in a specific PBL room design. Our research has developed a tool that can compare actual classrooms based on student engagement and outcomes. Research is required on whether stakeholders confronted at the planning stage with several alternate classroom designs can evaluate such designs, preferably as virtual reality models, by completing the scale of transactional distance for each in order to provide decision makers with a measure of how each design would generate student engagement and outcomes and use this as part of their design selection criteria.

Author Contributions: Each author has contributed equally to all aspects of this research. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Study ID: UMCIRB 18-000457 Date Approved 3/9/2018 Does not expire.

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

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

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

Research Ethics: Study ID: UMCIRB 18-000457. Date approved 3 September 2018. Does not expire.

Appendix A

Table A1. Scale of Transactional Distance.

FACTOR	Element	Description/Question
		Transactional Distance between Student and Instructor
TDST	tdsi1	The instructor generally answers the student's questions
	tdsi2	The instructor pays no attention to me
	tdsi3	I receive prompt feedback from the instructor on my academic performance
	tdsi4	The instructor was helpful to me
	tdsi5	The instructors is available to answer my questions
	tdsi6	The instructor can be turned to when I need help in the course
		Transactional Distance between Student and Content
TDSC	tdsc1	The content of this course is of great interest to me
	tdsc2	I do not know why I have to learn this
	tdsc3	The examinations in this course have challenged me to do my best work
	tdsc4	This course emphasized SYNTHESIZING and organizing ideas, information, or experiences into new, more complex interpretations and relationships
	tdsc5	This course emphasized MAKING JUDGEMENTS about the value of information, arguments, or methods such as examining how others gathered and incorporated data and assessing the soundness of their conclusions
	tdsc6	This course emphasized APPLYING theories and concepts to practical problems or in new situations

Table A1. *Cont.*

FACTOR	Element	Description/Question
		Transactional Distance between Student and Student
TDSS	tdss1	I learned a lot from observing the interactions among the students
	tdss2	The students in this online class challenged me to do my best work
	tdss3	I get along well with my classmates
	tdss4	I feel valued by the class members in this online class
	tdss5	My classmates in this online class value my ideas and opinions very highly
	tdss6	My classmates respect me in this online class
	tdss7	I am good at working with the other students in this online class
	tdss8	I feel a sense of kindred spirit with my fellow classmates
	tdss9	The class members can be turned to when I need help in the course
	tdss10	There are students I can turn to in this online class
	tdss11	The class members are supportive of my ability to make my own decisions
		Transactional Distance between Student and Instructional Technology
TDSTECH	tdstech1	It is difficult to pay attention to the instructor in the web environment
	tdstech2	I have adequate access to the web resources I need
	tdstech3	The fact that I am online does not inhibit my class participation
	tdstech4	An efficient system is provided for students and instructor to exchange materials
	tdstech5	I am comfortable using the computer
	tdstech6	I hate using the web
	tdstech7	It was easy for me to use the technology involved with this online class
	tdstech8	The technology used in this course is difficult to learn and use
		Zhang's Outcomes
LEARNING GOALS		I have learned a great deal in this class
		I have made tremendous progress towards my goal in the subject area of this course
SATISFACTION		Overall, I am satisfied with this course

References

1. Pickles, M. Shouldn't Lecture Be Obsolete by Now? *BBC NEWS*. 2016. Available online: https://www.bbc.com/news/business-38058477 (accessed on 12 May 2020).
2. Prince, M. Does active learning work? A review of the research. *J. Eng. Educ.* **2004**, *93*, 223–231. [CrossRef]
3. Bishop, J.; Verleger, M. The flipped classroom: A survey of the research. In Proceedings of the 120 ASEE Annual Conference and Exposition, Atlanta, GA, USA, 23–26 June 2013. Paper ID #6219.
4. Selingo, J. *College (Un)bound: The Future of Higher Education and What It Means for Students*; Houghton Mifflin Harcourt: New York, NY, USA, 2013.
5. Talbot, R. How Much Research Has Been Done on Flipped Learning? 2020 Update for 2020. Available online: http://rtalbert.org/how-much-research-has-been-done-on-flipped-learning-update-for-2020/ (accessed on 10 February 2021).
6. Noonoo, S. Why Flipped Learning Is Still Going Strong 10 Years Later. 2017. Available online: https://www.edsurge.com/news/2017-10-03-why-flipped-learning-is-still-going-strong-10-years-later# (accessed on 15 February 2021).
7. AP News. $1.9 Bn Flip Classroom Market—Global Industry Trends, Share, Growth, Opportunity and Forecast 2019–2024—Research and Markets.com. 8 July 2019. Available online: https://apnews.com/press-release/pr-businesswire/f29b1c293dcc47bbb2562a9380cb5d22 (accessed on 1 March 2021).
8. Swart, W.; Wuensch, K. Flipping Quantitative Classes: A Triple Win. *Decis. Sci. J. Innov. Educ.* **2016**, *14*, 67–89. [CrossRef]
9. Swart, W. *Extending the Principles of Flipped Learning to Achieve Measurable Results: Emerging Research and Opportunities*; IGI Global: Hershey, PA, USA.
10. Adedokun, O.A.; Parker, L.C.; Henke, J.N.; Burgess, W.D. Student Perception of a 21st Century learning space. *J. Learn. Spaces* **2017**, *6*, 1–13.
11. Lee, D.; Morrone, A.S.; Siering, G. From swimming pool to collaborative learning studio: Pedagogy, space, and technology in a large active learning classroom. *Educ. Technol. Res. Dev.* **2018**, *66*, 95–127. [CrossRef]
12. Walker, J.P.; Baepler, P. Measuring social relations in new classroom spaces: Development and validation of the Social Context and Learning Environments (SCALE) survey. *J. Learn. Spaces* **2017**, *6*, 34–41.
13. Clinton, V.; Kelly, A.E. Student attitudes toward group discussions. *Act. Learn. High. Educ.* **2020**, *21*, 154–164. [CrossRef]
14. Machemer, P.L.; Crawford, P. Student perceptions of active learning in a large cross-disciplinary classroom. *Act. Learn. High. Educ.* **2007**, *8*, 9–30. [CrossRef]
15. Wagner, S.C.; Garippo, S.J.; Lovaas, P. A Longitudinal Comparison of Online versus Traditional Instruction. *J. Online Learn. Teach.* **2011**, *7*, 68–78.

16. Leverett, H.M.; Zurita, F.I.; Kamery, R.H. The effects of adjunct instruction on the final grades in the principles of accounting course. In Proceedings of the Academy of Legal, Ethical and Regulatory Issues, Memphis, TN, USA, 13–16 April 2005; Volume 9, pp. 79–83.
17. Hereid, C.; Schiller, N. Case Studies and the Flipped Classroom. *J. Coll. Sci. Teach.* **2013**, *42*, 62–66.
18. Janz, K.; Gratez, K.; Kjorlien, C. Building Collaborative Learning Environments. In Proceedings of the SIGUCCS Annual Conference, SIGUCCS 12 ACM, Montreal, QC, Canada, 27–29 June 2012; pp. 121–126.
19. Lipmann, P. Designing Collaborative Spaces for Schools. *J. Digit. Ed.* January 2013. Available online: https://thejournal.com/articles/2013/02/13/design-ing-collaborative-spaces-for-schools.aspx (accessed on 21 November 2019).
20. Villano, M. 7 Tips for Building Collaborative Learning Spaces. 2010. Available online: http://campustechnology.com/Articles/2010/06/01/7 (accessed on 27 February 2015).
21. Weinstein, C.S. The Physical Environment of the School: A Review of the Research. *Rev. Educ. Res.* **1979**, *49*, 577–610. [CrossRef]
22. Oblinger, D. *Learning Spaces*; EDUCAUSE: Boulder, CO, USA, 2006.
23. Clinton, V.; Wilson, N. More than chalkboards: Classroom spaces and collaborative learning attitudes. *Learn. Environ. Res.* **2019**, *22*, 325–344. [CrossRef]
24. Peberdy, D. Editor, Active Learning Spaces and Technology Advances in Higher and Further Education. 2014. Available online: http://www.wolfvision.com/wolf/Active_Learning (accessed on 2 February 2020).
25. Crabb, M.; Clarke, D.; AlWaer, H.; Heron, M.; Laing, R. Inclusive Design for Immersive Spaces. *Des. J.* **2019**, *22*, 2105–2118. [CrossRef]
26. Cleveland, B.; Fisher, K. The evaluation of physical learning environments: A critical review of the literature. *Learn. Environ. Res.* **2014**, *17*, 1–28. [CrossRef]
27. Moore, M.G. Theory of Transactional Distance. In *Theoretical Principles of Distance Education*, 1st ed.; Keegan, D., Ed.; Routledge: London, UK, 1993; pp. 22–38.
28. Zhang, A. Transactional Distance in Web-Based College Learning Environments: Toward Measurement and Theory Construction. Ph.D. Dissertation, Virginia Commonwealth University, Richmond, VA, USA, 2003.
29. Schwab, K.; Moseley, B.; Dustin, D. Grading Grades as a Measure of Student Learning. *SCHOLE J. Leisure Stud. Recreat. Educ.* **2018**, *33*, 87–95. [CrossRef]
30. Axelson, R.D.; Flick, A. Defining Student Engagement. *Chang. Mag. High. Learn.* **2010**, *43*, 38–43. [CrossRef]
31. Paul, R.C.; Swart, W.; Zhang, A.M.; MacLeod, K.R. Revisiting Zhang's scale of transactional distance: Refinement and validation using structural equation modeling. *Distance Educ.* **2015**, *36*, 364–382. [CrossRef]
32. Swart, W.; MacLeod, K.; Paul, R.; Zhang, A.; Gagulic, M. Relative Proximity Theory: Measuring the Gap between Actual and Ideal Online Course Delivery. *Am. J. Distance Educ.* **2014**, *28*, 222–240. [CrossRef]
33. Allen, E.; Seaman, J.; Poulin, R.; Taylor, T. *Online Report Card: Tracking Online Education in the United States*; Babson Survey Research Group and Quahog Research Group, LLC: Babson Park, MA, USA, 2016.
34. Hall, C.; Swart, W.; Duncan, S. Balancing Customer Needs and Standards in Higher Education. *Qual. Approaches High. Educ.* **2013**, *3*, 2–7.
35. Grant, C.; Thompson, A. Revolutionizing School Design with Virtual Reality. HMC Architects Blog. 9 November 2017. Available online: https://www.bdcnetwork.com/blog/revolutionizing-school-design-virtual-reality (accessed on 4 April 2021).

education sciences

MDPI

Article

A Qualitative Assessment of the Pedagogical Process at Czech Public Universities

Jiří Bečica and Roman Vavrek *

Department of Public Economics, Faculty of Economics, VŠB—Technical University of Ostrava, Sokolská Třída 33, 70200 Ostrava, Czech Republic; jiri.becica@vsb.cz
* Correspondence: roman.vavrek@vsb.cz; Tel.: +420-597-322-334

Abstract: Quality is currently an often-used term in all areas of human activity. However, the measurement of quality is very problematic in the field of education, particularly if no specific, comprehensible criteria for its measurement, accepted by most subjects active in the specific sector, exist. Monitoring quality in the field of education is difficult because there is no long-term embedded quality standard and the established level can be affected not only by the selection of the chosen criteria for measurement, but also by determining specific weights when comparing the importance of the chosen criteria. The authors of this paper endeavour to point out one way of assessing the quality of publicly established universities in the Czech Republic during the academic years 2011/2012 and 2018/2019 on a basic sample of all 26 publicly established universities. The quality of the pedagogic apparatus and the converted number of students indicate that the classification of schools into categories according to the Ministry of Education, Youth, and Sports of the Czech Republic (MEYS) is possible, but the individual categories should be discussed and modified according to the assessment performed.

Keywords: education; quality; universities; Czech Republic

Citation: Bečica, J.; Vavrek, R. A Qualitative Assessment of the Pedagogical Process at Czech Public Universities. *Educ. Sci.* **2021**, *11*, 389. https://doi.org/10.3390/educsci11080389

Academic Editors: Sandra Raquel Gonçalves Fernandes, Marta Abelha and Ana Teresa Ferreira-Oliveira

Received: 1 July 2021
Accepted: 26 July 2021
Published: 29 July 2021

1. Introduction

The education sector is one of the most important public sectors and is included in the development services sector within the terms of the national economy and also among nonprofit organisations. A common trait of nonprofit organisations is that they are not primarily established for the purpose of generating a profit (even though they may report a profit), but for the purpose of providing services to broad sections of the public, and are mostly financed from public budgets. According to the Czech Statistical Institute, the greatest number of nonprofit organisations in the Czech Republic are registered in the spheres of education, culture, or social services. Their objective is to encourage the economic growth of the country, maintain and cultivate human potential, and provide high-quality services to the population.

However, the quality of services is very difficult to measure, particularly in a non-market environment and with the missing prerequisite of economic or other measurable outputs (inputs can be measured). The term "quality" is most often used to identify something that is optimal, desirable, or ideal, or meets a specific standard. Monitoring quality (not just of tuition) is generally considered relevant from the viewpoint of the top management of an organisation and also from the viewpoint of its founder or providers of funds (donors). Quality is of interest not only to consumers of services (pupils, students, and their parents), but the outputs should also be important to pedagogues (particularly teachers), employers, and politicians, whose steps decide on the state's educational policy in the long-term horizon and who should strive to require, guarantee, support, and increase the quality of tuition. This itself assumes that a standard (assessment reference framework) will be defined and observed in the long term, with measures implemented leading to remediation in the event of an undesirable (negative) finding. The quality of

services provided by individual organisations, public universities in the Czech Republic (hereinafter "PU"), can therefore be evaluated on the basis of various criteria.

The organisations assessed below, all active in the sphere of public university education in the Czech Republic, can provide purely public assets, mixed, and purely private assets. It is therefore natural that the costs for education, which can differ not only due to the long-term economic situation of the country (GDP), but also, for example, due to the historical development of the educational system of the specific country, are understood in terms of an international comparison.

The objective of this paper is to measure the quality of the pedagogic process during the provision of services at publicly established universities in the Czech Republic during the academic year 2011/2012, and to assess any changes in this field in the academic year 2018/2019. For the purposes of fulfilling the set objective, the submitted paper is structured as follows. Section 2 discusses quality, its perception in the educational process, and the approaches to its measurement. Section 3 focuses on presenting tertiary education in the Czech Republic and the method of its financing as the sector on which the executed research focuses. Section 4 defines the objective, the research hypotheses, the basic group of monitored quality indicators, and the method of their verification. Section 5 is devoted to the results of our own research, i.e., the assessment of quality. Section 6 contains a discussion of the results in the context of local conditions.

2. Quality and Its Measurement

The concept of quality is often used in a number of sectors. The answer to the question of "what is quality?" is very problematic, because, in order to be able to state that something is of high quality, there must be agreement on what the minimum acknowledged standard in the assessed area is. Economists most often imagine an effective expenditure of funds under the term of quality, or consider the price of goods or services in relation to the performance. A problem arises in situations when assets and services are provided with no link to the actual market price (for instance, compulsory school attendance), or the provider of the service does not conduct its activities on a profit-generating principle. The additional terms productivity, economy, and efficiency can therefore also be encountered in relation to the concept of quality.

Terhart [1] states that the concept of quality is used chiefly formally, for the purpose of introducing a difference. As a formal category, it allows for a differentiation between the less valuable and the more valuable. In order to understand how valuable things are, the concept of quality must be given content. In other words, the definition of quality is not an issue of the gathering of quantitative evidence as such; according to Terhart [1], it essentially consists in finding and justifying the content criteria. Fend [2] states that the concept of "high quality" fulfils the function of an assessment concept, i.e., objectivising the value or a general quality of some item. The term "excellence" is proposed in the effort to name a level of quality that is higher than high.

2.1. Perception of Quality in the Educational Process

The first research, the objective of which was to establish the cause of a pedagogue's success in teaching, was carried out roughly from the middle of the 20th century. Within the terms of research, psychometric methods were used to investigate so-called personality paradigms (research on teacher personality), i.e., which personality traits of a pedagogue had a positive impact on pupils and students and their performance. The research indicated that an enthusiastic and open teacher motivated pupils to learn more and perform better than a teacher who did not have these traits.

The concept of the quality of tuition began to appear in the literature in the 1960s, particularly thanks to the works of Caroll [3] and Bloom [4]. Einsiedler [5] states that there has been an effort to compensate for the bias of the approach by focusing only on the quantity of tuition, by devoting attention to the characteristics (of quality), such as the comprehensibility, structuring, and cohesion of tuition, in the background of Carroll's and Bloom's models. Caroll [3] worked with the factors of time needed to learn and time available to learn. The lower the quality of tuition, the more time the pupil needs to learn. In his concept of tuition and its quality, Bloom [4] took into consideration not only the cognitive performance of pupils, but also the motivational and affective objectives (interests, standpoints, motivation, and self-perception). Quality can be defined variously, even within the terms of the teaching process or the assessed teaching level (primary, secondary, or university). Harvey and Green [6] provided a definition of quality in their work, for instance. Research examining the relationship between the teacher's behaviour and the performance/results of pupils is presented, for example, by [7–9].

Janík [10] states that addressing the topic of quality in education requires courage. A discussion about what is qualitatively good cannot avoid the themes of what is not qualitatively good. In other words, a discussion about quality will only make sense if it also includes a discussion of poor quality, or the standard on which there is a consensus. Weinert et al. [11] stated that the quality of tuition can be defined like "any stable mode of behaviour, which enables substantial prediction as a whole or by means of individual components." According to Martensson et al. [12], good (high-quality) tuition is considered to be tuition that leads to excellent learning results on the part of the student.

Rýdl [13] points out that the concept of quality in education is subjective, often defined by immeasurable and vague phenomena as compared to existing standard quantitative methodological procedures. Starý and Chvál [14] stated that the formula "emphasis must be placed on the quality of tuition" often appears within the terms of the educational process, which applies to the high quality of tuition—as defined, for example, by Průcha [15]. In this regard, Janík [10] stated that quantitative and qualitative methodologists clash on an academic level, with regard to the orientation towards hard versus soft indicators when measuring the quality of education.

2.2. Approaches to and Measuring Quality in the Educational Process

Terhart [1] stated that a school that supports normative designations, implements these in its programme, and fulfils them in its work is considered a "good school." The normative approach to the definition of quality consists in determining and justifying the roles of education, as an institution in society. During this time, decisions are made about what effects of education can be considered an expression of higher quality.

On the contrary, the analytical approach is based on an examination of the various concepts of quality and their use in discussions about education. This is about deriving the perception of the quality of education from an analysis of discourse about a specific period or a specific (governing) educational culture or tradition. The analytical approach can be seen, for instance, in the works of Harvey and Green [6], who defined five different but mutually related concepts of quality on the basis of an analysis of the methods of perceiving quality:

- quality as uniqueness,
- quality as flawlessness,
- quality as effectiveness,
- quality as an adequate countervalue, and
- quality as transformation.

The work of Jürgens [16], who analysed discussions of reform pedagogy in the context of the quality of education, can serve as another example. Pupala [17] stated that it is very difficult to come to a unified understanding of the concept of quality in education, because various actors define quality differently.

Another way to assess quality is by means of an empirical approach, where attention is focused on the actual impact of educational institutions, which is placed into context through their official roles. Mincer [18] stated that this approach can be applied, for example, when measuring the unemployment rate of university graduates. The empirical approach focuses on proposed goals, available resources, programmes used, and effects achieved. On the basis of the above, it is possible to determine the relationship between costs and revenue and to approximate the economic perception of the concept of quality as closely as possible. An example of research that allows for the differentiation of effective schools from less effective schools on this basis is TIMSS research. TIMSS (Trends in International Mathematics and Science Study) research is organised by the International Association for the Evaluation of Achievement and focuses on evaluating the knowledge and skills of pupils in various types of school in terms of mathematics and natural sciences.

On the contrary, PISA (Programme for International Student Assessment) research is an expression of the determination of OECD (Organisation for Economic Co-Operation and Development) countries to monitor outputs from the educational systems of individual countries in terms of an international comparison, on the basis of measuring the educational results of pupils. This research focuses on the reading, mathematical, and natural science literacy of 15-year-old pupils, usually at the end of their compulsory school attendance.

There are a number of studies devoted to the issue of the quality of educational institutions: for instance [2,10,19–27].

3. Key Parameters of Education in the Czech Republic

In the Czech Republic, education is a public asset with collective consumption, which has positive social benefits. It is a preferred public asset, in whose consumption the state is interested. This is usually evidenced by the greater expenditure of funds from public sources in proportion to the GDP of the country. A 2016 OECD study that maps the state of education in the most advanced countries in the world (OECD countries) indicates that the Czech Republic expends an average of approx. 3–4% of its GDP on the field of education in the long term (Table 1).

Table 1. Budget of the Ministry of Education, Youth, and Sports of the Czech Republic in 2010–2020.

	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
TR	2.01	6.15	1.51	1.78	1.23	7.18	8.11	8.64	7.19	10.15	12.28
TE	125.2	127.0	137.8	140.4	137.3	135.9	142.3	156.5	176.1	205.7	226.4
SH	3.34	3.33	3.53	3.62	3.45	2.86	2.88	3.03	3.18	3.50	3.87

TR—Total revenue (in billion CZK), TE—Total expenditure (in billion CZK), SH—Share of MEYS chapter expenditures in GDP (%).

The average expenditure on education of all OECD countries is 5.2% of the GDP. The proportion of public and private funds invested into the educational system differs in individual countries, but public funds usually predominate and account for, on average, 83% of total expenditure on education in OECD countries. This percentage is reported as 87% in the Czech Republic.

Expenditure in the field of education is, on average, 11.6% of all public expenditure in OECD countries. At 8.9%, the Czech Republic is similar to countries such as Hungary, Italy, and Spain in this area and these also reported values of around 8%. In the past three years, the percentage of expenditure on education has increased, which is the result of more funds being funnelled into wages, particularly in regional education. Regional education in the Czech Republic consists of pre-primary, primary, and secondary level education according to the ISCED, on which a total of CZK 161.5 billion was expended in 2020. The decision to fund individual levels of the educational system differs in various countries and is the result of a political decision.

3.1. Funding of Tertiary Education in the Czech Republic

In the Czech Republic, education on a tertiary level at public and publicly established universities is funded from multiple sources and is free for the time being. The main sources of funding of public universities, which are established on the basis of specific laws, include subsidies from the budget of the Ministry of Education, Youth and Sports of the Czech Republic for activities, funds from scientific research activities, funds from various European funds, and own sources from business activities and donations. Financing from EU funds is also among the sources of funding.

Of the annual volume of approx. CZK 225 billion expended from the budget of the Ministry of Education, Youth, and Sports of the Czech Republic on education in 2020, 21.2% of expenses (a total of CZK 48.4 billion) was expended in the field of tertiary education; of this, CZK 21 billion went into the field of research, development, and innovation (approx. CZK 7.1 billion from programmes cofunded by the EU) and CZK 27.4 billion was for funding university education, particularly 26 publicly established universities, which are compared below in terms of the aforementioned quality of pedagogic and scientific workers (see Figure 1).

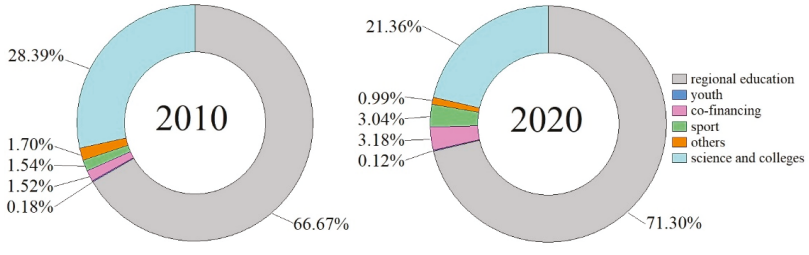

Figure 1. Structure of expenditure by the Ministry of Education, Youth, and Sports of the Czech Republic in 2010 and 2020.

3.2. The Quality of Tertiary Education in the Czech Republic

There are various approaches to the assessment of the quality of university institutions in the available specialist literature. Bloch et al. [28] stated that the concept of "quality" of university education can also be encountered in a number of political declarations, which do not necessarily have to be related to the provision of high-quality performance and measurable outputs on the part of providers of education, i.e., individual universities. The study by Mulder et al. [29] states that the assessment of the work of pedagogues and their students leads to an improvement of the quality of education. Berezvai et al. [30] stated that a better rating of students by pedagogues subsequently leads to a better rating of pedagogues by students. Stewart [31] assessed the influence of the professional growth of university pedagogues on the teaching process. The authors of this paper approach the assessment of public universities below in a similar spirit.

Černikovský [32] stated that there is no wide-ranging agreement on what the quality of universities means or should mean. Various actors work with various concepts of quality, often not reflected on more deeply and also mutually exclusive in many aspects. The authors of this paper approach the differentiation of the "quality" of university education in the Czech Republic according to the areas of pedagogy, scientific research, and accreditation activities.

Standards of quality within the terms of accreditation activities were established in university education in the Czech Republic in the 1990s, when the accreditation of study programmes was officially implemented on the basis of University Act No. 111/1998 Sb. The act established a national agency with the title of "Accreditation Committee", which was granted the competence and authority to grant accreditation for all levels of university studies, including the granting of rights to carry out rehabilitation and professorial

proceedings and their subsequent qualitative assessment. Cardoso et al. [33] questioned whether national agencies, which require a specific standard of "quality" for accreditation of university education in Europe, are actually capable of promoting an increase in the quality of education and stated that this remains an open question for discussion.

The competence of the Accreditation Committee in the Czech Republic was assumed by the "National Accreditation Office" [34] in 2016, on the basis of an amendment of the University Act. This organisation also newly discusses the granting of so-called institutional accreditation. On the basis of this accreditation, the university can subsequently approve new accreditations within the terms of the institution, on the basis of an established "Committee for internal assessment," for a period of 10 years, as well as changes to the accreditation that has already been granted. In this area, we can state that a number of set criteria are given vaguely and interpreted and assessed variously when executing accreditation files. The monitoring and reporting of some indicators is also difficult to understand, and universities consider this unnecessarily demanding from the point of view of the administrator. Tesar [35] stated that not enough research has been carried out yet in order to be able to state that the accreditation system itself will lead to increased quality.

However, according to the authors, the greatest problem with assessing "quality" at universities is in the field of assessing pedagogic activities. It would be desirable to focus more attention on these activities on the level of the management of individual universities and on a national level. The result could and should be a more objective assessment of academic and scientific workers at individual universities, which could be followed by a differentiated remuneration of such "quality" tuition executed by the specific workers. Kember [36] stated that many academics chiefly consider themselves specialists in their discipline, and it is difficult to convince them to try innovative forms of tuition. This should lead to a greater involvement of students in tuition, as confirmed by existing evidence about the higher efficacy of some new forms of tuition on learning [37].

In the field of scientific research (generally creative) activities, a detailed methodology for assessing "quality" has been created for pedagogic and scientific workers in the Czech Republic, which is constantly updated by the management of universities and also national authorities (including quantitative indicators). The monitored criteria change over time. For instance, the required minimum criteria for scientific research activities necessary for the successful completion of studies within the terms of doctoral study programmes or initiation of rehabilitation proceedings and proceedings for appointment as a professor increase constantly. The fact that these criteria differ substantially between the same fields and individual schools (faculties) is problematic. A frequent issue concerns the changes to the weights of reported and monitored criteria, which have a significant impact on the long-term predictability of funding the educational institution as a whole. Whether the growth of such a monitored "quality" in scientific research activities also has the effect of increasing the quality of an educational institution in terms of its pedagogic activities is addressed below.

4. Materials and Methods

As stated above, the objective of this paper is to measure the quality of the pedagogic process during the provision of services at publicly established universities in the Czech Republic during the academic year of 2011/2012, and to note any changes in this field in the academic year of 2018/2019.

A basic group of public universities consisting of 26 subjects, classified by [38] into four groups, as recorded in Table 2, is assessed.

Table 2. Classification of universities into groups.

Groups	Universities
Art universities (group 1)	AMU, AVU, JAMU, UMPRUM
Nonuniversities (group 2)	VŠTE, VŠPJ
Smaller universities (group 3)	VŠCHT, ČZU, JU, MEN, OU, SU, TUL, UHK, UJEP, UPa, UTB, VFU, VŠB-TUO, VŠE, ZČU
Larger universities (group 4)	MU, ČVUT, UK, UP, VUT

For this purpose, the quality of the provided services on the level of the assessed public universities is measured by means of:

- the number of students per pedagogic or scientific worker according to acquired qualifications (professor, docent, specialist assistant, assistant),
- the quality of the pedagogic apparatus, i.e., the pedagogic and scientific worker, according to acquired qualifications (professor, docent, specialist assistant, assistant).

For the calculation of both quality indicators specified above, it is necessary to quantify the "differences" between the qualifications acquired by pedagogic and scientific workers, i.e., the differences between a professor, docent, specialist assistant, and assistant. The Ministry of Education, Science, Research, and Sport of the Slovak Republic [39] quantifies these differences using a coefficient of the qualification structure as follows: professor—2, docent—1.66, specialist assistant—1.33, and assistant—1. Using this coefficient, the indicators in question for the individual public universities are calculated as follows:

$$Q1 = \frac{\text{students}}{Q2} \tag{1}$$

$$Q2 = \frac{\text{profesors} \times 2 + \text{docents} \times 1.66 + \text{specialist assistants} \times 1.33 + \text{assistants} \times 1}{\text{profesors} + \text{docents} + \text{specialist assistants} + \text{assistants}}. \tag{2}$$

Using the created quality indicators, the following research hypotheses are verified in the subsequent chapter using the created quality indicators:

Hypothesis 1. *Assumption of statistically significant differences in the number of students (Q1) between individual PU groups.*

Hypothesis 2. *Assumption of statistically significant differences in the quality of the pedagogic apparatus (Q2) between individual PU groups.*

Hypothesis 3. *Assumption of a statistically significant negative linear relationship between the number of students (Q1) and the quality of the pedagogic apparatus (Q2) on the level of individual PU groups.*

The definition of H1 and H2 is based on the classification of PU into four MEYS groups. Within the terms of H3, we assume that a smaller number of students taught by one pedagogue should lead to a higher quality of tuition, i.e., it translates into more opportunities for an individual approach by the pedagogue to the student.

The input data for assessment are taken from the annual reports of individual public universities within the terms of the monitored period, [40–42], and other documents available on the website of the Ministry of Education, Youth, and Sports of the Czech Republic.

The acquired results are supplemented by statistical verification using a Kruskal–Wallis test (Q), Levene test (LE), and Kendall rank correlation coefficient (r_K). A multicriteria assessment is executed using TOPSIS techniques (see [43,44] for the calculation procedure) under the condition of the equality of monitored quality indicators. The analyses are executed with MS Excel, Statistica 13.4, and the Statgraphics XVIII software.

5. Results

This section assesses the quality of universities in the Czech Republic in various aspects in the academic years 2011/2012 and 2019/2019. "Quality" is initially assessed on the basis of the number of students per standardised pedagogic worker. This is followed by the assessment of the "quality" of pedagogic workers by means of the qualifications they have acquired (professor, docent, specialist assistant with PhD, or specialist assistant without PhD). The authors integrated the monitored indicators of scientific research activities, and also the monitored indicators of the "quality" of tuition within the terms of accreditation activities, into the assessment of the quality of the pedagogic process by means of this indicator. The authors of this paper work on the assumption that university pedagogic workers have a high motivation to acquire higher qualifications and therefore move up the ranks of the university. The universities themselves have detailed internal regulations and guidelines for this matter, the objective of which is a lump sum reward on professional progress within the terms of qualification (successful acquisition of a scientific research grant, publication of articles in magazines with an impact factor, acquisition of a new accreditation or defence of an established accreditation, etc.). The pedagogic workers are also motivated to increase their qualifications by long-term financial compensation, as well as lump sum rewards. This is linked to their progress to higher qualifications and other benefits, which are linked to a rise in status (moral perception by the general public, etc.). This is followed by a section that shows the "quality" of the pedagogic process by combining both indicators, i.e., the number of students in combination with the number of pedagogues and their acquired qualifications.

5.1. Number of Students as a Quality Indicator (Q1)

The first assessed quality indicator is the number of students per standardised pedagogic worker. The differences at the beginning and end of the monitored period, i.e., in the academic year 2011/2012 and the academic year of 2018/2019, are illustrated in Figure 2.

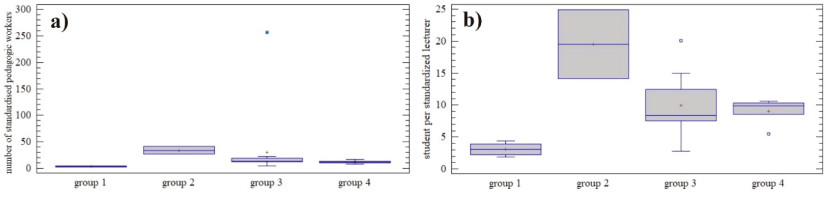

Figure 2. Number of students per standardised pedagogic worker according to MEYS groups in (**a**) academic year 2011/2012 and (**b**) academic year 2018/2019.

In the academic year 2011/2012, we observed different values of this indicator across groups of universities. During the illustration using a box plot (see Figure 1), which we needed to interpret in the context of the extent of variation influenced by extreme values, we noted a difference in the median (Q = 13.5462; $p \leq 0.01$). During this time, with the application of the Bonferroni post hoc method, we could identify the first group of universities, i.e., art universities, as a separate group. We can confirm H1 on the basis of these results.

Despite this difference, the homogeneity of their dispersal was confirmed (LE = 0.8543; $p = 0.4793$), and therefore the differences within individual groups are the same. Differences between individual groups of universities increased significantly at the end of the monitored period. Nonuniversity colleges (the second group) retained a dominant position; we recorded the fewest students in the academic year of 2018/2019 at art universities. The differences (Q = 12.0182; $p \leq 0.01$) and also uniform dispersal (LE = 2.6754; $p = 0.0722$) were maintained. Whether we can extrapolate this to the entire monitored period or whether it is relevant to just the two aforementioned academic years is reported in Figure 3.

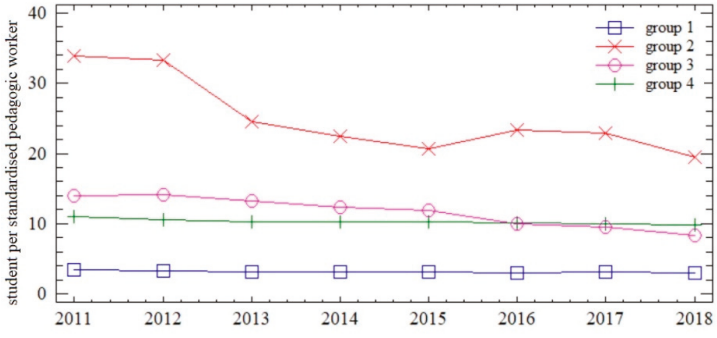

Figure 3. Evolution of the median number of students per standardised pedagogic worker according to the MEYS group from academic years 2011/2012 to 2018/2019.

When monitoring the change in the number of students per standardised pedagogic worker (Figure 3), we observed a stable development, or a slight reduction, in the case of art universities (the first group) and larger universities (the fourth group). In the case of these institutions, the shift in the number of pedagogic workers reflects the shift in the number of students and vice versa, whereas the year-on-year shift did not exceed 5% (negative in the majority of cases). In the other two groups, this year-on-year reduction was more marked, often exceeding 10% (six of 14). Nonuniversity colleges and smaller universities are unable to deal with a fall in the number of students, which is also reflected in the number of their pedagogic workers (see Figure 4).

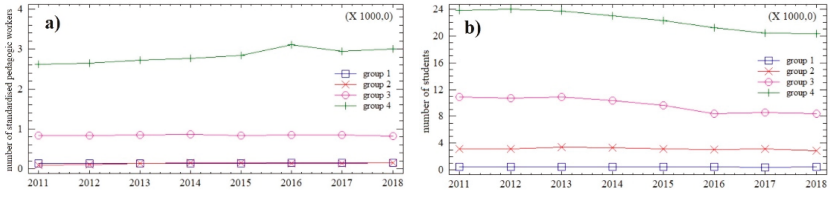

Figure 4. Evolution of the median number of standardised pedagogic worker (**a**) and students (**b**) according to the MEYS group from academic years 2011/2012 to 2018/2019.

5.2. Quality of the Pedagogic Apparatus as a Quality Indicator (Q2)

The second assessed indicator for evaluating the quality of the pedagogic process at universities is the quality of the pedagogic apparatus, taking into consideration qualifications in the form of academic titles. The state at the beginning and end of the monitored period, i.e., academic year 2011/2012 and academic year 2018/2019, is reported in Figure 5.

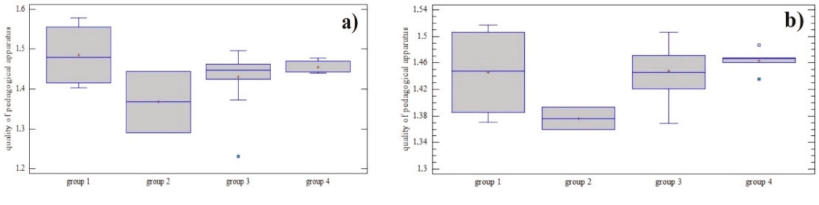

Figure 5. The quality of the pedagogic apparatus according to the MEYS group in academic years 2011/2012 (**a**) and 2018/2019 (**b**).

At the beginning of the monitored period, we can observe the quality of the pedagogic apparatus (Figure 5) through the variability within the terms of the group and also in the difference in quality between groups. The greatest differences are in the least numerous group of nonuniversity colleges (v_{group1} = 7.96%); on the contrary, the smallest differences can be observed for larger universities, i.e., within the terms of the fourth group (v_{group4} = 1.22%). However, the differences in the quality of the pedagogic apparatus cannot be called significantly different between individual groups (Q = 1.7487; p = 0.6261). On the basis of these results, we can disprove H2.

At the end of the monitored period, we repeatedly observe the uniformity of the group of larger universities (v_{group4} = 1.25%) during a reduction of the differences in other groups. Differences within the terms of the variability of individual groups (LE = 4.5110; p = 0.0130), not between them (Q = 4.8372; p = 0.1844), were confirmed. Whether we can attribute this state to the entire monitoring period is reported in Figure 6.

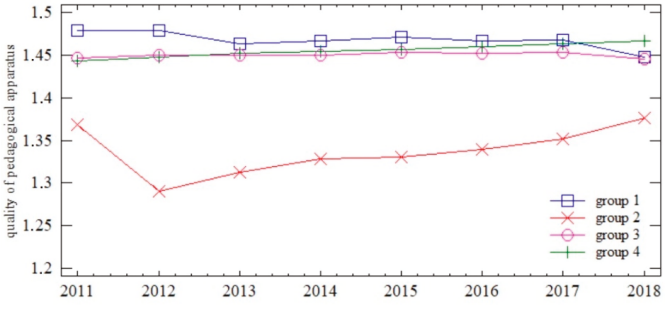

Figure 6. Development of the median quality of the pedagogic apparatus according to the MEYS group from academic years 2011/2012 to 2018/2019.

When monitoring the development of the quality of the pedagogic apparatus, we observed a different situation between groups of universities (Figure 6). Year-on-year changes, with the exception of one situation (academic year 2012/2013 in the second group), did not exceed 2% in the positive or negative sense. From this viewpoint, the differences between the three groups of universities are minimal. The quality of the pedagogic apparatus is lowest in the long term at nonuniversity colleges; however, these differences continued to decrease after academic year 2012/2013. This structure, expressed in terms of the number of individual pedagogic workers, is reported in Figure 7.

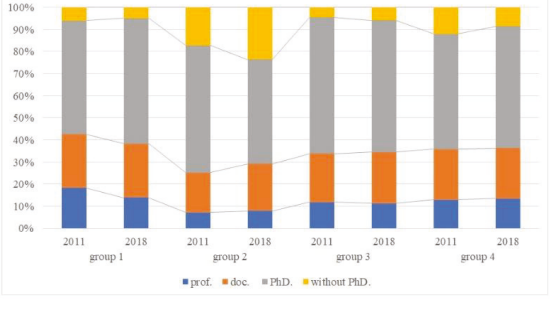

Figure 7. Structure of the pedagogic apparatus of the MEYS groups from academic years 2011/2012 to 2018/2019 (median).

5.3. Number of Students and the Quality of the Pedagogic Apparatus from the Viewpoint of Individual Universities

The last part of the analysis concerns both monitored indicators of the quality of the pedagogic process simultaneously. The results of monitoring the number of students per standardised pedagogue and also the quality of the pedagogic apparatus are as follows (see Figure 8).

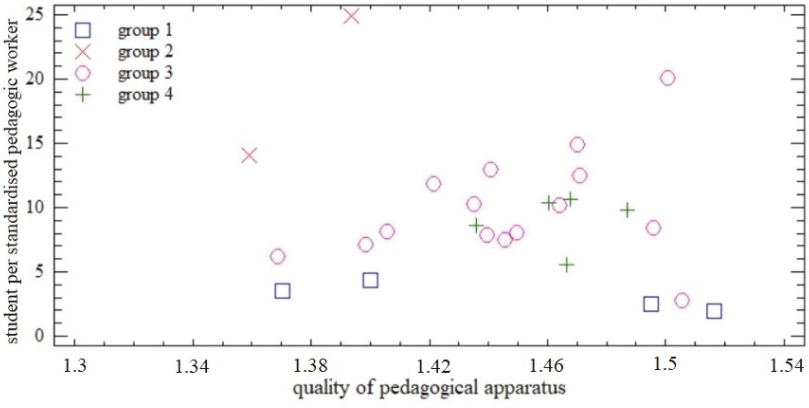

Figure 8. Quality of the pedagogic apparatus vs. number of students per standardised pedagogic worker according to the MEYS groups in academic year 2018/2019.

The results illustrate a considerable overlap regardless of the categorisation of the university. Art universities (the first group) can be characterised by a smaller number of students per standardised pedagogic worker. On the contrary, nonuniversity colleges (the second group) have a greater number of students. Within the third group of smaller universities, we observe differences more in the number of students than in the quality of the pedagogic apparatus. The results of larger schools (the fourth group) do not de facto differ from the results of the preceding group. If both quality indicators are taken into account simultaneously, without taking into account the PU group, the order is as follows (Table 3).

Table 3. Comparison of the quality of UP using TOPSIS techniques in the academic years 2011/2012 and 2018/2019.

	2011/2012			2018/2019		
Rank	**MEYS**	**University**	**C_i**	**MEYS**	**University**	**C_i**
1.	1	JAMU	1	1	JAMU	1
2.	1	AMU	0.992884	1	AMU	0.972143
3.	3	VŠCHT	0.985846	3	VŠCHT	0.960658
4.	1	AVU	0.977419	1	AVU	0.914956
5.	1	UMPRUM	0.973456	1	UMPRUM	0.888070
6.	4	ČVUT	0.971903	4	ČVUT	0.841761
7.	3	TUL	0.968333	3	VFU	0.809662
8.	4	UK	0.962300	3	TUL	0.770473
9.	4	UP	0.962186	3	OU	0.754386
10.	3	JU	0.961192	3	ZČU	0.737655
11.	3	OU	0.955287	3	JU	0.732830
12.	3	ZČU	0.954367	3	UPa	0.725362
13.	4	VUT	0.953012	3	VŠB-TUO	0.717977
14.	3	VŠB-TUO	0.951619	4	UP	0.708175

Table 3. *Cont.*

	2011/2012			2018/2019		
Rank	**MEYS**	**University**	**C_i**	**MEYS**	**University**	**C_i**
15.	3	MEN	0.950940	4	UK	0.655043
16.	3	UPa	0.949021	3	UJEP	0.641483
17.	3	UJEP	0.945110	3	MEN	0.637044
18.	4	MU	0.940218	4	MU	0.632998
19.	3	UHK	0.937717	4	VUT	0.621358
20.	3	SU	0.932894	3	SU	0.567808
21.	3	VŠE	0.931311	3	UHK	0.539711
22.	3	UTB	0.924072	3	UTB	0.517417
23.	3	ČZU	0.917987	2	VŠPJ	0.467984
24.	2	VŠPJ	0.902251	3	VŠE	0.434042
25.	2	VŠTE	0.842624	3	ČZU	0.213620
26.	3	VFU	0	2	VŠTE	0.011005

C_i—relative distance to PIS alternative (result of TOPSIS technique).

The Janáček Academy of Music and Performing Arts in Brno (JAMU) can be identified as the best-rated UP in the first and last assessed years, followed by the Academy of Performing Arts in Prague (AMU) and the University of Chemistry and Technology (VŠCHT). The highest quality of the pedagogic process is therefore observed in the first group of art UP. On the contrary, from the viewpoint of the second assessed indicator, the quality of the pedagogic process is the lowest at nonuniversity public colleges. The greatest shift in the order of placement can be attributed to the University of Veterinary and Pharmaceutical Sciences Brno (VFU).

Looking back, we can observe an increase in differences between individual UP, which is accompanied by a reduction in the skew of the acquired results ($\beta_{2011} = -4.878$; $\beta_{2018} = -1.113$). If we viewed the relationship between these quality indicators at the level of individual groups of UP, the results would be as follows (Table 4).

Table 4. Results of order linear correlation of quality indicators (Q1, Q2).

	Group 1	**Group 2**	**Group 3**	**Group 4**
r_K	−0.43 *	0.19	−0.01	0.17

* Significant at the level of importance $\alpha = 0.05$.

The quality of the pedagogic apparatus increases the smaller the number of students, i.e., with more time to devote to each student. There is also the opportunity for personal development, publication, or project activities. In other UP groups, these two indicators of quality do not correlate in a linear way. The assumption for H3 was therefore confirmed only in the case of art UP (the first group), i.e., we can disprove H3.

6. Discussion and Conclusions

Šebková et al. [45] stated that the "quality" of university education has been mentioned for several decades in a number of countries as a key priority and should lead to further development of university systems. Chvátalová et al. [46] stated that the quality of university education in the Czech Republic is monitored by a number of international organisations, which is also evidenced by the frequently published rankings of universities. These usually differ from each other by the number of assessed criteria, which subsequently corresponds to the different results in placement of the assessed institutions (e.g., QS, ARWU, or THE assessment). Other organisations monitoring "quality" include, for example, the International Network for Quality Assurance Agencies in Higher Education (INQAAHE), established in 1991, or the European Network for Quality Assurance in Higher Education (ENQUA) [47].

In his work, Neave [48] introduced the useful term conditional autonomy, which we understand to mean the conditions that enable a university to achieve the expected output level of graduates or enable a comparison of the number of workers compared to other parameters and to the nationwide standard. However, what is the nationwide standard? Is it developing? Is it good that it is developing? In their work, Šebková et al. [49] stated that there are a number of different views of the concept of "quality". According to them, quality can be defined as:

- perfection, excellence, the effort to be the best;
- compliance of a product with the defined standard;
- suitability for a specific purpose, eligibility for a defined purpose;
- threshold, i.e., fulfilment of at least the minimum defined standards;
- improvement of the monitored parameters, or growth of the institution.

How to appropriately define the quality of work of universities is quite a complex matter and depends on the point of view of the specific author. Stes et al. [50] stated that evidence of the impact of professional development of pedagogues on the quality of tuition of students is rare in university education. In his work, he stated that the quality of tuition by a pedagogue is not dependant on the size of the class or the number of students. Martensson et al. [12] stated that growing pressure on increasing "quality" by means of increasing the qualifications of pedagogues creates an unhappy gap between the formal rules of university institutions and routines and everyday procedures in the academic sphere, which are linked to teaching and learning.

The authors of this paper incline towards the long-term horizon of measuring the quality of university education according to De Weert [51], who was inspired by other sectors and defined three aspects of monitoring quality: quality of inputs (finance, qualification structure of the teaching body, spatial and technical equipment of the school, administrative processes, the quality of the admitted students); quality of processes (procedures by the management to achieve the set goals); and quality of outputs (compliance of individual goals with long-term and strategic goals), which will be assessed in other works. The objective of this paper was to point out one of the possible methods for assessing "quality" under the conditions of public universities in the Czech Republic. On the level of individual UP groups, differences in the number of students were confirmed (H1), but not in the different quality of the pedagogic apparatus (H2). However, there are differences between individual universities, as evidenced by Table 2 and the chart of universities for academic years 2011/2012 and 2018/2019. The table shows a shift in the placement of some universities over time that does not correspond to the order of universities in terms of international comparison according to the QS, ARWU, or THE assessment.

According to the authors, it is essential to ensure a greater involvement of students in tuition, innovate forms of tuition, and actively support all forms of discussion and interconnection of theory and practice, in order to improve the assessment of tuition from the viewpoint of students' perceptions. It is also important to support and systematically create an environment of mutual trust between all elements of the assessed institution (students, pedagogues, and management) and to realise various types of assessment (not just accreditation or science, but also tuition) and to evaluate this in an adequate manner. From our point of view, smaller study groups can be considered a "suitable climate for tuition and discussion". With the exception of the group of nonart UP, we must, however, state that smaller groups of students do not lead to a higher quality of tuition (H3). Discussion in tuition is also promoted more by a greater number of docents and professors, from whom we can expect more experience in how to connect theory with practice. These individuals should represent versatile personalities in the field of their speciality. The higher professional qualification of a university pedagogue is linked to scientific research activities and creative activities in the form of publishing in important magazines with an impact factor and the submission, management, and resolution of research tasks and projects. The management of the university or the faculty is responsible for assuring these prerequisites. They should endeavour to support the growth of the qualification structure

of their employees, on the one hand, and a regular active communication (discussion) with students, not only of the methods and form of tuition, but also of their comments and suggestions for improvement of the "quality" of tuition, on the other hand. In the long term, not just research activities, which are better expressed quantitatively and qualitatively than the assessment of the quality of pedagogic activities, should be assessed and evaluated regularly at public universities in the Czech Republic. The quality of tuition is discussed, but its objective evaluation is not very successful. Teaching activities at a number of public universities in the Czech Republic are left in the "shadow" of research activities, and there are tendencies by individual universities to categorise themselves as just teaching universities or research universities, depending on the reported publication activities. This is also partially confirmed by the classification of public universities by MEYS into four categories, according to which public universities were assessed above within the terms of this paper.

Author Contributions: Conceptualisation, R.V. and J.B.; methodology, R.V.; software, R.V.; validation, R.V.; formal analysis, R.V.; investigation, R.V. and J.B.; resources, J.B.; data curation, J.B.; writing—original draft preparation, R.V. and J.B.; writing—review and editing, R.V.; visualisation, R.V.; supervision, R.V.; project administration, J.B.; funding acquisition, J.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Student Grant Competition in VŠB—Technical University of Ostrava, grant number SP2021/18.

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

References

1. Terhart, E. Qualität und Qualitätssicherung im Schulsystem: Hintergründe—Konzepte—Probleme. *Zeit. Pädag.* **2000**, *41*, 809–829.
2. Fend, H. Qualität und Qualitätssicherung im Bildungsbereich. Schule, Sozialpadagogik, Hochschule. *Zeit. Pädag.* **2000**, *41*, 56–71.
3. Carroll, J.B. A model of school learning. *Teach. Coll. Rec.* **1963**, *64*, 723–733.
4. Bloom, B.S. *Human Characteristics and School Learning*; McGraw-Hill: New York, NY, USA, 1976; pp. 1–17.
5. Einsiedler, W. Das Konzept "Unterrichtsqualität". *Unterrichtswissenschaft* **2000**, *30*, 194–196.
6. Harvey, L.; Green, D. Defining quality. *Assess. Eval. High. Educ.* **1993**, *18*, 9–34. [CrossRef]
7. Brophy, J.E.; Good, T.L. Teacher behavior and student achievement. In *Handbook of Research on Teaching*; Wittrock, M.C., Ed.; Macmillan: New York, NY, USA, 1986; pp. 328–375.
8. Wang, C.M.; Haertel, D.G.; Walberg, H.J. Toward a knowledge base for school learning. *Rev. Educ. Res.* **1993**, *63*, 249–294. [CrossRef]
9. Seidel, T.; Shavelson, J.R. Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Rev. Educ. Res.* **2007**, *77*, 454–499. [CrossRef]
10. Janík, T. O potřebě integrace myšlenkových snah zaměřených k profesionalizaci učitelství: Diskuse na okraj monotematického čísla pedagogiky. *Pedagogika* **2010**, *60*, 329–341.
11. Weinert, E.F.; Schrader, W.F.; Helmke, A. Quality of instruction and achievement outcomes. *Int. J. Educ. Res.* **1989**, *13*, 895–914. [CrossRef]
12. Mårtensson, K.; Roxå, T.; Stensaker, B. From quality assurance to quality practices: An investigation of strong microcultures in teaching and learning. *Stud. High. Educ.* **2014**, *39*, 534–545. [CrossRef]
13. Rýdl, K. K pojetí kvality ve školství a jejímu hodnocení. *e-Pedagogium* **2002**, *2*, 104–123.
14. Starý, K.; Chvál, M. Kvalita a efektivita výuky: Metodologické přístupy. In *Výzkum Výuky: Tematické Oblasti, Výzkumné Přístupy a Metody*; Janíková, M., Vlčková, K., Eds.; Paido: Brno, Czech Republic, 2009; pp. 63–81.
15. Průcha, J. *Pedagogická evaluace*; Masaryk University: Brno, Czech Republic, 1996; pp. 1–23.
16. Jürgens, E. Was ist guter Unterricht aus der Perspektive "der" Reformpädagogik? Vom Aktivitätsparadigma zum "Schüleraktiven Unterricht". In *Was Ist "Guter" Unterricht?* Jürgens, E., Standop, J., Eds.; Klinkhardt: Bad Heilbrunn, Germany, 2010; pp. 39–82.
17. Pupala, B. Kvalita školy v spektre odlišných politických hodnôt a záujmových skupín. *Škol. Psych.* **2010**, *12*, 5–10.
18. Mincer, J. Education and unemployment. *Educ. Unempl.* **1991**, 3838. [CrossRef]
19. Purkey, C.; Smith, S.; Marshall, S. Effective schools: A review. *Elemen. Sch. J.* **1983**, *83*, 427–452. [CrossRef]
20. Jackson, W.P. *Life in the Classroom*; Teacher College Press: New York, NY, USA, 1990; pp. 20–42.
21. Kurt, A. *Gute Schulen—Worauf Beruht ihre Wirksamkeit?* Klinkhardt: Bad Heilbrunn, Germany, 1990; pp. 10–44.
22. Steffens, U.; Bargel, T. *Erkundungen zur Qualität von Schule*; Luchterhand: Berlin, Germany, 1993; pp. 30–51.
23. Pol, M. *Škola v Proměnách*; Masaryk University: Brno, Czech Republic, 2007; pp. 22–33.
24. Pol, M.; Hloušková, L.; Novotný, P.; Zounek, J. *Kultura Školy. Příspěvek k Výzkumu a Rozvoji*; Masaryk University: Brno, Czech Republic, 2005; pp. 1–19.

25. Blackmur, D. A critical analysis of the INQAAHE Guidelines of Good Practice for higher education quality assurance agencies. *High. Educ.* **2008**, *56*, 723–734. [CrossRef]
26. Trnková, K.; Knotová, D.; Chaloupková, L. *Málotřídní školy v České republice*; Paido: Brno, Czech Republic, 2010; pp. 1–12.
27. Dvořák, D.; Starý, K.; Urbánek, P.; Chvál, M.; Walterová, E. *Česká Základní Škola. Vícepřípadová Studie*; Karolinum: Prague, Czech Republic, 2010; pp. 22–45.
28. Bloch, C.; Degn, L.; Nygaard, S.; Haase, S. Does quality work work? A systematic review of academic literature on quality initiatives in higher education. *Assess. Eval. High. Educ.* **2020**, *46*, 701–718. [CrossRef]
29. Mulder, R.; Baik, C.; Naylor, R.; Pearce, J. How does student peer review influence perceptions, engagement and academic outcomes? A case study. *Assess. Eval. High. Educ.* **2013**, *39*, 657–677. [CrossRef]
30. Berezvai, Z.; Lukats, G.D.; Molontay, R. Can professors buy better evaluation with lenient grading? The effect of grade inflation on student evaluation of teaching. *Assess. Eval. High. Educ.* **2020**, *46*, 793–808. [CrossRef]
31. Stewart, M. Making sense of a teaching programme for university academics: Exploring the longer-term effects. *Teach. Teach. Educ.* **2014**, *38*, 89–98. [CrossRef]
32. Černikovský, P. Posedlost kvalitou a meze zlepšování: 10 tezí k diskusi. In Proceedings of the Konference Vysoké Školství a Česká Společnost v Minulých 100 Letech a Jak Dál, Prague, Czech Republic, 1 November 2018.
33. Cardoso, S.; Rosa, M.J.; Stensaker, B. Why is quality in higher education not achieved? The view of academics. *Assess. Eval. High. Educ.* **2015**, *41*, 950–965. [CrossRef]
34. National Accreditation Bureau for Higher Education of the Czech Republic. Zpráva o Vnitřním Hodnocení NAÚ za Období 2016–2019. Available online: https://www.nauvs.cz/attachments/article/151/RNAU-schvaleno-2019-ZpravaOVnitrnimHodnoceniNAU.pdf (accessed on 15 March 2021).
35. Tesar, G.; Hodnocení Kvality Vysokých škol. Annual Meeting Centre of Higher Education in the Czech Republic. Available online: https://www.csvs.cz/ (accessed on 5 May 2021).
36. Kember, D. Promoting student-centred forms of learning across an entire university. *High. Educ.* **2008**, *58*, 1–13. [CrossRef]
37. Bovill, C.; Cook-Sather, A.; Felten, P. Students as co-creators of teaching approaches, course design, and curricula: Implications for academic developers. *Int. J. Acad. Dev.* **2011**, *16*, 133–145. [CrossRef]
38. Ministry of Education, Youth and Sport of the Czech Republic. Rozpočet Kapitoly MŠMT. Available online: https://www.msmt.cz/vzdelavani/skolstvi-v-cr/ekonomika-skolstvi/rozpocet-kapitoly-msmt (accessed on 2 February 2021).
39. Ministry of Education, Science, Research and Sport of the Slovak Republic. Metodika rozpisu dotácií zo štátneho rozpočtu verejným vysokým školám. Available online: https://www.minedu.sk/rozpis-dotacii-zo-statneho-rozpoctu-verejnym-vysokym-skolam-na-rok-2020/ (accessed on 5 February 2021).
40. OECD. *Education at a Glance 2020: OECD Indicators*; OECD Publishing: Paris, France, 2020.
41. Ministry of Education, Youth and Sport of the Czech Republic. Registr Vysokých škol a Uskutečňovaných Studijních Programů. Available online: https://regvssp.msmt.cz/registrvssp/csplist.aspx (accessed on 22 February 2021).
42. Ministry of Education, Youth and Sport of the Czech Republic. Závěrečný účet Kapitoly MŠMT. Available online: https://www.msmt.cz/vzdelavani/skolstvi-v-cr/ekonomika-skolstvi/szu (accessed on 27 February 2021).
43. Vavrek, R.; Bečica, J. Capital city as a factor of multi-criteria decision analysis—Application on transport companies in the Czech Republic. *Mathematics* **2020**, *8*, 1765. [CrossRef]
44. Vavrek, R. Evaluation of the impact of selected weighting methods on the results of the TOPSIS technique. *Int. J. Inf. Technol. Decis. Mak.* **2019**, *18*, 1821–1843. [CrossRef]
45. Šebková, H.; Kohoutek, J.; Šturzová, J. Metodika komplexního hodnocení kvality. *Aula* **2005**, *13*, 110–125.
46. Chvátalová, A.; Kohoutek, J.; Šebková, H. *Zajišťování kvality v Českém vysokém školství*; Aleš Čeněk: Plzeň, Czech Republic, 2009; pp. 1–22.
47. European Network for Quality Assurance in Higher Education. Available online: https://enqa.eu/ (accessed on 5 February 2021).
48. Neave, G. On the cultivation of quality, efficiency and enterprise: An Overview of recent trends in higher education in Western Europe, 1986–1988. *Eur. J. Educ.* **1988**, *23*, 7–23. [CrossRef]
49. Šebková, H.; Munsterová, E. Akreditace a hodnocení kvality. *Aula* **2005**, *13*, 14–22.
50. Stes, A.; Coertjens, L.; van Petegem, P. Instructional development for teachers in higher education: Impact on teaching ap-proach. *High. Educ.* **2009**, *60*, 187–204. [CrossRef]
51. De Weert, E. A macro-analysis of quality assessment in higher education. *High. Educ.* **1990**, *19*, 57–72. [CrossRef]

education sciences

MDPI

Article

Effect on Procrastination and Learning of Mistakes in the Design of the Formative and Summative Assessments: A Case Study

Fidel Salas Vicente *, Ángel Vicente Escuder, Miguel Ángel Pérez Puig and Francisco Segovia López

Mechanical al Materials Engineering Department, Universitat Politècnica de València, Camino de Vera, s/n, 46022 Valencia, Spain; avicente@mcm.upv.es (Á.V.E.); mipepui@mcm.upv.es (M.Á.P.P.); fsegovia@mcm.upv.es (F.S.L.)

* Correspondence: fisavi@doctor.upv.es; Tel.: +34-963-877-628

Abstract: The design of the formative and summative assessment processes is of paramount importance to help students avoid procrastination and guide them towards the achievement of the learning objectives that are described in the course syllabus. If the assessment processes are poorly designed the outcome can be disappointing, including high grades but poor learning. In this paper, we describe the unexpected and undesirable effects that an on-demand formative assessment and the timetable of a summative assessment that left the most cognitively demanding part, problem-solving, to the end of the course, had on the behavior of students and on both grading and learning. As the formative assessment was voluntary, students procrastinated till the last minute. However, the real problem was that due to the design of the summative assessment, they focused their efforts mainly on the easiest parts of the summative assessment, passing the course with ease, but achieving a low learning level, as evidenced by the low scores of the problem-solving part of the summative assessment.

Keywords: formative assessment; summative assessment; learning; procrastination

Citation: Salas Vicente, F.; Escuder, Á.V.; Pérez Puig, M.Á.; Segovia López, F. Effect on Procrastination and Learning of Mistakes in the Design of the Formative and Summative Assessments: A Case Study. *Educ. Sci.* **2021**, *11*, 428. https://doi.org/10.3390/educsci11080428

Academic Editors: Sandra Raquel Gonçalves Fernandes, Marta Abelha and Ana Teresa Ferreira-Oliveira

Received: 4 June 2021
Accepted: 11 August 2021
Published: 12 August 2021

1. Introduction

The syllabus of a course includes, among other things, the learning objectives, a description of the assessment procedures and the means, including the formative assessment, the students are supposed to use to reach the established objectives. This guide is at the disposal of the students since the beginning of the course and, if properly followed, it is supposed to be a priceless tool for the students in their learning path. However, to reach the learning objectives, the assessment processes and the means at the disposal of students must be carefully chosen and used along the course.

Therefore, one of the most important actions teachers can take is to design cleverly the formative and summative assessment processes [1–3]. The first one will provide the students with feedback and useful information about their degree of understanding of the studied matter [4,5] so they can have a predictor of the outcome [6,7] and know where to direct their efforts, leading to a self-regulated learning [8,9]. The second one will provide a formal certification of the acquired knowledges and skills.

To design the assessment procedures, it must be taken into account that a good grade is not always synonymous with a good understanding of the matter. Usually, the main aim of students is to pass the exam optimizing the use of the resources at their disposal [10,11], of which time is the most important and the one that must be reduced to liberate hours that will be used in the extensive social life associated with their age. The main aim of the teacher is to help students in their learning process and that usually means getting the students to spend, in a continuous and well-organized way, as many resources as necessary in their education, as that will guarantee a good outcome. Combining both views is not an easy task [12,13] and calls for a good knowledge of the behavior of students

and how and when they will use the resources at their disposal. Lack of experience or unexpected mistakes can spoil easily all the work and, worse still, the consequences can even go unnoticed.

Part of the problem can be caused using automated online tools, mainly based on quizzes, which on the other hand also have multiple and great advantages [14]. In fact, online quizzes are displacing teachers in the task of marking written homework and in giving students the necessary feedback. Furthermore, this change has been accelerated worldwide due to the changes and restrictions imposed due to the COVID-19 pandemic.

Despite this change seems desirable and unstoppable, it has some disadvantages. Two of them are:

- Teachers can feel their involvement in the learning process of students is less and less necessary. The possibilities online quizzes offer can lead to the design of on-demand and automated formative assessment systems based on online quizzes where the participation of the teacher is very limited. This will not be a problem for the more motivated and skilled students, but can be a problem for the rest of them as no one with experience is helping them during the course.
- Quizzes and other online tools are not very well suited for time-consuming and cognitive-demanding activities like solving numerical problems. This does not mean a quiz cannot be used to test high-level cognitive skills [15], but if a quiz is defined as a set of short questions that must be answered in 1 or 2 min, essays and time-consuming questions that require a long calculation process are inevitably ruled out.

The first disadvantage leads immediately to procrastination or avoidance if completing the formative assessment is not mandatory, but voluntary.

Procrastination, which can be defined as an irrational tendency to postpone performing a task is recognized as one of the main threats to student performance [16,17], leading not only to a decrease in the quality and quantity of learning, but also to an increase of stress[18] that affects even private life. The causes of procrastination in an academic context have been studied during recent decades and include variables such as internet addiction [19], self-esteem, perfectionism, planning skills, extroversion, emotional intelligence and self-efficacy [18], although self-efficacy and procrastination are, by definition, antonyms and mutually exclusive. Nevertheless, the main causes of procrastination during the young age could be related to the inevitable preference of students for socializing and enjoying life and the absence of a pressing need to earn a living. The immediate reward associated with socializing and fun activities is hardly comparable to the distant payoff of working today in something that can be done tomorrow [20].

Another problem is that although students can receive the correct answer and even an explanation from the automated system, there is no guarantee that they will understand it, and as the interaction with the teacher is reduced (not on purpose), they could be more reluctant to ask for support, something that is anyway a widespread problem [21], even when encouraging them actively ameliorates the problem [22].

The second disadvantage does not mean online quizzes are not pertinent. As a matter of fact, online quizzes are a highly demanded option by students and seems to improve their summative assessment [23–26], although there is no advantage with regard to written quizzes [27] and some authors doubt whether doing online quizzes improves the final marks or, simply, the more skilled and willing students do more quizzes [28,29]. Some even doubt if in some cases the relation exists or if it is positive [30,31], possibly due to the detrimental effect of overconfidence [6]. That said, if complex and time-consuming questions are ruled out, it will be difficult for the students to reach a deep understanding of the studied matter.

This paper presents a study of the causes of the undesirable and unexpected consequences changes in the formative and summative assessment procedures had on the learning of the students of a materials science course and why they almost went undetected.

The changes, part of them requested by the students in previous years, consisted of changing from a weekly summative assessment that was also used as formative assess-

ment to an online, automated and on-demand system-based mainly in quizzes for the formative assessment and in using 2 online exams based on quizzes and one last written exam aimed at evaluating the numerical problem-solving skills of students for the summative assessment.

The new assessment procedures led to an increase in the scores obtained in the quizzes part of the summative assessment but also to a significant drop in the scores obtained in the written exam. Most of the students passed the course but with a low grade due to the low scores of the written exam that required a deeper knowledge, calculations and a slower and more thoughtful work.

The problem was the formative and summative assessment were poorly designed to force the activation of self-regulated learning, or rather, self-regulation consisted of a shortcut that avoided the effort of a profound learning. The identified mistakes give a word of advice against assessment procedures that do not prevent procrastination, allow students to pass the course completing only part of the assignments or only evaluate part of the skills students should have acquired.

2. Materials and Methods

2.1. Changes in Assessment Procedures

"Materials" is one of the required courses students must take during their second year of studies towards a bachelor's degree in Industrial Design Engineering and Product Development at the Valencia Polytechnic University in Spain. This course, with near 150 students enrolled every year, is divided into 14 chapters and the minimum score to pass the course is 5 out of 10. The course extends over 15 weeks.

The summative assessment procedure was changed for the 2019/2020 course to satisfy the demands of part of the students that expressed that a continuous summative assessment was too demanding and stressful, leaving less time that desirable for other courses. In fact, this feeling has also been observed by other authors [32,33] and affects teachers [34]. The old grading was:

- One online quiz with 10 questions for each chapter except the first one: 15%.
- One online problem for each chapter except the first one (by problem we mean time-consuming questions that require a numerical calculation process). The solution was sent scanned or photographed: 15%.
- Two written exams with 10 problems, one mid-semester and another one at the end of the course: 55%.
- Lab reports and post-lab questions and three online quizzes about the lab sessions: 15%
- There is no minimum passing grade for the different parts of the assessment.

The last rule is part of the assessment code established by the Higher Technical School of Design Engineering where the course is taught.

Regarding the formative assessment, the online quizzes and problems were supposed to give enough feedback to the students, as one was done almost every week.

The new grading was:

- Two online quizzes, one mid-semester and another one at the end of the course: 35%.
- One on-site written exam with 10 numerical problems at the end of the course: 40%.
- Lab reports and post-lab questions: 25%
- There is no minimum passing grade for the different parts of the assessment.

In this case, the students had at their disposal a series of online quizzes (10 questions each one) and some online problems involving mathematical calculations from the beginning of the course. The questions for the quizzes were randomly taken from a database that contains between 30 and 110 questions for each chapter. There was no time limit, and the students could take as many quizzes as wished. Once the quiz was finished, the students received their score immediately along with the correct answers. The online formative problems were scored automatically by comparing the numerical result introduced by the

students with the correct one, allowing some margin of error, but no description of the correct calculation process was given. This way, feedback was provided by the system without the intervention of the teacher unless students asked for a tutoring meeting. The questions for the summative quizzes were not taken from the same database than the questions for the formative quizzes.

The change was supposed to provide students with a means to obtain direct and continuous formative assessment during the course, but, despite previous years, when the quizzes were mandatory and part of the summative assessment, its use was voluntary and dependent only on the behavior and will of students.

The consequences of the changes showed there was a fault in the design of the assessment procedures that should be investigated. This fault was supposed to be related to how the students adapted their behavior to the new formative and summative assessment.

2.2. Participants

During the year of study, 135 students enrolled the course, 48% male and 52% female. Apart from gender, the students form a homogeneous group of mainly Caucasian Spanish students with a medium-high income and cultural level. Only 2 students were South America natives and other 2 were from North Africa. Those quantities are not high enough to consider an investigation on the student behavior based on ethnicity.

Except 1 north African student, all of them had studied in Spain since childhood and even this student had been living in Spain for 8 years. Therefore, in general, no language difficulties were expected. Additionally, no disabilities were reported by any student.

2.3. Methodology

The effects of the changes in the assessment procedures were evaluated by compared to the mean summative scores obtained by the students in the previous years, mainly, the year before the change, although that year is a good representative of all of them. As some consequences of the changes were undesirable, the changes are not going to be maintained any more. Therefore, there is not the possibility of obtaining new data from the scores of another year with the same assessment procedures, what could confirm the findings presented in this paper. Nevertheless, the differences in the problem-solving part of the summative assessment were considered big enough to be credible and a consequence of the changes. Additionally, once the behavior of students and the assessment timetable was studied, the consequences seem logical.

To study the behavior of students and disclose the relationship between that behavior and learning, we studied the data collected by the learning management system of the university, based on Sakai. The data available is divided in two databases: The first one, automatically collected, includes identification of the student, start and finish time for each formative quiz or problem and scores. This data were available for all chapters except chapter 8 due to computer problems that led to the loss of that data during the onset of the COVID-19 pandemic. The second part included identification of the students and scores of the summative assessment, which was introduced manually by the teachers. After retrieval, the data were analyzed using macros programmed in MS Excel. The students' ID served as a link to relate both databases.

As the changes consisted mainly of an on-demand formative assessment and of changes in the summative assessment, the relation between them and summative scores were analyzed. The first step was to study the behavior of students with regards to the voluntary formative assessment. The number of formative quizzes and numerical questions taken per course week and per student were counted. This simple analysis provided a rough view of when and how much the students used the formative assessment along the course.

The consequences of the design of the formative assessment were studied by relating the amount and type (quizzes or problems) of formative work done by the students along the course with the scores of the different parts of the summative assessment. The tem-

poral distribution of the work done showed a direct relationship with the summative assessment, what provided the main clue to find the most critical mistakes of the new assessment procedures.

3. Results and Discussion

132 students took a least one quiz or solved a numerical problem. In total, 20,358 quizzes and 1205 problems were done (excluding the ones concerning chapter 8 of the course), which implies each student did a mean of 154 quizzes and 9.1 problems, not even one per chapter. Figure 1 gives a better perspective on how many formative quizzes and problems did each student. Almost 60% of them did less than the mean value of quizzes (154 quizzes per student) while 7.5% students did more than 300 quizzes in total.

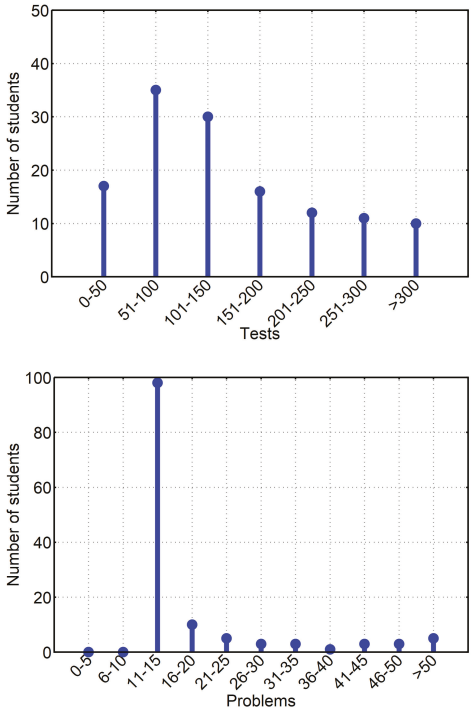

Figure 1. Number of students that have done a certain number of tests and problems.

There was a slight difference in the number of tests done between male and female students, while female students did a mean of 165.6 tests along the course, male students did a mean of 141.3 tests. On the other hand, male students did a mean of 9.9 numerical problems while female students did a mean of 8.3 numerical problems. The difference is not big, but seems to indicate female students are more conscious of what parts of the summative assessment are more important to pass the course and, so, they prepare for them with a bit more intensity.

The study of the time students used to complete the quizzes show that 65% of the quizzes were finished in 5 min or less, while 30% were finished the quizzes in 2 or less minutes. The commonly accepted value is 1 min per question [35], with being verified that longer times could lead to a decrease in performance in terms of scores [36]. This means the quizzes should be answered in around 10 minutes.

One of the potential explanations is that as the number of quiz questions in the database is 3 to 11 times the number of questions in the quizzes, the students used to repeat

the quizzes time and again in an attempt to view as many questions as possible in their attempt to prepare for the summative assessment. After the first quizzes, many questions repeat, what means the answer is known. If that is the case, the students can answer very quickly (they can even mark the correct answer without reading it if they recognize it visually). In any case, the time needed to complete the test is reduced considerably.

The number of numerical problems done by the students is much lower than that of quizzes. Not only were they available in less quantity, but required more time, were harder to solve and, overall, the capacity to solve such problems was not evaluated until the end of the course. The limited use of the available online numerical problems as a tool for formative assessment is unfortunate because problem-solving require a different set of skills and cognitive skills than common quizzes [37], where the work with data or formulas is very limited, and that work is important to reach all the goals listed in the syllabus.

Additionally important is to know if the effort of the students was distributed equally along the course, something that would show a high level of self-regulation, what is accompanied by a higher level of academic success [9]. Figure 2 shows that was not the case and that despite the great number of quizzes done; this kind of formative assessment was concentrated in the two weeks before the summative assessments.

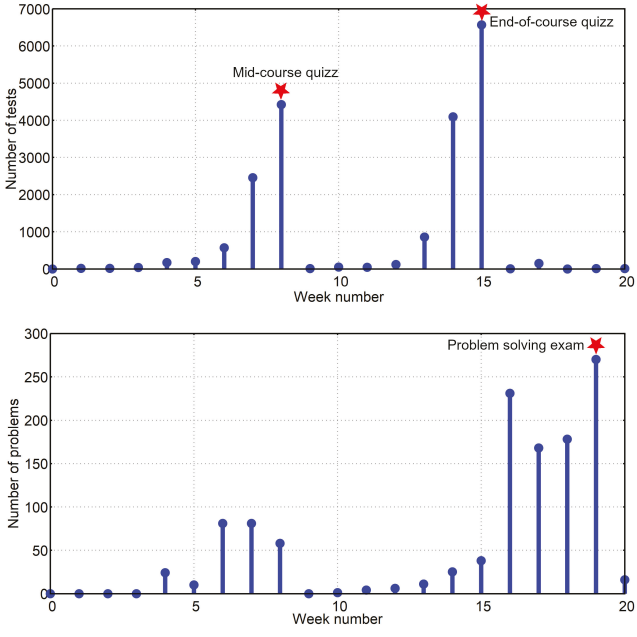

Figure 2. Distribution of the number of formative quizzes and problems done by week.

Although unfortunate, this is the typical behavior of students if effective measures to avoid academic procrastination are not implemented in the class [38]. The level of procrastination varies from 50 to 90% according to different authors [39,40] or even more [41]. In this case, if the level of procrastination is calculated as the number of quizzes done during the 2 weeks before each exam divided by the total number of quizzes, procrastination reaches nearly 90%. This is not how procrastination percentage is usually calculated (using self-reports about executive functioning [42]), but gives a good image of the problem because 4 weeks (two weeks before each quiz exam) accounts for around 23.5% of the course duration, including non-teaching weeks.

The effect of procrastination is usually stress [43] and a lower learning level [44] With these facts in mind, it is advisable to implement multiple deadlines along the course so

that the curve of learning and time dedicated to study changes significantly, as shown in Figure 3.

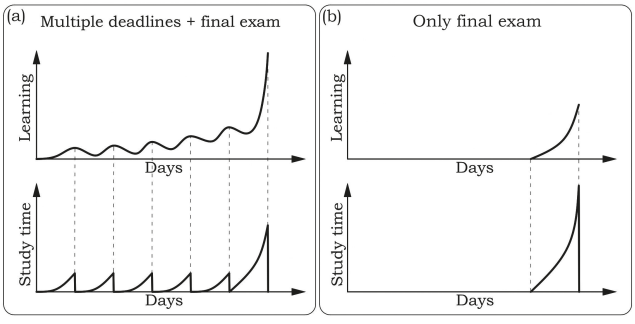

Figure 3. Evolution of learning (**a**) with and (**b**) without multiple deadlines along the course.

According to the model proposed in Figure 3, if students have several deadlines distributed along the course, even if they delay the work until the last minute, they will be forced to study before each deadline. Evidently, they are not studying for the final exam and the learning level could not be as high as desirable and some concepts will be partially forgotten as days go on, but not all of them and the scaffold that is created will help students gain a deeper, greater and faster understanding of the studied matter. With studied at a steady pace will also reduce the stress level and give them more confidence before the final exam.

Unfortunately, it seems self-imposed deadlines do not work too well [45], as they are easily delayed or cancelled. To be effective those deadlines should be somehow imposed by the teacher. This was one of the main problems of the design of the formative assessment. Students were not forced to take the quizzes after each lesson and, as a consequence, they did not take them until the days before the summative assessment.

Nonetheless, online formative quizzes were useful, even despite procrastination, as a means to prepare the students for the summative assessment. In fact, the mean score for the summative quizzes after the change was 7.33, while in previous years it ranged from 6.5 to 7.0 (in the preceding course it was 6.89).

The usefulness of formative quizzes is directly related to the amount of quizzes done by the students. Figure 4 shows how the mean number of right answers in the formative assessment evolves with the number of quizzes done per chapter. That figure shows clearly how the score obtained by students in the formative quizzes increase until a certain number of quizzes is reached. In this case, this number is 11 quizzes. After that threshold the profit of doing more quizzes is small. The mean score of the summative quizzes show a more continuous growth, without a saturation threshold. This is somewhat expected as after some formative quizzes the answers to the questions are known. The summative quizzes do not show this effect, but the outcome of the effort and time of study of each student.

The evolution of the mean score with the number of formative quizzes done is very similar for the two summative quizzes, with a continuous increase that boosts the mean score from 5.97 to 9.20 (out of 10) for the first summative assessment and from 5.44 to 8.62 for the second summative assessment. This accounts for a 54% and a 58% increment, respectively. Nevertheless, a slight decrease in the score seems to take place for students who have done more than 200 tests (about 18 quizzes per chapter). There is no simple explanation, but it could be these students do not know when to stop and dedicate time to other learning activities once they have done enough quizzes.

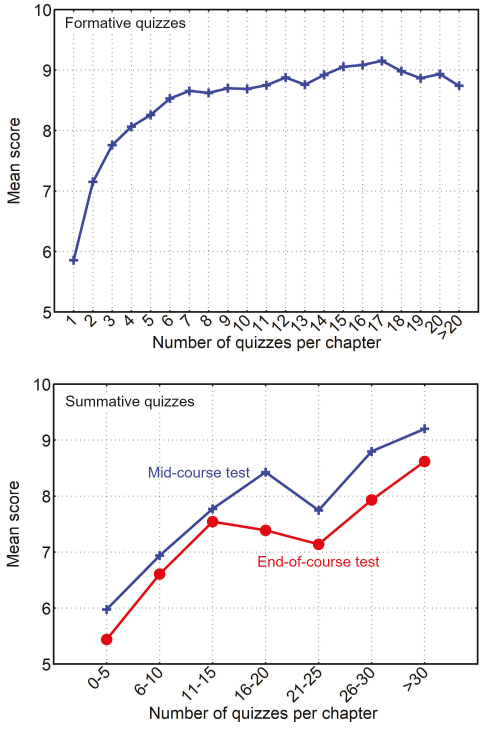

Figure 4. Evolution of the mean score in the formative (up) and summative quizzes (down) with the number of formative quizzes done.

The mean score for the formative assessment-based in problem-solving was 2.57 out of 10, although it must be taken into account that the score for each problem is 10 or 0 as only the final result is evaluated by comparing it with the correct value. Most of the students were unable to solve the problems correctly but they did not ask for advice or a tutoring session despite being encouraged to do so many times along the course and not with may possibilities of knowing where they had failed without the detailed solutions to the problems. This fact indicates a serious flaw in the behavior of the students.

The mean score for the summative numerical problems was 1.69. This score is usually the lowest of the different parts of the summative assessment (in the preceding course it was 4.30), but 1.69 is extremely low. In fact, only two students passed that exam, while in the previous year 35.5% passed that part of the assessment (not counting the second-chance exam).

Furthermore, concerning is that if the problem-solving exam had not been passed at the end of the course, the final grades would have shown an improvement compared to previous years and would have led to the conclusion that the changes implemented in the course were greatly successful in improving learning. This would have been an unfortunate mistake, as actually learning levels had decreased due to the fact that part of the educational goals was far from being attained.

Regarding the influence of gender, no significant differences were found, as Table 1 shows. The differences are very small and could be attributed to the quantity of formative quizzes and problems done by each gender.

Table 1. Gender differences in the summative mean scores obtained by female and male students.

Male		Female	
Quizzes	**Problems**	**Quizzes**	**Problems**
7.28	1.76	7.38	1.62

The change in the scores associated with the new assessment is not due to part of the students, but to a general change in student self-regulation that had to be studied and was not due to procrastination. This was the first year the students had online quizzes and problems at their disposal to use as formative assessment, so, an increase in scores was expected (as usual, written questions in the textbook were also available). This happened in the quizzes, but not in the numerical problems. Figure 5 shows this effect in a comparison with the scores of the prior course.

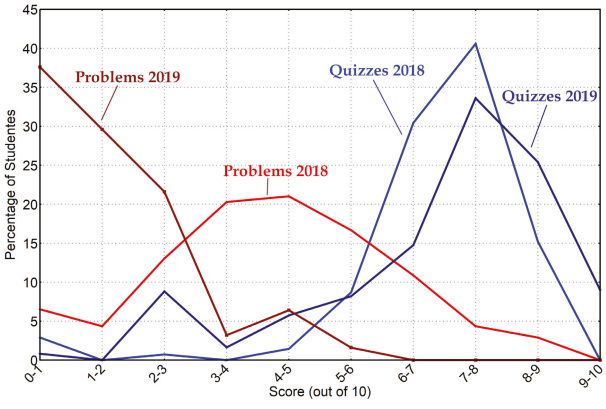

Figure 5. Percentage of students that have gotten a certain score in the summative quizzes and the numerical problems exam in the year of study and the previous one.

The teachers consider the problems were not different nor more difficult than the ones of previous years.

The cause of the low scores in problems lies in two facts:

- The low interest the students have had in doing that kind of numerical problems along the course. That is an issue, but one that has not affected the scores in summative quizzes.
- The formative goals related to the practical application of theory in problem-solving were deemed as secondary and not worthy of attention by the students.

This second fact is the most important one. The source of the preference of the students can be found in the timetable for the summative assessment and in the percentage of the final grade associated with each assessment activity. The problem-solving exam was done after the two summative quizzes had been done and graded. This means the students knew their scores for the 60% of the final grade (quizzes were automatically scored once they were finished and the score for the lab reports and post-lab questions, which are usually high, were also known). As the scores for the two already graded parts were high enough to assure a final grade over 5 and there was not a minimum passing score for the different parts of the assessment, most of the students knew they had passed the course before the problem-solving exam. Knowing this and with other exams to think about, they decided to devote very little time to prepare the last exam of the summative assessment. There was an obvious increment in the online activity before this last exam, but proportionally much lower than the activity associated with quizzes, and, it must be assumed that with far less interest.

Therefore, the mistakes in the design of the formative and summative assessment can be summarized as follows:

- No mechanism was implemented to avoid procrastination and assure the correct and frequent use of the formative assessment.
- The description of the correct calculation procedure for the problems was not given to the students.
- Problem-solving abilities were evaluated only one time at the end of the course, after the formative quiz tests.
- The students knew if they had passed the course before the problem-solving exam.
- There was no minimal passing grade for the problem-solving exam.

The study has one main limitation, and that is the fact that the data analyzed corresponds to only one year. Nevertheless, the differences in the scores obtained with respect to previous years are too big to think they are the result of pure chance.

According to the scores, procrastination was not the problem it could have been, but nonetheless, it is a behavior that should be avoided. The same cannot be said about the design of the summative assessment that certainly should be changed. Some proposals that should work are:

- The formative online quizzes and problems should remain available to the students, but their availability could be limited to 2 weeks after the corresponding theory has been studied in class. Students will have to do them in those 2 weeks or lose the opportunity. This should reduce considerably procrastination. It also could simply be made mandatory.
- Explanations about the correct answer should be given to the students for both quizzes and problems.
- The summative assessment could be changed to two written exams along the course (one mid-course and another one at the end of the course), both including quizzes and problems. This should also reduce procrastination and will assure no goals will be forgotten, while will reduce the continuous stress associated with the weekly summative assessment.
- All parts of the summative assessment should have a minimum score for the students to pass the course, although right now that is not possible.

Evidently, many other assessment procedures could be used in the course provided all the detected mistakes, which give a word of advice to teachers who are going to change their assessments, are corrected.

Author Contributions: Conceptualization, Á.V.E.; software, F.S.V.; validation, Á.V.E. and F.S.L.; formal analysis, F.S.V.; investigation, Á.V.E. and F.S.V.; resources, M.Á.P.P.; data curation, F.S.V.; writing—original draft preparation; writing—review and editing, M.Á.P.P.; visualization, M.Á.P.P.; supervision, M.Á.P.P. and F.S.L.; project administration, Á.V.E.; funding acquisition, none All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the study can be obtained from the authors.

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

References

1. Dick, W. Summative evaluation. In *Instructional Design: Principles and Applications*; Universidad de Salamanca: Salamanca, Spain, 1977; pp. 337–348.
2. Migueláñez, S.O. Evaluación formativa y sumativa de estudiantes universitarios: aplicación de las tecnologías a la evaluación educativa. *Teoría De La Educación. Educ. Y Cult. En La Soc. De La Inf.* **2009**, *10*, 305–307.

3. Martínez Rizo, F. Dificultades para implementar la evaluación formativa: Revisión de literatura. *Perfiles Educ.* **2013**, *35*, 128–150. [CrossRef]
4. Rolfe, I.; McPherson, J. Formative assessment: How am I doing? *Lancet* **1995**, *345*, 837–839. [CrossRef]
5. Iahad, N.; Dafoulas, G.; Kalaitzakis, E.; Macaulay, L. Evaluation of online assessment: The role of feedback in learner-centered e-learning. In Proceedings of the 37th Annual Hawaii International Conference on System Sciences, Big Island, HI, USA, 5–8 January 2004; doi:10.1109/hicss.2004.1265051. [CrossRef]
6. Bell, P.; Volckmann, D. Knowledge Surveys in General Chemistry: Confidence, Overconfidence, and Performance. *J. Chem. Educ.* **2011**, *88*, 1469–1476. [CrossRef]
7. Dobson, J.L. The use of formative online quizzes to enhance class preparation and scores on summative exams. *Adv. Physiol. Educ.* **2008**, *32*, 297–302. [CrossRef]
8. Paris, S.G.; Paris, A.H. Classroom Applications of Research on Self-Regulated Learning. *Educ. Psychol.* **2001**, *36*, 89–101._4. [CrossRef]
9. Zimmerman, B.J. Self-Regulated Learning and Academic Achievement: An Overview. *Educ. Psychol.* **1990**, *25*, 3–17._2. [CrossRef]
10. Meece, J.L.; Blumenfeld, P.C.; Hoyle, R.H. Students' goal orientations and cognitive engagement in classroom activities. *J. Educ. Psychol.* **1988**, *80*, 514. [CrossRef]
11. McCoach, D.B.; Siegle, D. Factors That Differentiate Underachieving Gifted Students From High-Achieving Gifted Students. *Gift. Child Q.* **2003**, *47*, 144–154. [CrossRef]
12. Bardach, L.; Yanagida, T.; Schober, B.; Lüftenegger, M. Students' and teachers' perceptions of goal structures—Will they ever converge? Exploring changes in student-teacher agreement and reciprocal relations to self-concept and achievement. *Contemp. Educ. Psychol.* **2019**, *59*, 101799. [CrossRef]
13. Rettinger, D.A.; Kramer, Y. Situational and Personal Causes of Student Cheating. *Res. High. Educ.* **2008**, *50*, 293–313. [CrossRef]
14. Marcell, M. Effectiveness of Regular Online Quizzing in Increasing Class Participation and Preparation. *Int. J. Scholarsh. Teach. Learn.* **2008**, *2*. [CrossRef]
15. Cox, K.; Clark, D. The use of formative quizzes for deep learning. *Comput. Educ.* **1998**, *30*, 157–168.
16. van Eerde, W. A meta-analytically derived nomological network of procrastination. *Personal. Individ. Differ.* **2003**, *35*, 1401–1418. [CrossRef]
17. Hayat, A.A.; Jahanian, M.; Bazrafcan, L.; Shokrpour, N. Prevalence of Academic Procrastination Among Medical Students and Its Relationship with Their Academic Achievement. *Shiraz E-Med. J.* **2020**, *21*. [CrossRef]
18. Hen, M.; Goroshit, M. Academic self-efficacy, emotional intelligence, GPA and academic procrastination in higher education. *Eurasian J. Soc. Sci.* **2014**, *2*, 1–10.
19. Aznar-Díaz, I.; Romero-Rodríguez, J.M.; García-González, A.; Ramírez-Montoya, M.S. Mexican and Spanish university students' Internet addiction and academic procrastination: Correlation and potential factors. *PLoS ONE* **2020**, *15*, e0233655. [CrossRef] [PubMed]
20. Dewitte, S.; Schouwenburg, H.C. Procrastination, temptations, and incentives: The struggle between the present and the future in procrastinators and the punctual. *Eur. J. Personal.* **2002**, *16*, 469–489. [CrossRef]
21. Ryan, A.M.; Pintrich, P.R.; Midgley, C. Avoiding Seeking Help in the Classroom: Who and Why? *Educ. Psychol. Rev.* **2001**, *13*, 93–114.:1009013420053. [CrossRef]
22. Pellegrino, C. Does telling them to ask for help work?: Investigating library help-seeking behaviors in college undergraduates. *Ref. User Serv. Q.* **2012**, *51*, 272–277. [CrossRef]
23. Pennebaker, J.W.; Gosling, S.D.; Ferrell, J.D. Daily Online Testing in Large Classes: Boosting College Performance while Reducing Achievement Gaps. *PLoS ONE* **2013**, *8*, e79774. [CrossRef]
24. Cantor, A.D.; Eslick, A.N.; Marsh, E.J.; Bjork, R.A.; Bjork, E.L. Multiple-choice tests stabilize access to marginal knowledge. *Mem. Cogn.* **2014**, *43*, 193–205. [CrossRef]
25. Roediger, H.L.; Agarwal, P.K.; McDaniel, M.A.; McDermott, K.B. Test-enhanced learning in the classroom: Long-term improvements from quizzing. *J. Exp. Psychol. Appl.* **2011**, *17*, 382–395. [CrossRef]
26. McDermott, K.B.; Agarwal, P.K.; D'Antonio, L.; Roediger, H.L.; McDaniel, M.A. Both multiple-choice and short-answer quizzes enhance later exam performance in middle and high school classes. *J. Exp. Psychol. Appl.* **2014**, *20*, 3–21. [CrossRef]
27. Anakwe, B. Comparison of Student Performance in Paper–Based Versus Computer–Based Testing. *J. Educ. Bus.* **2008**, *84*, 13–17. [CrossRef]
28. Johnson, G. Optional online quizzes: College student use and relationship to achievement. *Can. J. Learn. Technol.* **2006**, 32. [CrossRef]
29. Peat, M.; Franklin, S. Has student learning been improved by the use of online and offline formative assessment opportunities? *Australas. J. Educ. Technol.* **2003**, *19*. [CrossRef]
30. Merello-Giménez, P.; Zorio-Grima, A. Impact of students' performance in the continuous assessment methodology through Moodle on the final exam. In *EDUCADE—Revista de Educación en Contabilidad, Finanzas y Administración de Empresas*; 2017; pp. 57–68. [CrossRef]
31. Yin, Y.; Shavelson, R.J.; Ayala, C.C.; Ruiz-Primo, M.A.; Brandon, P.R.; Furtak, E.M.; Tomita, M.K.; Young, D.B. On the Impact of Formative Assessment on Student Motivation, Achievement, and Conceptual Change. *Appl. Meas. Educ.* **2008**, *21*, 335–359. [CrossRef]

32. Soe, M.K.; Baharudin, M.S.F. Association of academic stress & performance in continuous assessment among pharmacy students in body system course. *IIUM Med. J. Malays.* **2016**, *15*. [CrossRef]
33. Ghiatău, R.; Diac, G.; Curelaru, V. Interaction between summative and formative in higher education assessment: Students' perception. *Procedia Soc. Behav. Sci.* **2011**, *11*, 220–224. [CrossRef]
34. Trotter, E. Student perceptions of continuous summative assessment. *Assess. Eval. High. Educ.* **2006**, *31*, 505–521. [CrossRef]
35. Brothen, T. Time Limits on Tests. *Teach. Psychol.* **2012**, *39*, 288–292. [CrossRef]
36. Portolese, L.; Krause, J.; Bonner, J. Timed Online Tests: Do Students Perform Better With More Time? *Am. J. Distance Educ.* **2016**, *30*, 264–271. [CrossRef]
37. Litzinger, T.A.; Meter, P.V.; Firetto, C.M.; Passmore, L.J.; Masters, C.B.; Turns, S.R.; Gray, G.L.; Costanzo, F.; Zappe, S.E. A Cognitive Study of Problem Solving in Statics. *J. Eng. Educ.* **2010**, *99*, 337–353. [CrossRef]
38. Zacks, S.; Hen, M. Academic interventions for academic procrastination: A review of the literature. *J. Prev. Interv. Community* **2018**, *46*, 117–130. [CrossRef] [PubMed]
39. Özer, B.U.; Demir, A.; Ferrari, J.R. Exploring Academic Procrastination Among Turkish Students: Possible Gender Differences in Prevalence and Reasons. *J. Soc. Psychol.* **2009**, *149*, 241–257. [CrossRef]
40. Steel, P. The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychol. Bull.* **2007**, *133*, 65–94. [CrossRef]
41. Ellis, A.; Knaus, W. *Overcoming Procrastination*; Institute for Rational Living: New York, NY, USA, 1977.
42. Klingsieck, K.B. Procrastination. *Eur. Psychol.* **2013**, *18*, 24–34. [CrossRef]
43. Sirois, F.M. Procrastination and Stress: Exploring the Role of Self-compassion. *Self Identity* **2013**, *13*, 128–145. 15298868.2013.763404. [CrossRef]
44. Goda, Y.; Yamada, M.; Kato, H.; Matsuda, T.; Saito, Y.; Miyagawa, H. Procrastination and other learning behavioral types in e-learning and their relationship with learning outcomes. *Learn. Individ. Differ.* **2015**, *37*, 72–80. [CrossRef]
45. Bisin, A.; Hyndman, K. Present-bias, procrastination and deadlines in a field experiment. *Games Econ. Behav.* **2020**, *119*, 339–357. [CrossRef]

education sciences

MDPI

Article

Organizational Differences among Universities in Three Socioeconomic Contexts: Finland, Spain and Ecuador. Relational Coordination Approach

Cristina Checa-Morales [1,2,3,*], Carmen De-Pablos-Heredero [2,4], Angela Lorena Carreño [5], Sajid Haider [6] and Antón García [3]

[1] International Doctoral School (EID), Rey Juan Carlos University, 28008 Madrid, Spain
[2] Department of Business Economics (Administration, Management and Organization), Applied Economics II and Fundamentals of Economic Analysis, Rey Juan Carlos University, 28032 Madrid, Spain; carmen.depablos@urjc.es
[3] Animal Science Department, University of Cordoba, Rabanales University Campus, 14071 Cordoba, Spain; pa1gamaa@uco.es
[4] Area of Business Economics, ESIC University, 28223 Madrid, Spain
[5] Faculty of Administrative and Economic Sciences, Technical University of Manabí, Portoviejo 130105, Ecuador; angela.carreno@utm.edu.ec
[6] Department of Management Sciences, COMSATS University Islamabad, Vehari Campus, Vehari 61100, Pakistan; sajidhaider@ciitvehari.edu.pk
* Correspondence: c.checa.2019@alumnos.urjc.es

Citation: Checa-Morales, C.; De-Pablos-Heredero, C.; Carreño, A.L.; Haider, S.; García, A. Organizational Differences among Universities in Three Socioeconomic Contexts: Finland, Spain and Ecuador. Relational Coordination Approach. *Educ. Sci.* **2021**, *11*, 445. https://doi.org/10.3390/educsci11080445

Academic Editors: Sandra Raquel Gonçalves Fernandes, Marta Abelha and Ana Teresa Ferreira-Oliveira

Received: 22 July 2021
Accepted: 18 August 2021
Published: 20 August 2021

Abstract: The knowledge of local culture is essential to establish competitive strategies in higher education. The objective of this research was to identify the organizational differences among three universities with different international contexts and satisfaction level. An approach was made regarding Relational Coordination (RC) attributes: accurate, frequent and problem-solving communication, shared knowledge, mutual respect and shared goals, by discriminant analysis method. A random sample of 300 students, 100 belonging to each university, was surveyed on the 23 RC variables in 2017–2018. First, the RC variables were evaluated by general linear model (GLM). The three universities—Arcada University of Applied Science (ARCADA) in Finland, University of Cordoba (UCO) in Spain and Agricultural Polytechnic of Manabi "MFL" (ESPAM) in Ecuador—and the two levels of student satisfaction—Low and High—were used as fixed factors. Second, a discriminant model was built with RC variables. A higher level of RC practices concerning to accurate, frequent and problem-solving communication achieved higher levels of satisfaction, regardless of the universities' socioeconomic context. RC differentiation among three universities showed that shared goals with lecturers and administrative officers and problem-solving communication among classmates were the variables with the highest discriminant power. Two clusters were obtained, where UCO was the most differentiated university. In conclusion, organizational practices made a difference among the three universities. Discriminant analysis can be adapted and extended to different universities to improve quality.

Keywords: relational coordination; student satisfaction; higher education; communication; discriminant analysis

1. Introduction

As a key to value creation in modern societies, improvement in higher education has received considerable attention from policy makers [1,2]. The way Higher Education Institutions have put into practice organizational learning is considered a key element [3]. Organizational learning is interpersonal and relational, and it has often involved learning to coordinate work in new ways [4]. Coordination has been explained by organization design and contingency theorists such a Kundu et al. [5], as an information-processing problem. Some authors such a Faraj and Sproull [6] and Margalina et al. [7] have perceived

coordination as shared understanding of work, and the context in which the activity is carried out has been defined as a relational process. According to Fu et al. [8], coordination is an important management strategy which helps organizations to improve efficiency and effectiveness.

Gittell [9] defined relational coordination (RC) as a mutually reinforcing process of communicating and relating for the purpose of task integration. RC is a mechanism based on human factors. Even though excellence is measured by results, it is generated in intangible, personal and human processes [10]. The RC model could become a helpful tool to measure and encourage effective coordination. RC is based on human relationships among emotional beings [11]. For this reason, it attempts to group together all the connections between them, not merely as tasks, tools or technical needs, but rather in real ways to make it possible for people to work efficiently. RC has been built around work coordination, by considering all aspects of team's relationships. It recognizes, therefore, the need for the relational side of coordination to achieve organizational effectiveness [12,13].

Gittell et al. [14] structured the RC model around two dimensions: communication and relationships. The communication dimensions are: (i) frequent communication helps to establish relationships via roles through the proximity generated because of repeated interaction; (ii) timely communication, communication provided on time; (iii) accurate communication, in the context of relevant information, this plays a critical role in the effectiveness of a group's tasks performance; (iv) problem-solving communication, referred to effective coordination to solve problems. The relationship dimensions are: (i) shared goals: these play a key role in the coordination of highly interdependent tasks; (ii) shared knowledge: communication among those involved in the various tasks that constitute a process is not always effective because of different social backgrounds, training and experience; (iii) mutual respect, that generates an effective coordination, because participant's profiles in the same process value the contribution of others and consider the impact of their own actions in others too.

Understanding the RC factor relationships helped to know how resources can be organized best in order to maximize an institution's performance [10,11]. Existing research showed that RC was positively linked to organizational performance in several sectors. Gittell et al. [11] applied it to different medical units inside hospitals and observed that units with higher levels of RC produced best performance. Havens et al. [15] explained higher levels of job satisfaction, work engagement and decrease of burnout from the RC perspective. Haider et al. [13] applied the RC to the banking industry to explain the relationship between high performance work systems and job satisfaction. Gallego et al. [16,17] and Margalina et al. [7] applied the RC model to explain best results in online systems in higher education. The model was also applied to face-to-face learning. At Quevedo State Technical University (Ecuador), a typology of organizational models was built [18]. In addition, the level of quality in education of Agricultural Polytechnic of Manabi "MFL" and Quevedo State Technical University [19] was estimated. Furthermore, Checa et al. [20] located RC factors oriented towards sustainability in higher education.

Quality is an important performance indicator for education, and it is one of the main issues examined by modern scholars and practitioners from the international education market [21]. The main problem lies in the subjectivity of the concept of quality, which makes it difficult to measure [22]. According to Gallego et al. [16], an indicator to measure the quality was the degree of students' satisfaction. Satisfaction showed a customer orientation, linking what it is expected from one student with the obtained result [23]. Student's perceived satisfaction showed the efficiency of organizations at different areas of activity: Academic services, administrative services, teaching staff, training programs, etc. [24]. However, there is still a lack of empirical research that examines the relationship between RC and students' satisfaction to solve some questions such us: How deep is that relationship? Does an improvement in RC increase student satisfaction? Furthermore, in the current literature, the findings are based on cases that can hardly be extended to other universities. Addi-Raccah and Gavish [25], Lee and Yu [26] and Noël et al. [27] identified

organizational differences through discriminant analysis. A comparison among universities allows identifying the key organizational factors that differentiate them. This can enable the design of strategic measures oriented to improve the quality of performances.

We pose the following research questions: (RQ1) Do universities located in developed countries showed a higher level of RC? (RQ2) Do the most satisfied students have a higher level of RC? (RQ3) Is it possible to build an organizational model that differentiates the three institutions?

Therefore, the objective of this study was to identify the organizational differences among three universities with different socioeconomic contexts and two levels of student satisfaction, from the RC perspective. The cases of Arcada University of Applied Science (ARCADA) in Finland, University of Cordoba (UCO) in Spain and Agricultural Polytechnic of Manabi "MFL" (ESPAM) in Ecuador with different satisfaction levels and socioeconomic contexts were selected. The analysis was developed in two stages. In the first stage, considering the 23 organizational variables proposed, those variables with significant differences among the three universities and between the two levels of satisfaction by general linear model (GLM) were identified. In the second stage, the organizational differences among the three universities were explained by discriminant analysis.

Socioeconomic Contexts

The knowledge of local culture and the socioeconomic situation are essential to establish competitive educational leadership and management strategies [28]. In this research, three universities were selected as representative instruments of three international socioeconomic contexts with different organizational structures. ARCADA represented Finland, which has a high Gross Domestic Product (GDP) (Table 1). This university is in the position 18 out of 35 in the ranking of Finnish Universities [29]. UCO represented Spain, which has a medium GDP (Table 1). UCO is a non-private university and appears in position 53 out of 120 in the ranking of Spanish universities [30]. Ecuador presented the lowest GDP (Table 1) and was represented by ESPAM. This is university is in the position 41 out of 61 universities in Ecuador [31]. In Table 1, the main differences between the three socioeconomic contexts and the positions in the higher education rankings of the three universities were shown.

The three countries were compared using the Program for International Student Assessment (PISA) report. This study carried out by the countries belonging to the Organization for Economic Cooperation and Development (OECD), measures the academic performance of students according to subjects such as mathematics, science and reading. Finland and Spain showed their data in the regular PISA report [32,33], while Ecuador presented their data in the "PISA for development" or PISA-D report [34], an OECD initiative for low-and middle-income countries. Ranking Web of Universities showed the position of each university worldwide [35].

Table 1. Socioeconomic contexts (pre-COVID 19).

Characteristics of the Country						PISA Report [32,33]/PISA-D Report [34]								Ranking [3]
Country	Population	GDP [1]	Public univ.	Non-public univ.	Political System	Mathematics and science average (pts.)	High performance students (%)	Socioeconomic impact (%) [2]	Student performance trend	Staff and resources quality	Qualified teaching by type of centre (%)	Bullying cases (%)	Lifestyle satisfaction (%)	Student growth mindset (%)	
Finland	5,515,525	48,280	13	22	Parliamentary republic	Mathem. 520 Science 552	12	11	Decrease	Remain	Favoured: 94 Disadvantaged: 91	18	78	67	ARCADA = 4589
Spain	46,797,754	29,350	64	56	Parliamentary monarchy	Mathem. 481 Science 483	4	10	Stable	Shortage	Favoured: 94 Disadvantaged: 98	17	74	62	UCO = 679
Ecuador	17,084,359	6090	30	31	Presidential constitutional republic	Mathem. 377 Science 399	1.4	25	Decrease	Shortage	No data	No data	87	No data	ESPAM = 15,330

[1] Gross Domestic Product: $/person/year. [2] Status students over performance. [3] Ranking web universities [35].

Arcada University of Applied Science (ARCADA) is a private university located in Finland, in Northern Europe. It is composed of 2443 students, 165 employees, 4 educational departments, 17 grade programs and 10 Master programs. U-Multirank ranking evaluates five dimensions of higher education centers (teaching and learning, research, knowledge transfer, international orientation and regional engagement). ARCADA overall profile showed higher performance on several indicators, with "A" scores (very good) overall. According to U-Multirank [36], ARCADA was strongest in Regional Engagement. ARCADA was in position 4589 out of 30,585 in the Ranking Web of Universities [35]. Within the three world rankings that classify the top 1000 universities, University of Cordoba (UCO) was ranked at position 800 in The Higher Education World Universities Ranking [37], position 700–800 in the Shanghai Ranking [38] and position 101–150 in QS World Universities Ranking [39]. It presented a medium size dimension with 21,000 students, 1200 lecturers, 700 workers, 47 undergraduate studies and more than 50 postgraduate studies. Finally, it was classified in the position 686 out of 30,585 in the Ranking Web of Universities [35]. Moreover, Agricultural Polytechnic of Manabi, ESPAM, is a public institution located in Ecuador. It was graded with "C" category by the Council of Evaluation, Accreditation and Quality Assurance of Upper Education (CEAACES) [40]. This is a ranking applicable to Ecuadorian universities exclusively. This classification is distributed in a decreasing way from category "A" to "D". In this case, 2811 students and 176 employees compose ESPAM, and it offers 8 grade programs. In addition, it was ranked 15,330 out of 30,585 in the Ranking Web of Universities [35].

2. Materials and Methods

2.1. Data Collection and Survey

A stratified random sample composed of 300 surveyed students, 100 from each university, was collected during the period 2017–2018. The initial data started from a database for each university, made up of 200–1000 data each one. Incomplete surveys and those that showed logical inconsistencies were deleted. Finally, a group of 100 surveys from each university was randomly selected with the random function of the spreadsheet software, making up the complete database with 300 surveys. The survey included 33 items: 4 socioeconomic (age, gender, character public/private, size), 6 of students' perceived satisfaction and 23 related to RC. The survey's reliability was verified by means of Cronbach's alpha, with values greater than 0.7, acceptable to confirm internal consistency: communication dimension (0.703), relationship dimension (0.831) and satisfaction (0.793) (Table 2). The complete survey showed a Cronbach's alpha of 0.87 [20,41].

The 23 items of the RC model focused on the mechanisms involved in organizational practices are shown in Table 2. 11 variables of the communication dimension, 12 of relationship dimension and 6 related to the level of student satisfaction were used. The students answered each question of the survey (Table S1) as many times as profiles were observed at the university. Then, each relational coordination variable was disaggregated into the following profiles: lecturers, administrative officers, classmates, student representatives and me (myself), as a control variable. A Likert scale metric was used, from 1 (infrequent) to 5 (very frequent). In this case, the intervals between the points on the scale corresponded to empirical observations in the metric sense [42]. A visual analog scale was displayed on each survey question presented to the students.

Table 2. Relational coordination and satisfaction variables.

Dimension	α Cronbach	Code	Question/Variable
COMMUNICATION	0.703	ACCURATE COMMUNICATION: Do the people who belong to these areas have the need to offer you information at certain times?	
		1. $ACCU_{Admin}$	Accurate communication with administrative officers
		2. $ACCU_{Lect}$	Accurate communication with lecturers
		3. $ACCU_{Class}$	Accurate communication with classmates
		FREQUENT COMMUNICATION: Do people who belong to the following work areas communicate with you frequently?	
		4. $FREQ_{Admin}$	Frequent communication with administrative officers
		5. $FREQ_{Lect}$	Frequent communication with lecturers
		6. $FREQ_{Class}$	Frequent communication with classmates
		SOLVING PROBLEM COMMUNICATION: When any type of problem appears (study, logistics, documentation . . .), how much did the following profiles help you to solve your problem?	
		7. $SOLPRO_{Myself}$	Problem-solving communication with myself
		8. $SOLPRO_{Lect}$	Problem-solving communication with lecturers
		9. $SOLPRO_{Repres}$	Problem-solving communication with students' representatives
		10. $SOLPRO_{Admin}$	Problem-solving communication with administrative officers
		11. $SOLPRO_{Class}$	Problem-solving communication with classmates
RELATIONSHIP	0.831	SHARED KNOWLEDGE: How well do the following profiles know about your role in the university and the problems that arise?	
		12. $SKNOW_{Lect}$	Shared knowledge with lecturers
		13. $SKNOW_{Repres}$	Shared knowledge with students' representatives
		14. $SKNOW_{Admin}$	Shared knowledge with administrative officers
		15. $SKNOW_{Class}$	Shared knowledge with classmates
		MUTUAL RESPECT: How much do the following profiles respect your role at the university?	
		16. $RESPE_{Lect}$	Mutual respect with lectures
		17. $RESPE_{Repres}$	Mutual respect with students' representatives
		18. $RESPE_{Admin}$	Mutual respect with administrative officers
		19. $RESPE_{Class}$	Mutual respect with classmates
		SHARED GOALS: How well do the following profiles share your goals at the university?	
		20. $SHARGOAL_{Lect}$	Shared goals with lecturers
		21. $SHARGOAL_{Repres}$	Shared goals with students' representatives
		22. $SHARGOAL_{Admin}$	Shared goals with administrative officers
		23. $SHARGOAL_{Class}$	Shared goals with classmates
SATISFACTION	0.793	STUDENT SATISFACTION: Indicate your degree of satisfaction with the following profiles.	
		24. $SATIS_{Lect}$	Satisfaction with lectures
		25. $SATIS_{Represent}$	Satisfaction with students' representatives
		26. $SATIS_{Admin}$	Satisfaction with administrative officers
		27. $SATIS_{Materials}$	Satisfaction with materials
		28. $SATIS_{Communic}$	Satisfaction with communication channels
		29. $SATIS_{Contents}$	Satisfaction with training contents

The proposed indicator of satisfaction was based on the student's satisfaction level [7,16,43]. This indicator was obtained from variables 24–29, related to profiles of conferences, student representatives, administrative officers, materials, communication channels, training con-

tent. The descriptive statistics of trend, dispersion and position of the satisfaction variable were calculated (Figure S1). In each university the median ranges between 18–20 points and for the total sample of 19 points. Therefore, two levels were determined: 19 points was used as border: less than 19, "Low satisfaction" and more than 19, "High satisfaction" [19,20]. Later, satisfaction level was understood as fixed or independent variable.

2.2. Statistical Analysis

The normality of the data distribution was evaluated using the Kolmogorov-Smirnov test (with the Lilliefords correction) and a Levene test was used to evaluate the homogeneity of variance. For those variables that did not show a normal distribution, the Bartlett test was applied to assess if the data had equal variances.

In the first stage, to answer RQ1 and RQ2, the RC variables influenced by the university (socioeconomic context) and the level of satisfaction were identified. 23 variables of RC were compared using the general linear model (GLM). The three universities (ARCADA, ESPAM and UCO) and two satisfaction levels (Low and High) were used as fixed factors. The interactions between both factors were also considered [43]. Three levels: * p-value < 0.05; ** p-value < 0.01 and *** p-value < 0.001 were considered. The test allowed determining which pairs of means differed significantly and to study data whose error did not conform to the normal distribution and non-constant variances. The test allowed determining which pairs of means differed significantly and to study data whose error did not conform to the normal distribution and non-constant variances [43].

Secondly, an organizational model was built using a canonical discriminant analysis to answer to RQ3. This analysis allows studying the concrete relationships that exist among discriminated groups (universities) and their degree of association [44]. The coefficients of the discriminant model show the relative contribution of the variables to the model. The higher the value of the F-remove coefficient, the greater the contribution to group discrimination [45,46]. Therefore, variables with a p-value < 0.05 were accepted and a model with highest percentage of correctly classified cases was selected. The most discriminant variables were calculated applying the F of Snedecor, Wilks' Lambda and the 1-Tolerance. High values of F for each variable indicated that the means of each group were widely separated. Small Lambda values showed that the variable discriminates well among groups. Variables with a high percentage of tolerance (1-Toler) were selected [26]. Statistical analyses were performed using the STATISTICA 12.

3. Results

The three universities showed an average age of students less than 25 years in 86% of the sample (p-value < 0.05). Regarding gender, the distribution was uniform in ARCADA. However, in UCO most of the students were women (p-value < 0.001) and in ESPAM most of the students were men (p-value < 0.05). Regarding the field of knowledge, significant differences were found among the three universities. In ARCADA 100% of the data corresponded to the Social Sciences (p-value < 0.001), in UCO the Health Sciences predominated (90%; p-value < 0.001) and in ESPAM the Engineering Areas obtained the highest percentage (72%; p-value < 0.05). The sociodemographic indicators of the sample are shown in Table 3.

Table 3. Sociodemographic distribution of the sample (%).

	Age			Gender			Field of Knowledge			
	<25	≥25	*p*-Value	Male	Female	*p*-Value	Social sciences	Engineering	Health sciences	*p*-Value
ARCADA	86	14	ns	54	46	ns	100	-	-	***
UCO	88	12	ns	30	70	***	-	10	90	***
ESPAM	83	17	ns	56	44	*	15	72	13	*
TOTAL	86	14	ns	46	54	***	38	27	35	***

* *p*-value < 0.05; ** *p*-value < 0.01; *** *p*-value < 0.001; ns = not significantly different.

The three universities reached relational coordination values close to the average (69.87 ± 0.78; CV = 0.19). Regarding satisfaction, UCO obtained the lowest level (18.25 ± 0.44; CV = 0.24) and ARCADA (19.44 ± 0.53; CV = 0.27), the highest one. The dispersion coefficient was low in the three universities (data not presented).

3.1. Identification of Organizational Differences

GLM results are shown in Table 4. Significant differences were found in most of the variables of RC by university and satisfaction (*p*-value < 0.05). 82.61% of the RC variables showed significant differences by university. The highest RC values were observed at ESPAM and UCO, while ARCADA showed lower values. ARCADA showed significant differences in the variables related to solving problem communication and shared knowledge with the administrative officers. In UCO significant differences were found, highlighting the frequent communication, shared knowledge and mutual respect related to classmates. Lastly, ESPAM showed significant differences in accurate and frequent communication variables, and shared knowledge and goals were the variables that stood out, relating to lectures.

Table 4. Relational coordination by university and satisfaction level (Mean ± SE).

Variable	University (A)			Satisfaction Level (B)		*p*-Value		
	ARCADA	UCO	ESPAM	Low	High	University (A)	Satisfaction level (B)	Interactions (A × B)
1. $ACCU_{Admin}$	$2.75^{a} \pm 0.10$	$2.60^{a} \pm 0.10$	$3.38^{b} \pm 0.10$	2.67 ± 0.09	3.15 ± 0.08	***	***	ns
2. $ACCU_{Lect}$	$3.57^{a} \pm 0.09$	$3.77^{ab} \pm 0.09$	$3.95^{b} \pm 0.09$	3.60 ± 0.07	3.92 ± 0.07	**	**	**
3. $ACCU_{Class}$	$3.67^{a} \pm 0.10$	$3.76^{a} \pm 0.10$	$3.84^{a} \pm 0.10$	3.67 ± 0.08	3.84 ± 0.08	ns	ns	ns
4. $FREQ_{Admin}$	$2.54^{b} \pm 0.09$	$1.99^{a} \pm 0.09$	$2.81^{c} \pm 0.10$	2.16 ± 0.08	2.74 ± 0.07	***	***	**
5. $FREQ_{Lect}$	$3.44^{a} \pm 0.09$	$3.54^{a} \pm 0.09$	$3.83^{b} \pm 0.09$	3.36 ± 0.08	3.84 ± 0.07	**	***	ns
6. $FREQ_{Class}$	$4.08^{ab} \pm 0.09$	$4.30^{b} \pm 0.09$	$4.00^{a} \pm 0.09$	4.08 ± 0.08	4.17 ± 0.07	ns	ns	ns
7. $SOLPRO_{Myself}$	$4.20^{ab} \pm 0.08$	$4.43^{b} \pm 0.08$	$4.18^{a} \pm 0.08$	4.29 ± 0.07	4.25 ± 0.07	ns	ns	ns
8. $SOLPRO_{Lect}$	$3.04^{a} \pm 0.08$	$3.04^{a} \pm 0.08$	$3.11^{a} \pm 0.08$	2.87 ± 0.07	3.26 ± 0.07	ns	***	ns
9. $SOLPRO_{Repres}$	$2.13^{a} \pm 0.11$	$2.51^{b} \pm 0.11$	$2.47^{b} \pm 0.11$	2.18 ± 0.09	2.56 ± 0.09	*	**	ns
10. $SOLPRO_{Admin}$	$2.50^{b} \pm 0.09$	$2.05^{a} \pm 0.09$	$2.46^{b} \pm 0.09$	2.09 ± 0.08	2.59 ± 0.07	***	***	ns
11. $SOLPRO_{Class}$	$3.43^{b} \pm 0.10$	$3.78^{c} \pm 0.10$	$2.95^{a} \pm 0.10$	3.37 ± 0.08	3.40 ± 0.08	***	ns	ns
12. $SKNOW_{Lect}$	$3.04^{ab} \pm 0.11$	$2.85^{a} \pm 0.10$	$3.16^{b} \pm 0.10$	2.81 ± 0.09	3.23 ± 0.08	ns	**	ns
13. $SKNOW_{Repres}$	$2.36^{a} \pm 0.12$	$2.21^{a} \pm 0.11$	$2.51^{a} \pm 0.11$	2.23 ± 0.10	2.48 ± 0.09	ns	ns	**
14. $SKNOW_{Admin}$	$2.40^{b} \pm 0.10$	$1.73^{a} \pm 0.10$	$2.47^{b} \pm 0.10$	1.96 ± 0.09	2.44 ± 0.08	***	***	***
15. $SKNOW_{Class}$	$3.35^{a} \pm 0.11$	$4.02^{b} \pm 0.10$	$3.21^{a} \pm 0.10$	3.47 ± 0.09	3.59 ± 0.08	***	ns	ns
16. $RESPE_{Lect}$	3.30 ± 0.11	3.62 ± 0.10	3.58 ± 0.10	3.27 ± 0.09	3.73 ± 0.08	ns	***	**
17. $RESPE_{Repres}$	$2.52^{a} \pm 0.12$	$2.96^{b} \pm 0.12$	$3.06^{b} \pm 0.12$	2.60 ± 0.10	3.09 ± 0.09	**	***	**
18. $RESPE_{Admin}$	$2.65^{b} \pm 0.12$	$2.33^{a} \pm 0.11$	$3.00^{c} \pm 0.11$	2.33 ± 0.10	2.99 ± 0.09	***	***	**
19. $RESPE_{Class}$	$3.42^{a} \pm 0.10$	$4.15^{b} \pm 0.09$	$3.53^{a} \pm 0.09$	3.62 ± 0.08	3.78 ± 0.07	***	ns	ns
20. $SHARGOAL_{Lect}$	$2.84^{a} \pm 0.10$	$3.09^{a} \pm 0.10$	$3.43^{b} \pm 0.10$	2.83 ± 0.08	3.41 ± 0.08	***	***	ns
21. $SHARGOAL_{Repres}$	$2.46^{a} \pm 0.12$	$2.65^{ab} \pm 0.12$	$2.85^{b} \pm 0.12$	2.33 ± 0.10	2.97 ± 0.09	ns	***	*
22. $SHARGOAL_{Admin}$	$2.55^{b} \pm 0.10$	$1.77^{a} \pm 0.10$	$2.85^{c} \pm 0.10$	2.09 ± 0.08	2.68 ± 0.08	***	***	*
23. $SHARGOAL_{Class}$	$3.46^{a} \pm 0.10$	$4.03^{b} \pm 0.10$	$3.46^{a} \pm 0.10$	3.59 ± 0.08	3.71 ± 0.08	***	ns	**

* *p*-value < 0.05; ** *p*-value < 0.01; *** *p*-value < 0.001; ns = not significantly different. ^{a}, ^{b} Within row, averages with different superscript differ significantly.

Significant differences by level of student satisfaction were found in 65.21% of the organizational variables. 26.09% of the variables showed differences according to both criteria (Table 4). The non-significant variables were those related to the classmates and representatives of the students in the two dimensions of RC. The interactions between university and satisfaction were found in six RC variables. Most of the variables were

related to the profile of administrative officers. The interactions found between both factors in these six significant variables are shown in Figure 1.

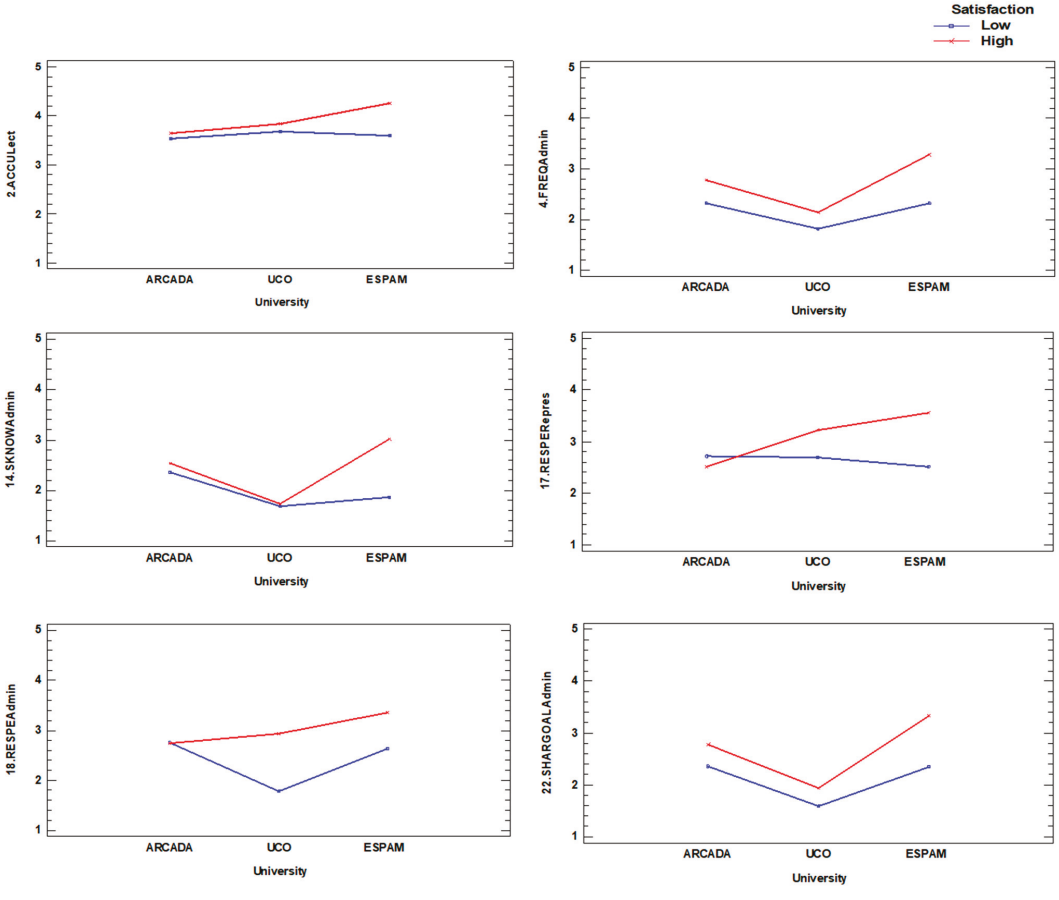

Figure 1. Interactions between university and level of satisfaction.

3.2. Discriminant Model Building

Discriminant analysis model among universities was carried out. As predictors, 23 RC variables were used. The eight significant variables, which showed a p-value < 0.05, were selected for the construction of the discriminant model: Three related to the communication and five to the relationship dimension (Table 5). Accurate communication with administrative officers and lecturers, shared knowledge and mutual respect with classmates; and shared objectives with the representatives of the students belonged to the discriminant model. Additionally, shared goals with lectures and administrative officers, and the communication for solving problems among classmates were the variables with the three highest discriminant powers, showing a higher F-remove coefficient.

Table 5. Discriminant function for the organizational variables of three universities (ARCADA, UCO and ESPAM).

Variable	Wilks'	Partial	F-Remove	*p*-Value	Toler	1-Toler
2. $ACCU_{Lect}$	0.491	0.974	3.088	*	0.603	0.397
3. $ACCU_{Class}$	0.494	0.967	4.029	*	0.681	0.319
11. $SOLPRO_{Class}$	0.516	0.927	9.279	***	0.737	0.263
15. $SKNOW_{Class}$	0.493	0.971	3.568	*	0.549	0.451
19. $RESPE_{Class}$	0.494	0.967	3.960	*	0.643	0.357
20. $SHARGOAL_{Lect}$	0.518	0.924	9.726	***	0.558	0.442
21. $SHARGOAL_{Repres}$	0.491	0.973	3.290	*	0.457	0.543
22. $SHARGOAL_{Admin}$	0.550	0.869	17.672	***	0.480	0.520

* *p*-value < 0.05; *** *p*-value < 0.001.

The classification matrix offered a correct ascription percentage of 69.32, obtaining assignment errors only in ESPAM (data not presented). The organizational differences of the three analyzed universities are shown in Figures 2 and 3. In the first one, in which the Mahalanobis distances obtained from the relational coordination indicators were graphically represented, a first cluster grouped ARCADA and ESPAM University and second cluster made up UCO. The students from UCO showed greater separation, and, therefore, greater relational coordination differentiation, due to its lower RC rating. The existence of different relational models for each university were observed in Figure 3, which showed a spatial distribution of each university with overlap of some individuals of ARCADA and ESPAM, but strong distance from UCO showed a clear differentiation.

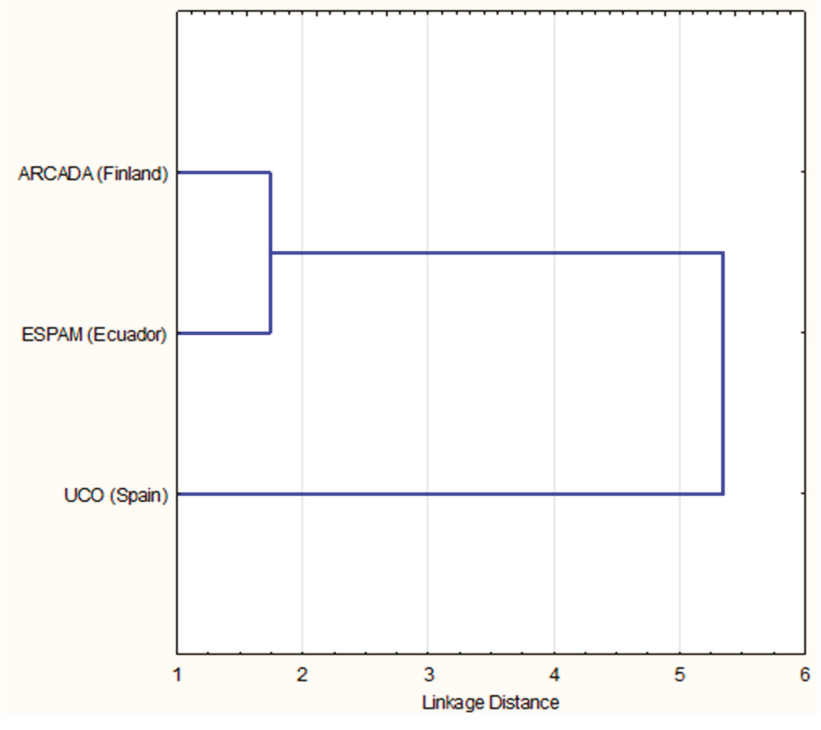

Figure 2. Cluster from Mahalanobis distances for three universities.

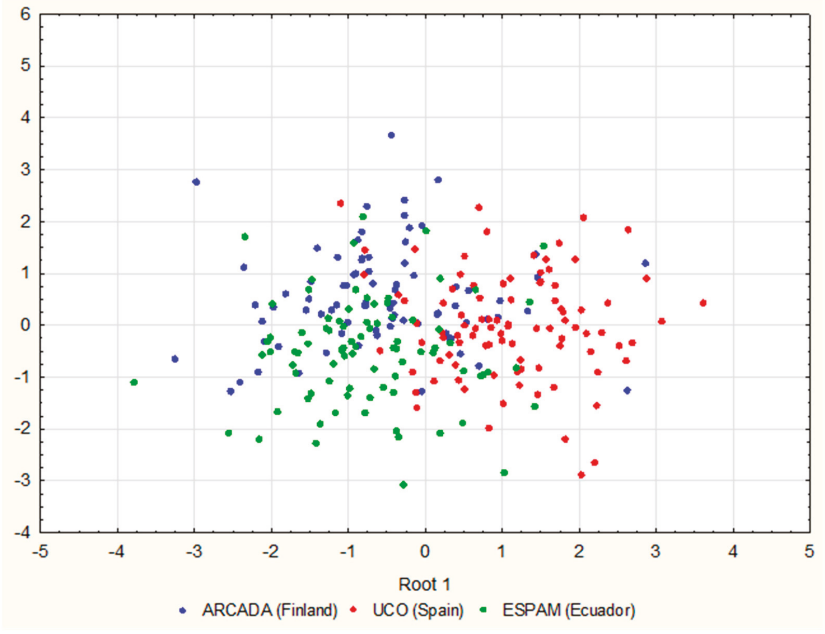

Figure 3. Plot of the individual observation discriminant scores obtained with the canonical discriminant function for three universities.

4. Discussion

The relational coordination framework provides an excellent basis for investigating the types of organizational models at universities [7,16,17,47]. According to [47], higher levels of relational coordination improve results. RC model can be useful to achieve excellent results in higher education where high levels of task interdependence, uncertainty, time restrictions and tacit knowledge are required [7]. In the case of higher education, it is important to identify best organizational practices to apply at universities, as well as the differences among universities, which contribute to the global knowledge of the importance of RC on the results of the organization [16,20]. The methodology developed in this research has allowed, in a first step, identifying the relational coordination variables that promote differences among universities and satisfaction levels. In a second stage, according to Addi-Raccah and Gavish [25], Lee and Yu [26] and Noël et al. [27], a canonical discriminant function for the ARCADA, UCO and ESPAM universities, in three countries and very different socioeconomics contexts has been built.

RQ1 was not validated in this study. According to De-Pablos-Heredero et al. [19], an improvement in organizational practices leads to an improvement in results regardless of the socioeconomic context.

RQ2 was validated, finding a positive relationship between RC and student satisfaction level. In the three universities there is a positive effect between RC and satisfaction. This link is more prominent in the case of ESPAM (Figure 1). In ESPAM, with high levels of RC, the highest values of satisfaction have been obtained. In ESPAM, which is a small size public university in a developing country with low economic growth, the level of satisfaction is very sensitive to the modifications in RC in the administrative officers profile [7]. According to the Pisa-D report [34] Ecuador requires an improvement in digital literacy, so there is a greater dependence on administrative officers [7]. Therefore, the different social contexts could explain part of the differences in organizational patterns [28].

Accurate and solving problem communication, mutual respect and shared knowledge and goals are strategic factors to improve de RC. The results obtained show that the personalized service to the student is positively valued by considering individual circumstances. Gallego et al. [16,17] and Margalina et al. [7] proofed how in universities with high quality levels, the institutional coordination with students was stronger. Havens et al. [15] and Haider et al. [13] paid attention to the similarities between teamwork quality and RC. Lacayo-Mendoza and De-Pablos-Heredero [48] indicated that the majority of students highly value the facilities provided by educational staff. Finally, results show that other outstanding attributes are shared goals with students' representatives and with administrative officers. Gallego et al. [16,17] and Margalina et al. [7] concluded how in universities exhibiting high quality levels, the institutional coordination with students is strong.

The construction of a discriminant model verified RQ3. Knowing the variables with the greatest discriminant power, it is possible to propose concrete, simple and economic measures to improve educational quality. The results of this research allow establishing the organizational differentiation among three Universities though discriminant analysis. Shared goals, with lectures and administrative officers, and the communication for solving problems among classmates were the variables with the highest discriminant power. UCO was the most differentiated university according to RC (Figures 2 and 3). This differentiation explains the fact that it is the highest ranked university in the world ranking of universities (Table 1).

Three different universities could be discriminated by the organizational model generated. Shared goals are a key piece for university excellence [17], therefore measures that allow sharing the objectives of the students with lectures and administrative officers are crucial. In order to enhance this, improvements are proposed in digital literacy for communication with administrative officers [7] and changes in the teaching guides, where the lectures establish specific objectives for the students in each subject, are welcomed. Solving problem communication shows that the students use the educational ecosystem in moments of lack of information [16,17]. This way, the creation of direct communication mechanisms among students and other profiles is proposed to solve the problems of university life.

Apart from this, it would be of great interest to develop prediction models for each set of organizational variables over satisfaction. This issue could be developed in future research lines by applying structural equation models. This approach could be extended to different universities and contexts.

5. Conclusions

This research contributes to a novel approach since it allows identifying the organizational differences among three universities with different socioeconomic contexts.

In each university, as the relational coordination dimensions are improved, the level of satisfaction increases. However, an association among universities located in countries with a higher level of economic resources and a higher level of relational coordination, has not been verified. Those universities that implement a program of best practices in relational coordination will achieve higher levels of quality in terms of student satisfaction, regardless the socioeconomic context.

The canonical discriminant model built according to the relational coordination dimensions showed that three organizational variables were enough to explain differences among universities. These variables were shared goals, with lectures and administrative officers, and the communication oriented to solve problems among classmates. Therefore, the discriminant analysis is useful for designing the improvement of the relational practices in each university.

The proposed model can easily be adapted and applied to different contexts and, therefore, they can be of great interest for the improvement of quality at universities. The results were validated but are conditioned in each university by its standard of satisfaction values.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/educsci11080445/s1, Figure S1: Statistical parameters of the student satisfaction value, Table S1: Relational coordination survey.

Author Contributions: Conceptualization and methodology, all authors. Formal analysis, software, data curation, data processing, A.L.C.; statistical analysis, C.C.-M. and S.H.; validation and investigation, C.D.-P.-H. and A.G.; supervision, A.G., C.D.-P.-H. and C.C.-M.; project administration, C.D.-P.-H. and A.G.; data acquisition, A.L.C. and A.G. All authors have been involved in developing, writing, commenting, editing and reviewing the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: This is not applicable as the data are not in any data repository of public access, however if editorial committee needs access, we will happily provide them, please use this email: pa1gamaa@uco.es.

Acknowledgments: We want to thank the RED-RC, the open research and innovation network for the improvement in higher education among Quevedo State Technical University (Ecuador), University of Cordoba (Spain) and Agricultural Polytechnic of Manabi "MFL" (Ecuador). The authors are thankful for the funding provided to the research project by the competitive scientific and technological research funds of Agricultural Polytechnic of Manabi "MFL" (Ecuador).

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

References

1. Bartimote-Aufflick, K.; Bridgeman, A.; Walker, R.; Sharma, M.; Smith, L. The Study, Evaluation, and Improvement of University Student Self-Efficacy. *Stud. High. Educ.* **2016**, *41*, 1918–1942. [CrossRef]
2. The Future of Jobs Report 2020. Available online: https://www.weforum.org/reports/the-future-of-jobs-report-2020/in-full/ (accessed on 15 July 2021).
3. Gast, I.; Schildkamp, K.; Van-der-Veen, J.T. Team-Based Professional Development Interventions in Higher Education: A Systematic Review. *Rev. Educ. Res.* **2017**, *87*, 736–767. [CrossRef]
4. Santos, J.M.; Horta, H. The Research Agenda Setting of Higher Education Researchers. *High. Educ.* **2018**, *76*, 649–668. [CrossRef]
5. Kundu, A.; Sarkar, D.N.; Bhattacharya, A. The Effect of Uncertainty on the Formulation of Strategies: A Study of Selected Indian Organizations. *SN Bus. Econ.* **2021**, *1*, 7. [CrossRef]
6. Faraj, S.; Sproull, L. Coordinating Expertise in Software Development Teams. *Manag. Sci.* **2000**, *46*, 1554–1568. [CrossRef]
7. Margalina, V.M.; De-Pablos-Heredero, C.; Montes-Botella, J.L. Achieving Quality in E-Learning through Relational Coordination. *Stud. High. Educ.* **2017**, *42*, 1655–1670. [CrossRef]
8. Fu, N.; Bosak, J.; Flood, P.C.; Ma, Q. Chinese and Irish Professional Service Firms Compared: Linking HPWS, Organizational Coordination, and Firm Performance. *J. Bus. Res.* **2019**, *95*, 266–276. [CrossRef]
9. Gittell, J.H. Coordinating Mechanisms in Care Provider Groups: Relational Coordination as a Mediator and Input Uncertainty as a Moderator of Performance Effects. *Manag. Sci.* **2002**, *48*, 1408–1426. [CrossRef]
10. Asadullah, M.A.; Haider, S.; Heredero, C.D.P.; Musaddiq, M. Effect of Ingratiation on Supervisor Satisfaction through Helping Behavior: A Moderated Mediation Model. *Intang. Cap.* **2016**, *12*, 1157–1191. [CrossRef]
11. Gittell, J.H.; Logan, C.; Cronenwett, J.; Foster, T.C.; Freeman, R.; Godfrey, M.; Vidal, D.C. Impact of Relational Coordination on Staff and Patient Outcomes in Outpatient Surgical Clinics. *Health Care Manag. Rev.* **2020**, *45*, 12–20. [CrossRef] [PubMed]
12. Haider, S.; Fernandez-Ortiz, A.; De-Pablos-Heredero, C. Organizational Citizenship Behavior and Implementation of Evidence-Based Practice: Moderating Role of Senior Management's Support. *Health Syst.* **2017**, *6*, 226–241. [CrossRef]
13. Haider, S.; De-Pablos-Heredero, C.; De-Pablos-Heredero, M. A Three-Wave Longitudinal Study of Moderated Mediation Between High-Performance Work Systems and Employee Job Satisfaction: The Role of Relational Coordination and Peer Justice Climate. *Front. Psychol.* **2020**, *11*, 792. [CrossRef] [PubMed]
14. Gittell, J.H.; Seidner, R.; Wimbush, J. A Relational Model of How High-Performance Work Systems Work. *Organ. Sci.* **2010**, *21*, 490–506. [CrossRef]
15. Havens, D.S.; Gittell, J.H.; Vasey, J. Impact of Relational Coordination on Nurse Job Satisfaction, Work Engagement and Burnout: Achieving the Quadruple Aim. *J. Nurs. Adm.* **2018**, *48*, 132–140. [CrossRef] [PubMed]
16. Sánchez, M.d.G.; De-Pablos-Heredero, C.; Merodio, J.A.M. Coordinación relacional en la educación en línea. *Interciencia Rev. De Cienc. Y Tecnol. De Am.* **2015**, *40*, 869–874.
17. Gallego, M.d.C.; De-Pablos-Heredero, C.; Medina-Merodio, J.-A.; Robina-Ramírez, R.; Fernandez-Sanz, L. Relationships among Relational Coordination Dimensions: Impact on the Quality of Education Online with a Structural Equations Model. *Technol. Forecast. Soc. Chang.* **2021**, *166*, 120608. [CrossRef]

18. De-Pablos-Heredero, C.; Díaz-Ocampo, E.; Torres, Y.; Checa, C.; Barba, C.; García, A. Typology of relational coordination model in higher education: The case of universidad tecnica estatal de quevedo-ecuador. In *EDULEARN19 Proceedings*; IATED: Valencia, Spain, 2019; pp. 9616–9620.
19. De-Pablos-Heredero, C.; Díaz-Ocampo, E.; Torres, Y.; Checa, C.; Barba, C.; García, A. Organizational quality level of universidad tecnica estatal de quevedo-ecuador. In *INTED2020 Proceedings*; IATED: Valencia, Spain, 2020; pp. 4921–4925.
20. Checa, C.; De-Pablos-Heredero, C.; Torres, Y.G.; Montes-Botella, J.L.; Barba, C.; García, A. Focused Coordination Models towards Sustainability in Higher Education. Case of Quevedo State Technical University (Ecuador). *Sustainability* **2020**, *12*, 5760. [CrossRef]
21. Caffarel-Serra, C.; Ortega-Mohedano, F.; Gaitán-Moya, J.-A. Investigación en Comunicación en la universidad española en el período 2007–2014. *Prof. De La Inf.* **2017**, *26*, 218–227. [CrossRef]
22. Zee, M.; Koomen, H.M.Y. Teacher Self-Efficacy and Its Effects on Classroom Processes, Student Academic Adjustment, and Teacher Well-Being: A Synthesis of 40 Years of Research. *Rev. Educ. Res.* **2016**, *86*, 981–1015. [CrossRef]
23. Ortega-Mohedano, J.; Rodríguez-Conde, M.J. Which Is or Should Be the Ideal Level of Participation of Students of Higher Education in Their Own Education?A Theoretical Model of Relations between the University and Their Students and the Quality of Higher Education. In Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality, Salamanca, Spain, 24 October 2018; ACM: New York, NY, USA; pp. 190–198.
24. Kilgo, C.A.; Culver, K.C.; Young, R.L.; Paulsen, M.B. The Relationship Between Students' Perceptions of "Good Practices for Undergraduate Education" and the Paradigmatic Development of Disciplines in Course-Taking Behavior. *Res. High. Educ.* **2017**, *58*, 430–448. [CrossRef]
25. Addi-Raccah, A.; Gavish, Y. The LEA's Role in a Decentralized School System: The School Principals' View. *Educ. Manag. Adm. Leadersh.* **2010**, *38*, 184–201. [CrossRef]
26. Lee, S.; Yu, J. Discriminant Model of BIM Acceptance Readiness in a Construction Organization. *KSCE J. Civ. Eng.* **2017**, *21*, 555–564. [CrossRef]
27. Noël, N.M.; Trocchia, P.; Luckett, M. A Predictive Psychometric Model to Identify Personality and Gender Differences of College Majors. *Int. J. Manag. Educ.* **2016**, *14*, 240–247. [CrossRef]
28. Bush, T. Preparation and Induction for School Principals: Global Perspectives. *Manag. Educ.* **2018**, *32*, 66–71. [CrossRef]
29. Finlandia | Ranking Web de Universidades: Webometrics Clasifica 30000 Instituciones. Available online: https://webometrics.info/es/Europe_es/Finlandia (accessed on 8 August 2021).
30. España | Ranking Web de Universidades: Webometrics Clasifica 30000 Instituciones. Available online: https://webometrics.info/es/Europe_es/Espa%C3%B1a (accessed on 8 August 2021).
31. Ecuador | Ranking Web de Universidades: Webometrics Clasifica 30000 Instituciones. Available online: https://webometrics.info/es/Latin_America_es/Ecuador (accessed on 8 August 2021).
32. Organisation for Economic Co-operation and Development. *The Programme for International Student Assessment (PISA). Results from PISA 2018, Finland*; Organisation for Economic Co-operation and Development: Paris, France, 2019; p. 10.
33. Organisation for Economic Co-operation and Development (OECD). *The Programme for International Student Assessment (PISA). Results from PISA 2018, Spain*; Organisation for Economic Co-operation and Development (OECD): Paris, France, 2019; p. 9.
34. Organisation for Economic Co-operation and Development (OECD). *Educación En Ecuador. Resultados Del Informe PISA Para El Desarrollo 2018*; Organisation for Economic Co-operation and Development (OECD): Paris, France, 2019; p. 152.
35. World | Ranking Web of Universities: Webometrics Ranks 30000 Institutions. Available online: https://www.webometrics.info/en/WORLD (accessed on 15 July 2021).
36. Ranking & Subjects. Available online: https://www.umultirank.org/study-at/arcada-university-of-applied-science-rankings/ (accessed on 15 July 2021).
37. University of Córdoba. Available online: https://www.timeshighereducation.com/world-university-rankings/university-cordoba (accessed on 15 July 2021).
38. ShanghaiRanking-Univiersities. Available online: http://www.shanghairanking.com/institution/university-of-cordoba (accessed on 15 July 2021).
39. Universidad de Córdoba. Available online: https://www.topuniversities.com/universities/universidad-de-cordoba/undergrad (accessed on 15 July 2021).
40. Ecuador Categorías de Universidades LISTADO 2021. Available online: https://informacionecuador.com/listado-categorias-de-universidades-ceaaces/ (accessed on 15 July 2021).
41. Méndez, D.; Macía, F. Análisis factorial confirmatorio de la escala de actitudes hacia la estadística. *Cuad. De Neuropsicol.* **2007**, *1*, 337–345.
42. Reips, U.-D.; Funke, F. Interval-Level Measurement with Visual Analogue Scales in Internet-Based Research: VAS Generator. *Behav. Res. Methods* **2008**, *40*, 699–704. [CrossRef]
43. Gutiérrez-Cordero, M.L.; Segovia-Vargas, M.J.; Escamilla, M.R. Análisis del Riesgo de Caída de Cartera en Seguros: Metodologías de "Inteligencia Artificial" vs "Modelos Lineales Generalizados". *Econ. Inf.* **2017**, *407*, 56–86. [CrossRef]
44. Rivas, J.; Manuel Perea, J.; De-Pablos-Heredero, C.; Morantes, M.; Angon, E.; Barba, C.; García, A. Role of Technological Innovation in Livestock Breeding Programmes: A Case of Cereal-Sheep System. *Ital. J. Anim. Sci.* **2019**, *18*, 1049–1057. [CrossRef]
45. Legendre, P.; Legendre, L. Chapter 11—Canonical analysis. In *Developments in Environmental Modelling*; Legendre, P., Legendre, L., Eds.; Numerical Ecology; Elsevier: Amsterdam, The Netherlands, 2012; Volume 24, pp. 625–710.

46. Rencher, A.C. Interpretation of Canonical Discriminant Functions, Canonical Variates and Principal Components. *Am. Stat.* **1992**, *46*, 217–225. [CrossRef]
47. Lacayo-Mendoza, A.; De-Pablos-Heredero, C. Managing Relationships and Communications in Higher Education Efficiently through Digital Social Networks: The Importance of the Relational Coordination Model. *DYNA Rev. De La Fac. De Minas. Univ. Nac. De Colombia. Sede Medellín* **2016**, *83*, 138–146. [CrossRef]
48. De-Pablos-Heredero, C.; Haider, S.; Martinez, A.G. Relational Coordination as an Indicator of Teamwork Quality: Potential Application to the Success of e-Learning at Universities. *Int. J. Emerg. Technol. Learn.* **2015**, *10*, 4–8. [CrossRef]

education sciences

MDPI

Article

Moodle Quizzes as a Continuous Assessment in Higher Education: An Exploratory Approach in Physical Chemistry

Isabel López-Tocón

Departamento de Química Física, Facultad de Ciencias, Campus de Excelencia Andalucía Tech, Universidad de Málaga, E-29071 Málaga, Spain; tocon@uma.es

Abstract: The use of Moodle quizzes as a continuous assessment and an integral part of the educational methodology in higher education has been analyzed in a case study of physical chemistry subject. Two types of quiz designed with different item types and different settings, called basic quiz (BQ) and thematic block quiz (TBQ), were elaborated making use of a question bank with more than 450 items. BQ has true/false items, while TBQ has randomly mixed items (multiple choice, numerical and matching). The effect of the type of quiz on the student scores is analyzed according to statistical and psychometric data such as the degree of participation, the facility index and the discrimination index of each item, and the average score, calculated according to the classical test theory. This allows us to discern which type of quiz has an enough quality to use it as an assessment tool. Moreover, the effect of this educational activity, developed during the last six academic years from 2014 to 2020, just before of the pandemic situation, is evaluated considering the scores of the students in the Ordinary Calls of exams and comparing them with previous courses taught with a traditional education based on master classes. The statistic results indicate that TBQs are more discriminative than BQs and could be used as an assessment tool, while BQs could be only useful as formative activity. Moodle quizzes turn out to be a reliable strategy for learning of contents in scientific matter, with a high participation in the knowledge tests, with good marks in the average score and a greater number of pass degrees in the Ordinary Calls.

Keywords: moodle quizzes; continuous assessment; statistical analysis; physical chemistry

Citation: López-Tocón, I. Moodle Quizzes as a Continuous Assessment in Higher Education: An Exploratory Approach in Physical Chemistry. *Educ. Sci.* **2021**, *11*, 500. https://doi.org/10.3390/educsci11090500

Academic Editors: Sandra Raquel Gonçalves Fernandes, Marta Abelha and Ana Teresa Ferreira-Oliveira

Received: 22 June 2021
Accepted: 30 August 2021
Published: 3 September 2021

1. Introduction

The adaptation of the subjects of any graduate degree to the new European Space for Higher Education, where new capacities and abilities are evaluated in students, implies a significant change in the traditional teaching methodology that has been developed, mainly as master classes. This change in the pedagogical model of teaching–learning is in turn conditioned and reinforced by the new model of current digital society and by the new information and communication techniques available in any space-time framework [1,2]. The availability of information on everyday electronic devices such as smartphones or tablets, in addition to the connectivity of user groups to the internet, allows establishing other more active work dynamics, giving to the student a greater participation in the teaching–learning process of a subject [3].

The studies carried out to evaluate the global impact of the use of technology on student performance are not conclusive, yielding different results [4,5] given that these teaching researches may depend on other factors not identified in the analysis itself, such as the educational method or strategy [6] that is carried out through the electronic medium, or the student's commitment to the learning methodology [7]. However, these multimedia and interactive technologies can be of great help in offering quality comprehensive education [8] based on current computer tools that facilitate cognitive learning processes and reinforce the capacities of abstract reasoning and study of a specific subject, in addition to complete the traditional forms of learning [9].

Different teaching methodologies integrate the technological devices in the educational environment such as the blended-learning [10–12] and gamification [13–15]. Blended-learning, b-learning, combines face-to-face lessons in the classroom, required for any subject in university study plans, and virtual training teaching activities through learning platforms, while the gamification techniques try to create similar experiences to those experienced when playing games in order to motivate and engage users. Both methodologies profit from the presence of the teacher as a transmitter of knowledge and guide of educational activities and from the communication technology that facilitates independent and collaborative learning. In particular, b-learning has been applied in subjects from different areas of knowledge such as education sciences [12], natural sciences [16], economics [9], engineering [7], etc. as a proposal for European convergence, given that it allows the student's noncontact work hours to be completed with virtual activities, as established in the new university teaching guides for the convergence of the European Space for Higher Education.

Moodle platform [17] is a virtual learning environment that offers very attractive functionalities from the pedagogical point of view by promoting the philosophy of constructivist social education [18,19], and where the subjects can be accommodated with easy handling, at the editing and user level by teachers and students, respectively. In this virtual environment, teaching resources of different characteristics can be included, such as links to web pages, chats, forums, messages, and other specific documents like notes, tutorials and question relationships elaborated by the teacher. Moreover, it offers the possibility of carrying out online activities through quizzes, which could allow the continuous assessment of students' learning. A great variety of quizzes can be designed with different item types and settings, but not all quizzes can differentiate the skills and competences of student, and thus, they could not be used as assessment tools. The quality of these quizzes can be analyzed by statistical and psychometric data reported by Moodle platform [20,21].

Concerning evaluation methods by using online quizzes, there are studies in diverse disciplines such as engineering, biology, medicine and the social sciences [22–24]. Although, there are some objections to the implementation of such systems related to the confidentiality of the identity of the student, the use of the information and its possible impact on the educational process [25,26]; these offer some advantages such as the efficient management of results in a huge students' group, the speed by which the evaluation can be performed, and the save of paper [27]. However, the design of quizzes must be adequately elaborated in order to be used as an assessment tool. Two important points must be considered in the design, such as the writing of different questions using different item type and the own quiz settings. The statistical and psychometric data derived from a particular quiz can a great help us know the quality of the quiz. There are some studies regarding the analysis of information generated from test-type quiz evaluations in other scientific subjects [20,21,28–30], yielding how such results could be useful for professors and students. No statistical and psychometric studies on physical chemistry quizzes have been found in the bibliography.

In this work, two types of Moodle quizzes are designed in physical chemistry subject. The main objective is to establish which type of quiz can be used as an assessment tool on the basis of statistical and psychometric data. Here, it highlights how Moodle statistics can be used to measure the effectiveness and reliability of a quiz. In addition, the effect of these online activities on the final scores of the students are compared with those obtained in a traditional education.

2. Materials and Methods

The research is designed in three stages: first, the student population was surveyed by a brief poll to inquire about their entry into the university; second, the students answered the quizzes during the teaching semester; and finally, the statistical and psychometric parameters of the quizzes was analyzed on the basis of the classical test theory [31–33]

(See Supplementary Materials). A brief survey is carried out at the end of the teaching period to know the opinion of the students about this experience. The scores obtained in the two Ordinary Calls of exams are compared with those obtained in previous courses where the teaching methodology corresponds to a traditional education based exclusively on master classes.

This research is performed in the general physical chemistry subject during six years, from the 2014–2015 to the 2019–2020 academic years, just before of the pandemic situation. This matter is included in the Basic Module of the Degree in Chemistry at the University of Málaga. It consists of six theoretical credits, and it is taught during the first semester of the first year of the degree.

This subject was chosen because it is a difficult matter for novel students in the Degree of Chemistry. It includes themes like thermodynamic, electrochemistry and kinetics that are the starting point of other physical chemistry subjects in higher courses, in which a significant dropout of students has been detected. Thus, it seems convenient to apply a new educational methodology, or new activities using technological devices, in the first course in order to consolidate and strengthen the basic concepts of this matter.

2.1. Sample

The average number of students in general physical chemistry was about 80 students during the last academic years, with a parity proportion of men and women in the last four years. All students can freely participate in the quizzes as a unique experimental group. No specific sampling method and no control group is established with the aim that all students were evaluated in a homogenous way so that there are no discrepancies in the final evaluation.

It was not possible to perform a similar study in other courses or scientific areas, even in other degrees, because there were no other teachers implied in the project using a similar educational strategy with Moodle quizzes. Although this sample is not representative of the higher education context, the similar results obtained in this experience along different years with different students population point out that it would not be expected to see significant changes in another similar scientific scene, giving probably a similar trend.

At the beginning of the course, a brief survey is carried out to explore the admission at the university, such as academic background on chemistry knowledges and the enrollment in the degree. Considering an average of the last six academic years, practically all the students, 86–88%, are 18 years old, and the rest, 11–12%, are in the range of 21 to 25 years old, which could probably be due to repeaters in the secondary or bachelor cycle, or students who come from other degrees. Most of the students, 86–95%, have studied a chemistry subject during high school, but it should be noted that about 4–5% of students have not studied any chemistry subject in any official degree before to their admission to the university, although they indicate that they have basic knowledge of chemistry. Only a small proportion, 1–2%, have no knowledge of chemistry. Moreover, a high proportion, around 70–85%, has enrolled in the Chemistry degree because it is their vocation, being the first option in university pre-registration. Only 12–25% of students recognize that it is not their vocation and it has not been the first option in the university pre-registration. In addition, this degree was not the first choice of about 1–2% of students, but it was the only option for their admission to the university.

2.2. Development of the Experience: Didactic Strategy

Within the Moodle platform, a question bank has been created and divided into five thematic blocks that involve all topics of the teaching program (Table 1). Each block has more than 50 questions or items, even over 100 items in the cases of the Matter and Thermodynamics blocks. The question bank has over 450 items belonging to four types of Moodle questions: true/false, multiple choice (with multi-responses and single response), matching and numerical. All these items were elaborated according to the scientific competencies required for passing this subject.

Table 1. Teaching program of the general physical chemistry subject developed in eleven lessons and distributed in five thematic blocks. Available at https://oas.sci.uma.es:8443/ht/2020/ProgramasAsignaturas_Titulacion_5004_AsigUMA_51635.pdf (accessed on 22 July 2021).

Theme	Brief Description of the Topic	Thematic Block	N° Moodle Items
1 2 3	The matter: Laws of chemical combination Gaseous state Liquid and solid states	1. Matter (102 items)	27 True/False 53 Multiple choice 9 Matching 10 Numerical
4 5	Solutions Colligative properties of solutions	2. Solutions (98 items)	34 True/False 49 Multiple choice 6 Matching 9 Numerical
6 7 8	Zeroth Law and First Law of Thermodynamics Second and Third Law of Thermodynamics Chemical equilibrium	3. Thermodynamics (122 items)	31 True/False 62 Multiple choice 14 Matching 15 Numerical
9 10	Electrolytes and electrolytic solutions Electrochemistry: electrolytic and galvanic cells	4. Electrochemistry (84 items)	20 True/False 50 Multiple choice 7 Matching 7 Numerical
11	Kinetics chemistry	5. Kinetics (53 items)	13 True/False 29 Multiple choice 2 Matching 9 Numerical

The set of items is classified, in turn, into two categories, one with the questions that collect the basic knowledge of the subject, while the other contains more elaborate questions, in order to check the skills and abilities of the students in practical reasoning about physical chemistry. In this way, two types of quiz are developed. First, a "basic" quiz (BQ) is proposed for each of the eleven topics. It consists of ten true/false type items, with a time limit of one hour. The BQ contains the same questions for all students and is active for a period of one week after finishing the topic in class. Second, another type of "thematic block" quiz (TBQ) is proposed corresponding to each of the five thematic blocks, which are made up of several topics in the teaching program, except those dedicated to chemistry kinetics (see Table 1). It has ten items of different type (multiple choice, numerical, matching) chosen at random from a category of question bank, so it is practically an individual and different test for each student. The multiple choice items have a particular characteristic, the correct/incorrect answers score positively/negatively, with a proportional value to the number of item options. These quizzes are held in a scheduled day.

In both types of quizzes, each item has the same statistical weight of 10% in the final mark. All quizzes are performed outside the classroom and have a delayed feedback; that is, the correct answers can be only checked once the test is over for all students. All these activities are carried out continuously throughout the semester according to the physical chemistry program.

All students were informed about the characteristic of Moodle quizzes and how the platform works before doing the activities. In this way, any bias factor due to students' attitudes towards technology along the time would be diminished.

3. Results and Discussion

3.1. Participation in the Quizzes

There is a high participation during the last six academic courses (Figure 1), higher than 50% in any BQ or TBQ quizzes, with the exception of the last BQ carried out in the 2017–2018 academic year with a participation of 40%.

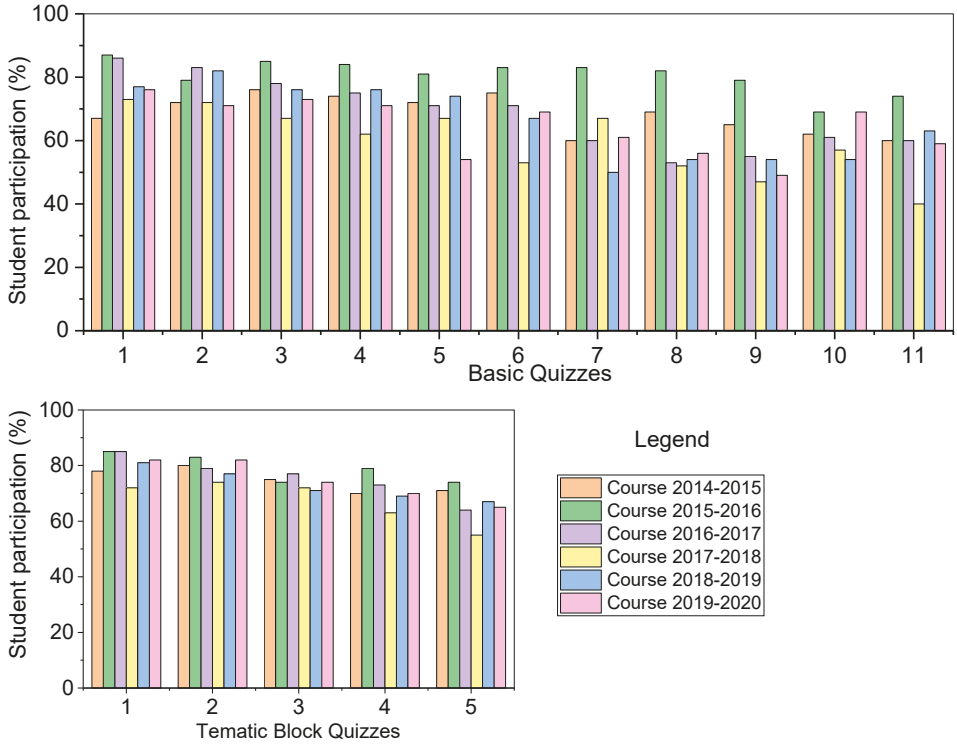

Figure 1. Participation in basic (BQ) and thematic block (TBQ) quizzes during the last six academic years. Own elaboration based on this study.

A detailed analysis by academic year allows us to know the dynamics and evolution of the participation. The participation falls down in the last quizzes, being always slightly lower than the first ones. This decrease is more striking in the 2016–2017 and 2017–2018 academic years, which go from approximately 85% and 75% in the first BQ to 60% and 40% in the last quiz, respectively, while it goes from 85% and 70% to 60% and 55% in the TBQs, respectively.

The general trend is a progressive decrease in participation throughout the semester in any academic year. This is due to several factors, such as possible changes in the enrolment of the subject given that some students are waiting for a possible change to another degree at the beginning of the course, and this process does not materialize until after a month, but in the meantime, they have been taking the quizzes. Moreover, mid-semester partial exams of other matters are held, so students are immersed in the study of other subjects and end up not doing the quizzes, either because the time has passed to do it, or because they have not studied. Additionally, at the end of the semester, a large number of students have decided to drop out of the degree in chemistry and are not involved in the training activities. The dropout rate in this first-degree course is approximately 20–25%. In the initial survey of the class, 25% of the students consider that the degree in chemistry is not their vocation and it was not their first option in the university pre-registration.

3.2. Statistical and Psychometric Data of the Quizzes and Each Item

The results provided directly by the Moodle platform (https://docs.moodle.org/dev/Quiz_statistics_calculations (accessed on 22 July 2021)) [20,21] have been analyzed and calculated according to the classical test theory [32,33]. Supplementary Materials summarize the definition of psychometric parameters. Tables 2 and 3 collect, for each quiz,

statistical data such as the average score, the standard deviation (SD), the range of correct answers (maximum and minimum percentage), also called the facility index (FI), and the asymmetry in the distribution of the scores, also called bias, together with the internal consistency coefficient (ICC), or Cronbach's alpha, which gives an idea of the quality of the tests and allows to recognize if the whole exam is homogeneous.

Table 2. Statistical data corresponding to the eleven basic quizzes (BQ) in different academic years.

	2014–2015					2015–2016				
BQ	**FI (%)**	**Average Score ± Error**	**SD (%)**	**Bias**	**ICC (%)**	**FI (%)**	**Average Score ± Error**	**SD (%)**	**Bias**	**ICC (%)**
1	68–94	8.24 ± 1.20	24.36	−1.29	78.85	31–95	6.98 ± 1.43	21.34	−0.28	54.55
2	78–97	9.04 ± 0.80	16.50	−2.90	76.50	83–97	9.35 ± 0.63	13.40	−2.38	74.30
3	88–98	9.47 ± 0.65	10.26	−2.20	59.52	80–100	9.24 ± 0.76	12.94	−1.90	65.23
4	86–97	9.11 ± 0.79	15.66	−2.38	74.45	85–100	9.46 ± 0.66	9.50	−2.18	50.73
5	90–99	9.41 ± 0.61	14.89	−2.99	82.97	90–98	9.44 ± 0.65	12.04	−2.33	70.83
6	81–95	8.36 ± 0.78	21.20	−2.50	86.40	15–97	8.48 ± 0.73	17.90	−2.62	83.60
7	71–100	9.22 ± 0.59	10.50	−1.66	67.90	47–97	8.53 ± 0.73	15.40	−2.86	77.60
8	85–98	9.36 ± 0.66	13.90	−2.46	77.30	93–97	9.56 ± 0.53	13.20	−5.47	84.00
9	95–100	9.77 ± 0.45	6.80	−3.38	56.40	89–98	9.62 ± 0.56	9.10	−2.93	62.30
10	80–92	8.46 ± 0.85	20.80	−2.08	83.50	37–90	6.51 ± 0.10	27.80	−0.54	84.70
11	93–100	9.68 ± 0.48	10.00	−5.42	76.90	58–100	8.93 ± 0.89	10.50	−1.38	27.80
	2016–2017					**2017–2018**				
BQ	**FI (%)**	**Average Score ± Error**	**SD (%)**	**Bias**	**ICC (%)**	**FI (%)**	**Average Score ± Error**	**SD (%)**	**Bias**	**ICC (%)**
1	81–98	9.35 ± 0.78	13.07	−2.85	64.11	84–98	8.93 ± 0.88	19.54	−2.07	79.34
2	85–100	9.54 ± 0.63	8.30	−2.09	42.50	87–98	9.38 ± 0.63	14.80	−2.78	81.60
3	97–100	9.86 ± 0.37	3.50	−2.10	−15.84	74–98	9.01 ± 0.85	14.09	−1.53	62.9
4	77–100	9.25 ± 0.73	13.16	−1.98	68.83	87–96	9.22 ± 0.75	14.19	−1.83	71.78
5	91–98	9.63 ± 0.53	10.03	−3.47	72.16	85–98	9.04 ± 0.78	17.62	−1.78	80.33
6	52–95	8.48 ± 0.88	21.30	−2.11	82.90	77–98	8.66 ± 0.82	23.60	−1.96	87.90
7	62–100	9.11 ± 0.68	10.30	−1.24	56.00	83–100	8.98 ± 0.71	16.80	−1.76	82.10
8	86–96	9.10 ± 0.82	14.70	−1.75	68.70	92–98	9.48 ± 0.64	11.60	−2.20	70.10
9	82–100	9.35 ± 0.72	10.90	−1.71	56.10	91–100	9.53 ± 0.62	10.00	−2.49	61.30
10	51–94	7.73 ± 1.02	27.00	−1.39	85.80	74–97	8.36 ± 0.79	25.00	−1.90	90.10
11	88–100	9.55 ± 0.60	9.60	−2.96	60.80	82–100	9.43 ± 0.68	10.10	−1.53	54.90
	2018–2019					**2019–2020**				
BQ	**FI (%)**	**Average Score ± Error**	**SD (%)**	**Bias**	**ICC (%)**	**FI (%)**	**Average Score ± Error**	**SD (%)**	**Bias**	**ICC (%)**
1	81–100	9.16 ± 0.78	18.01	−2.36	80.85	95–65	8.59 ± 0.89	21.37	−1.57	76.43
2	90–97	9.54 ± 0.59	11.10	−2.90	71.50	99–89	9.34 ± 0.67	12.40	−2.38	66.50
3	93–98	9.65 ± 0.53	8.88	−2.90	64.26	93–100	9.46 ± 0.62	10.81	−2.72	62.93
4	89–97	9.47 ± 0.64	11.37	−2.64	68.22	86–97	9.26 ± 0.61	16.98	−3.36	84.93
5	93–100	9.48 ± 0.61	12.41	−2.77	76.22	91–100	9.52 ± 0.52	13.70	−3.21	84.02
6	82–97	8.88 ± 0.79	22.40	−2.26	87.60	75–99	9.05 ± 0.65	16.20	−2.07	80.30
7	96–100	9.83 ± 0.31	7.50	−5.40	82.90	85–97	9.25 ± 0.53	15.70	−2.40	86.30
8	92–100	9.69 ± 0.53	7.50	−3.05	50.50	93–100	9.79 ± 0.42	6.20	−4.47	51.90
9	92–100	9.83 ± 0.39	5.00	−3.96	40.50	84–100	9.45 ± 0.55	13.90	−3.56	81.80
10	60–100	9.29 ± 0.50	10.80	−4.08	78.70	70–98	9.05 ± 0.59	18.90	−3.26	87.70
11	50–100	8.24 ± 0.96	11.90	−0.54	34.90	93–100	9.90 ± 0.31	3.60	−3.80	25.50

FI (Facility Index); SD (Standard Deviation); ICC (Internal consistency coefficient, Cronbach's alpha). Own elaboration based on this study.

The average score of any BQ in any academic year is high, between remarkable and outstanding (6.51 for BQ-10 and 9.83 for BQ-7 in the 2015–2016 and 2018–2019 academic years, respectively), with a high percentage of correct answers in each quiz that ranges from 70% to 100%, except in the 2015–2016 academic year where a minimum success rate of 15%, 47% and 37%, was obtained in the BQ-6, BQ-7 and BQ-10, respectively, which correspond to the two most difficult topics to assimilate: thermodynamics and electrochemistry. This large range in the correct answers yields an asymmetric distribution with a negative bias greater than −1 in all academic years. That indicates the lack of discrimination among those students who do better than the average ratio, and it is due to the fact that most of the items are classified as basic knowledge, and also to the type of question (true/false)

which shows a random response of 50%. The standard deviation is practically around 20%, except in some cases with a slightly higher value, between 22 and 28%, in those quizzes corresponding to the topics of thermodynamics and electrochemistry.

Table 3. Statistical data corresponding to the five thematic block quiz (TBQ) in different academic years.

	2014–2015					2015–2016				
TBQ	**FI (%)**	**Average Score ± Error**	**SD (%)**	**Bias**	**ICC (%)**	**FI (%)**	**Average Score ± Error**	**SD (%)**	**Bias**	**ICC (%)**
1	62–76	6.93 ± 1.57	20.20	−0.76	39.60	53–78	6.36 ± 1.66	17.70	−0.49	12.40
2	52–69	6.08 ± 1.66	19.00	−0.14	23.40	50–64	5.74 ± 1.68	21.60	0.17	39.60
3	48–53	5.16 ± 1.62	20.90	−0.30	39.80	48–58	5.18 ± 1.59	21.60	−0.09	45.40
4	34–52	4.48 ± 1.63	25.30	−0.13	58.70	39–66	5.30 ± 1.64	22.60	−0.24	47.50
5	28–51	3.93 ± 1.70	21.90	0.18	40.00	38–54	4.70 ± 1.68	24.30	−0.37	52.20
	2016–2017					**2017–2018**				
TBQ	**FI (%)**	**Average Score ± Error**	**SD (%)**	**Bias**	**ICC (%)**	**FI (%)**	**Average Score ± Error**	**SD (%)**	**Bias**	**ICC (%)**
1	61–81	6.97 ± 1.55	19.80	−0.58	38.60	62–74	7.01 ± 1.52	21.60	−0.78	50.40
2	61–76	5.80 ± 1.57	18.50	−0.72	28.20	50–64	5.23 ± 1.72	21.50	−0.66	35.90
3	40–60	5.23 ± 1.62	22.90	−0.03	50.00	42–61	5.14 ± 1.62	23.70	−0.07	53.10
4	48–66	5.86 ± 1.68	22.40	−0.30	43.80	46–76	5.82 ± 1.65	21.40	−0.73	40.40
5	42–64	5.68 ± 1.67	19.40	−0.35	26.50	50–73	5.88 ± 1.57	23.10	−0.33	54.10
	2018–2019					**2019–2020**				
TBQ	**FI (%)**	**Average Score ± Error**	**SD (%)**	**Bias**	**ICC (%)**	**FI (%)**	**Average Score ± Error**	**SD (%)**	**Bias**	**ICC (%)**
1	55–78	6.81 ± 1.57	19.50	−0.64	35.50	61–75	6.99 ± 1.09	19.4	−1.02	35.40
2	51–70	6.24 ± 1.65	20.60	−0.50	35.20	50–68	6.00 ± 1.00	21.0	−0.28	38.40
3	43–58	5.15 ± 1.67	19.40	0.07	25.70	49–57	5.31 ± 0.86	22.3	−0.51	46.50
4	41–62	5.32 ± 1.76	16.80	0.03	−10.30	49–65	5.73 ± 0.93	19.9	−0.07	32.30
5	44–56	4.96 ± 1.69	18.70	0.02	18.30	43–56	4.95 ± 0.81	20.6	−0.48	36.50

FI (Facility Index); SD (Standard Deviation); ICC (Internal consistency coefficient, Cronbach's alpha). Own elaboration based on this study.

The ICC in most quizzes at any academic year is higher than 65%, the minimum value proposed as indicator of an overall homogeneity of the quiz [34]. However, in some cases, values lower than 65% have been obtained, for example in the BQ-3 and BQ-9 in the 2014–2015 academic year, even significantly lower values of 27.80% in the BQ-11 of the 2015–2016 academic year or even negative values of −15.84% in the BQ-3 of the 2016–2017 academic year. These results show the limitation of this parameter in considering that the quiz measures with the same precision all the students evaluated when it really depends on the level of each student and, ultimately, on the population used to calculate it.

Moreover, the dispersion of the IF and the discrimination index (DI) for each item of any quiz have been analyzed in order to know the item effectiveness to discern between students with different cognitive ability (Figure S1, left). Most of the questions have an adequate discrimination, with a DI above 30%. A more detailed analysis of the discriminative efficiency (DE) for any item of the different BQ (Figure S2) shows that the effectiveness of the items depends on the academic year and, therefore, on the student population. For example, in the academic year 2015–2016, all items of the BQ1 do not reach 30% of DE, while in the 2018–2019 academic year, they are all above 30%. The same behavior has been found in other items corresponding to other quizzes. Therefore, the same item may or may not be discriminatory for a population of students depending on the level of knowledge they have, and therefore, questions that have a low DI should not be discarded. It is concluded that quizzes made only with true/false items serve as continuous training activities in the teaching–learning process of a matter, not being feasible as assessment activities because they are not discriminatory for students.

Different results are obtained for the TBQs (Table 3). The average score drops significantly with respect to the BQ quizzes, from 5.16 in the TBQ-3 of the 2014–2015 academic year to 7.01 in the TBQ-1 of the academic year 2017–2018, even in some cases reaching

values of 3.9 or 4.5 in the TBQ-4 and CBT-5 of the 2014–2015 academic year, for example. Moreover, the FI index drops significantly and ranges from 28% in the TBQ-5 of the 2014–2015 academic year to 81% in the TBQ-1 of the 2016–2017 academic year, but in no case does it reach 100% in any of the items in any quiz. The dispersion in the average scores oscillates around 20%, being slightly high in the last three CBTs of certain academic courses. The asymmetry of the distribution in the scores, the so-called bias, is still negative, but now with a value lower than −1, reaching a slightly positive value and with an almost symmetric distribution, with a bias close to zero in the last three quizzes of the 2018–2019 academic year with a value between 0.07 and 0.02.

As a general trend, the bias in the first two quizzes, TBQ-1 and TBQ-2, is somewhat higher than in the rest of the TBQ-3, TBQ-4 and TBQ-5 ones that have a bias close to zero. This indicates that the topics corresponding to the first two blocks are better assimilated than the rest of the topics corresponding to the blocks of thermodynamics, electrochemistry and kinetics. This effect is probably due to the fact that the first topics are already studied in the bachelor grade, while the topics of the last blocks are totally new, which means an effort in learning process.

Therefore, these TBQs are more discriminative between students than BQs, that is, here the cognitive abilities of each student are tested. The FI-DI scatter diagrams are shown in Figure S1 right. Although a low DI value is obtained due to the random characteristic of the quiz, the detailed analysis of the FI-DI diagrams for all items in any TBQ (Figure S3) reveals that most of items are discriminative with a DI value above 30% and with a wide range in the FI. However, the ICC is always lower than the reference value of 65% [34]. This is because the items of the quiz have been randomly selected by the Moodle platform and show different questions for each student.

3.3. Students' Opinion on the Educational Activities

A survey was performed among the students to find out their opinion on these two types of virtual quizzes. It is a short survey with only eight statements, four for each type of quizzes, in which students have to answer YES or NO to the proposed statement. Although this questionnaire was not contrasted by the scientific community and its experimental validity was not demonstrated, it only tries to show the students' opinion on Moodle quiz activities in few ideas. Table 4 shows the average values obtained in the BQ and TBQ, respectively, during the last six academic years. These average values practically coincide with those obtained for each academic course; therefore, the students' opinion about how they apperceive the virtual quizzes, practically, does not change with the years.

Table 4. Students' opinion on the BQ and TBQ. Average results of the last six academic years. Own elaboration based on this study.

	Statements	**Responses (Average Values)**		
	Basic Quizzes (BQ)	**YES**	**NO**	**NK/NA**
1.-	The difficulty of the question in these quizzes has been adequate for a basic level	89.9%	5.4%	4.7%
2.-	The time limit of one hour is sufficient to take the quiz	89.2%	5.4%	5.4%
3.-	The quizzes allow you to study the subject continuously	59.4%	32.4%	8.2%
4.-	The quizzes are used for self-evaluation of the subject and know the level of knowledge acquired	67.6%	27.1%	5.3%
	Statements	**Responses (Average Values)**		
	Thematic Block Quizzes (TBQ)	**YES**	**NO**	**NK/NA**
1.-	The level of difficulty of these quizzes is higher than that of the basic ones	91.9%	5.4%	2.7%
2.-	The time limit of one hour is sufficient to take the quiz	51.3%	45.9%	2.8%
3.-	These quizzes carried out on dates programmed allow you to study the topics before evaluation	62.1%	35.1%	2.8%
4.-	These quizzes serve to strengthen the acquired knowledge	59.4%	37.8%	2.8%

Regarding BQ, most of the students, about 90%, tell that the difficulty of the items is adequate for a basic level and the time of one hour is more than enough. These BQs help a lower percentage 60% to study the matter continuously, while the rest, 30%, think that they do not study continuously, even though they have to do the online quizzes. However, the BQs serve as a self-assessment of the knowledge acquired in class for almost 70%. Only 5–8% of students do not know or do not answer (NK/NA) to each of the proposed statements. The results obtained for the TBQs are a little bit different.

Regarding the TBQ, practically all students, 92%, say that the level of difficulty in these quizzes is greater than that of the BQ. This was predictable because one of the goals of these TBQ was to measure the level of abstract reasoning and assimilation of theoretical concepts. Only 50% of the students say that time limit of one hour is sufficient. On the other hand, these quizzes scheduled in fixed days allow more than 60% of students to study the days before doing the test. This fact consolidates the knowledge acquired in the classroom. Only 3% of the students answer NK/NA to the proposed statements.

In short, both types of quiz fulfill the goal of favoring a continuous study of the matter along the semester, avoiding the general trend to study only the days before the written exam. Moreover, all these educational activities into an electronic environment promote an online-accessible medium for the student that serves as a self-assessment of the level of knowledge acquired.

3.4. Final Scores in General Physical Chemistry

Figure 2 shows the final scores in the matter of general physical chemistry corresponding to the two Ordinary Calls of exams (February and September), between 2011 to 2014 courses following a traditional teaching methodology, and between 2014 to 2020 courses following the new methodology with quizzes.

In the first Ordinary Call (February), a slight decrease in the percentage of students not presented and also in that of failed following the new methodology is observed. However, it is interesting to note the increase in the number of passes, not only with the minimum score but also with remarkable, outstanding and even honors (H) in the last five academic years, highlighting the academic year 2016–2017 where the percentage of outstanding is higher than in the rest of the courses.

In the second Ordinary Call (September), the percentage of students not presented remains almost constant in all the courses developed with the methodology based on quizzes and is practically the same as that of the first courses taught with the traditional methodology. There is not much variation in the percentage of failures and approved following one methodology or another, except in the last academic year 2019–2020, in which the number of not presented decreases while the number of approved increases. In this September call, there is a low percentage of outstanding students and the absence of scores above the remarkable following the new methodology with quizzes.

These online education activities help not only students with high cognitive capacity to attain good scores but also those with a medium level to pass the exam in the first call of February. The percentage of students that remain for the second call in September are really those who have not assimilated the physical chemistry knowledge and those who find it difficult to make scientific reasoning or deductions.

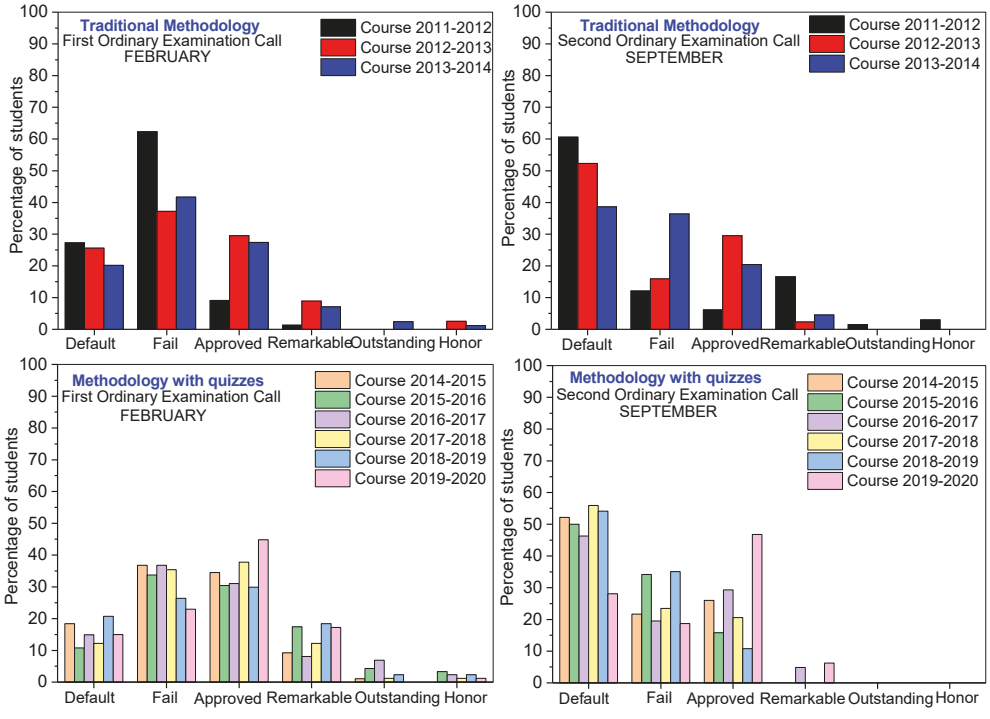

Figure 2. Scores of General Physical Chemistry in Ordinary Calls of exams. Own elaboration based on this study.

4. Conclusions

The use of Moodle quizzes, as online activities, favors the implementation of a different educational methodology in the subject of general physical chemistry of the Degree in Chemistry. However, not all quizzes can be used as assessment tools given that the item type and the quiz settings play an important role. In this work, two different types of quiz are proposed, basic (BQ) and thematic block (TBQ). BQ has only true/false items and these are the same for all students, while TBQ has multiple choice, matching and numerical items that are randomly selected from a category of the question bank.

Statistical and psychometric data provided by Moodle platform were analyzed. Most of items show a discrimination (DI) and facility (FI) index according to the proposed Moodle reference range. The FI-DI dispersion graphs, together with the average scores and the bias values, show the different quality of the two types of quiz. The BQs can be used as formative teaching activities because they have not enough evaluative quality to distinguish the different capacities and abilities of students, yielding in any quiz higher average scores and FI values, and also a large negative bias. TBQs are more discriminative, showing lower values of both FI and average scores, with a bias near to zero value, and thus, these can discern competencies and skills among students, so they could be used as assessment tools. As a conclusion, the true/false item must not be used as evaluative item. In the future, it is necessary to study how each type of item (multiple choice, numerical and matching) contributes individually to the final scores of the quiz in a similar performance. This would allow us to select the best type of item to use it in a particular assessment quiz.

Although the methodology applied is weak, the performance developed along the years indicates that TBQs work quite well to evaluate the physical chemistry knowledge and the different capacities of students, independent of the student population, and it could even be extrapolated to other similar scientific scenes.

The analysis of statistic data in TBQs, like FI index, allow the teacher to know the topics that were not well understood by the students and gives a feedback on learning process of the students. Moreover, these quizzes work as a self-assessment for the students, providing them a better preparation for the written exam. The performance of these quizzes has not supposed an excessive workload for the students, allowing improved scores obtained in the first Ordinary Call of February, unlike those obtained following a traditional methodology in previous courses.

In few words, the Moodle statistics point out that a particular quiz has assessment quality if the bias is symmetric, around zero value, and the range of FI and DI is about 40–70%. This feature can be obtained using a huge question bank with the aim to perform a random quiz. Mixed items (multiple choice, numerical and matching) also contribute to achieving a quiz in which different skills and abilities of physical chemistry have been inquired. In addition, the similar results obtained in this experience along the years, with different students' population, would also prove the validity and reliability of the designed quizzes. This study shows how the analysis of statistical and psychometric parameters allows checking if the design of an educative activity based on Moodle quizzes can be used as an assessment tool to evaluate the skill and competences of a particular subject along the semester.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/educsci11090500/s1, Definition of psychometric parameters, Figure S1: FI-DI diagrams corresponding to the two quiz types, Figure S2: DE diagrams corresponding to the ten items (true/false) of each BQ, Figure S3: FI-DI diagrams of the different items in each TBQ.

Funding: This research and the APC were funded by University of Málaga, through Educational Innovation Projects, PIE15-027 and PIE19-051.

Institutional Review Board Statement: Not applicable.

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

Data Availability Statement: All data reported in this work are based on the present study with an own elaboration.

Acknowledgments: The author thanks to Servicio Central de Informática of University of Málaga for technical support.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Valverde-Berrocoso, J.; Garrido-Arroyo, M.C.; Burgos-Videla, C.; Morales-Cevallos, M.B. Trends in Educational Research about e-Learning: A Systematic Literature Review (2009–2018). *Sustainability* **2020**, *12*, 5153. [CrossRef]
2. Linuma, M. *Learning and Teaching with Technology in the Knowledge Society*; Springer: Singapore, 2016.
3. Luisa Sevillano-García, M.A.; Vázquez-Cano, E. The Impact of Digital Mobile Devices in Higher Education. *J. Educ. Technol. Soc.* **2015**, *18*, 106–118. Available online: http://www.jstor.org/stable/jeductechsoci.18.1.106 (accessed on 10 March 2021).
4. Delialioglu, O.; Yildirim, Z. Design and development of a technology enhanced hybrid instruction bases on MOLTA model: Its effectiveness in comparison to traditional instruction. *Comput. Educ.* **2008**, *51*, 474–483. [CrossRef]
5. Ryberg, T.; Niemczik, C.; Brenstein, E. Methopedia-Pedagogical Design for European Educators. In Proceedings of the 8th European Conference on e-Learning, Bari, Italy, 29–30 October 2009; pp. 503–511.
6. Tamin, R.M.; Bernard, R.M.; Borokhovski, E.; Abrami, P.C.; Schmid, R.F. What Forty Years of Research Says About the Impact of Technology on Learning: A Second-Order Meta-Analysis and Validation Study. *Rev. Educ. Res.* **2011**, *81*, 4–28. [CrossRef]
7. Francis, R.; Shannon, S.J. Engaging with blended learning to improve students' learning outcomes. *Eur. J. Eng. Educ.* **2013**, *38*, 1–11. [CrossRef]
8. Kesici, S.; Sahin, I.; Akturk, A.O. Analysis of cognitive learning strategies and computer attitudes, according to college students' gender and locus of control. *Comput. Hum. Behav.* **2009**, *25*, 529–534. [CrossRef]
9. López-Pérez, M.V.; Pérez-López, M.C.; Rodríguez-Ariza, L. Blended learning in higher education: Students' perceptions and their relation to outcomes. *Comput. Educ.* **2011**, *56*, 818–826. [CrossRef]
10. Randy Garrison, D.; Vaughan, N.D. *Blended Learning in Higher Education: Framework, Principles, and Guidelines*; John Wiley & Sons, Inc.: New York, NY, USA, 2008; ISBN 978-1-118-26955-8.

11. Anthony, B.; Kamaludin, A.; Romli, A.; Raffei, A.F.M.; Phon, D.N.; Abdullah, A.; Ming, G.L. Blended Learning Adoption and Implementation in Higher Education: A Theoretical and Systematic Review. *Technol. Knowl. Learn.* **2020**, *25*, 1–48. [CrossRef]
12. Bidarra, J.; Rusman, E. Towards a pedagogical model for science education: Bridging educational contexts through a blended learning approach. *Open Learn. J. Open Distance E-Learn.* **2017**, *32*, 6–20. [CrossRef]
13. Jamshidifarsani, H.; Tamayo-Serrano, P.; Garbaya, S.; Lim, T. A three-step model for the gamification of training and automaticity acquisition. *J. Comput. Assist. Learn.* **2021**, *37*, 994–1014. [CrossRef]
14. Kirzner, R.S.; Alter, T.; Hughes, C.A. Online Quiz as Exit Ticket: Using Technology to Reinforce Learning in Face to Face Classes. *J. Teach. Soc. Work* **2021**, *41*, 151–171. [CrossRef]
15. Manzano-Leon, A.; Camacho-Lazarraga, P.; Guerrero, M.A.; Guerrero-Puerta, L.; Aguilar-Parra, J.M.; Trigueros, R.; Alias, A. Between Level Up and Game Over: A Systematic Literature Review of Gamification in Education. *Sustainability* **2021**, *13*, 2247. [CrossRef]
16. Lo, C.; Tang, K. Blended Learning with Multimedia e-Learning in Organic Chemistry Course. In Proceedings of the 2018 International Symposium on Educational Technology (ISET), Osaka, Japan, 31 July–2 August 2018; pp. 23–25. [CrossRef]
17. Nash, S.S.; Rice, W. *Moodle 3 E-Learning Course Development*; PACKT Publisher: Birmingham, UK, 2015.
18. Brophy, J.; Alleman, J. Assessment in a Social Constructivist Classroom. *Soc. Educ.* **1998**, *62*, 32–34.
19. Gold, S. Constructivist approach to online training for online teachers. *J. Asynchronous Learn. Netw.* **2001**, *5*, 35–57. Available online: https://wikieducator.org/images/f/fb/ALN_Constructivist_Approach.pdf (accessed on 2 March 2021). [CrossRef]
20. Gamage, S.H.; Ayres, J.R.; Behrend, M.B.; Smith, E.J. Optimising Moodle quizzes for online assessment. *Int. J. STEM Educ.* **2019**, *6*, 1–14. [CrossRef]
21. Gómez-Soberón, J.; Gómez-Soberón, M.C.; Corral-Higuera, R.; Arredonde-Rea, S.P.; Almaral-Sánchez, J.L.; Cabrera-Cavarrubias, F.G. *Calibrating Questionnaires by Psychometric Analysis to Evaluate Knowledge*; SAGE: Thousand Oaks, CA, USA, 2013. [CrossRef]
22. Jaeger, M.; Adair, D. Time pressure in scenario-based online construction safety quizzes and its effect on students' performance. *Eur. J. Eng. Educ.* **2017**, *42*, 241–251. [CrossRef]
23. Krause, C.M.A.; Krause, R.B.S.; Krause, R.B.S.; Gomez, N.G.M.D.; Jafry, Z.M.D.; AmDinh, V. Effectiveness of a 1-Hour Extended Focused Assessment with Sonography in Trauma Session in the Medical Student Surgery Clerkship. *J. Surg. Educ.* **2017**, *74*, 968–974. [CrossRef] [PubMed]
24. Sullivan, D.P. An Integrated Approach to Preempt Cheating on Asynchronous, Objective, Online Assessments in Graduate Business Classes. *Online Learn.* **2016**, *20*, 195–209. Available online: http://onlinelearningconsortium.org/read/online-learning-journal/ (accessed on 4 May 2021). [CrossRef]
25. Burden, P. ELT teacher views on the appropriateness for teacher development of end of semester student evaluation of teaching in a Japanese context. *System* **2008**, *36*, 478–491. [CrossRef]
26. Nulty, D.D. The adequacy of response rates to online and paper surveys: What can be done? *Assess. Eval. High. Educ.* **2008**, *33*, 301–314. [CrossRef]
27. Shepard, L.A. *The Assessment in the Classroom: Texts Evaluation*; National Institute for Educational Evaluation: Distrito Federal, Mexico, 2006.
28. Ferrao, M. E-assessment within the Bologna paradigm: Evidence from Portugal. *Assess. Eval. High. Educ.* **2010**, *35*, 819–830. [CrossRef]
29. Blanco, M.; Ginovart, M. Moodle quizzes for assessing statistical topics in engineering studies. In Proceedings of the Joint International IGIP-SEFI Annual Conference 2010, Trnava, Slovakia, 19–22 September 2010. Available online: https://upcommons.upc.edu/bitstream/handle/2117/9992/blanco_ginovart_igip_sefi_2010.pdf (accessed on 9 August 2021).
30. Blanco, M.; Ginovart, M. On How Moodle Quizzes Can Contribute to the Formative e-Assessment of First-Year Engineering Students in Mathematics Courses. *RUSC Univ. Knowl. Soc. J.* **2012**, *9*, 166–183. [CrossRef]
31. Crocker, L.; Algina, J. *Introducction to Classical and Modern Test Theory*; Holt Rinehart and Winston: New York, NY, USA, 1986.
32. Wang, Z.; Osterlind, S.J. Classical Test Theory. In *Handbook of Quantitative Methods for Educational Research*; Teo, T., Ed.; Sense Publishers: Rotterdam, The Netherlands, 2013. [CrossRef]
33. Shrout, P.E.; Lane, S.P. Psychometrics. In *Handbook of Research Methods for Studying Daily Life*; Mehl, M.R., Conner, T.S., Eds.; The Guilford Press: New York, NY, USA, 2012; pp. 302–320.
34. GNU. General Public License. 2013. Available online: https://docs.moodle.org/dev/Quiz_report_statistic (accessed on 7 July 2020).

education sciences

MDPI

Article

Authentic Assessment Implementation in Natural and Social Science

Eddy Sutadji [1,*], Herawati Susilo [2], Aji Prasetya Wibawa [1], Nidal A. M. Jabari [3] and Syaiful Nur Rohmad [1]

[1] Faculty of Engineering, Universitas Negeri Malang, Malang 65145, Indonesia; aji.prasetya.ft@um.ac.id (A.P.W.); syaifulnurrohmad@gmail.com (S.N.R.)
[2] Department of Biology, Faculty of Mathematics and Natural Science, Universitas Negeri Malang, Malang 65145, Indonesia; herawati.susilo.fmipa@um.ac.id
[3] Department of Applied Computing, Palestine Technical University Kadoorie, Arroub Camp St., Hebron P.O. Box 7, Palestine; nidal.jabari@ptuk.edu.ps
* Correspondence: eddy.sutadji.ft@um.ac.id

Abstract: Assessment methods are important to create qualified graduates who are ready to face the real world. Authentic assessment is considered to be the most effective method to achieve this. The application of authentic assessment is often universal. However, there is a difference between natural sciences and social sciences. If it is used for different scientific constructions, then the authentic assessment should also be different. Therefore, there is a need for authentic implementation research in these two fields of science. This research is survey research with quantitative descriptive method. This study focuses on the analysis of differences in implementation of the assessment carried out, assignment techniques, assessment components, and post-assessment at the State University of Malang in two different fields of science, namely natural sciences and social sciences. The population in this study was 1069 lecturers represented by 270 sample lecturers. There are 106 (39.26%) samples of lecturers representing 388 (36.3%) lecturer populations from 2 natural fields and 164 (60.74%) samples representing 681 (63.7%) lecturer populations from 6 social fields. The analysis is carried out by comparing the results of each aspect of the assessment implementation in the two fields. Almost all aspects of authentic assessment between the natural and social sciences had no difference. The only differences were in the assessment form and individual assignment techniques that were performed. Social science conducted non-test assessment only higher than the natural science. Measured tests were primarily used in the natural science using Higher-Order Thinking Skills questions. Performance test was mostly conducted in social science.

Keywords: authentic assessment; education evaluation; natural science; social science

Citation: Sutadji, E.; Susilo, H.; Wibawa, A.P.; Jabari, N.A.M.; Rohmad, S.N. Authentic Assessment Implementation in Natural and Social Science. *Educ. Sci.* **2021**, *11*, 534. https://doi.org/10.3390/educsci11090534

Academic Editors: Sandra Raquel Gonçalves Fernandes, Marta Abelha and Ana Teresa Ferreira-Oliveira

Received: 14 July 2021
Accepted: 7 September 2021
Published: 13 September 2021

1. Introduction

The world is developing very rapidly. This causes the education world to evolve very quickly too. Students are not only required to understand and memorize but also able to analyze critically and practice well in the real world. In addition to the teaching and learning condition factors, the assessment method factor is essential to create qualified graduates who are ready to face the real world [1,2]. Authentic assessment is considered to be the most effective assessment method to achieve this.

The selection of authentic assessment over traditional assessment is based on the advantages of it. Moreover, there are several differences between authentic and traditional assessment. The traditional assessment is product oriented, while the authentic assessment is process oriented. It means that traditional assessment used to evaluate the subject knowledge by comparing it against standards or other learners. While authentic assessment aims to evaluate the subject proficiency by asking them to perform real-life tasks. Based on this purpose, an authentic assessment has an advantage in which it provides a true picture of

the students' learning conditions, gives more information about students' strengths, weaknesses, needs, and preferences that can assist in adjusting instruction toward improving learning activities.

Moreover, authentic assessment is regarded as a better and real approach in the assessment. This approach associates learning in real and quite complex situations and contexts [3–5]. The assessment is based on student practice in the real world [6,7] that cannot be done by traditional testing. Traditional testing cannot explore real changes in student knowledge. On the other hand, an assessment approach that emphasizes the learning process and encourages students to carry out cognitive and reflective activities follows the constructivist concepts. In contrast, authentic assessment reflects these alternative assessment techniques. This assessment is based on authentic learning assignments and not a separate test and focuses as much on the process, as is the product [4]. Authentic assessment requires skills implementation that is needed in the classroom and the utilization to support further learning [8,9]. In this assessment, students demonstrate their skills toward an attitude in a real-life context and are assessed based on this fundamental performance [10].

The authentic assessment shows the application of specific student skills and assessments and focuses more on problem-solving, understanding, critical thinking, reasoning, and metacognition. Therefore, students are expected to handle meaningful material and problem solving through authentic assessment [11]. In other words, authentic assessment requires students to use prior knowledge, current teaching, and skills to solve real and complex problems. For example, students can create projects on their chosen topic, prepare research reports, and present their final products to evaluators [9]. On the other hand, research of [12] provided evidence that collaborative concepts in the peer assessment process, self-assessment, and assessment conducted by lecturers provide valid and authentic data on student performance.

The application of authentic assessment is often formulated into universal terms even though there is a difference between natural science and social science. The social and natural sciences tend to approach the question of truth and objective reality in different manners. Social science emphasizes that all human representations of reality are contingent constructs; some will even argue against it and feel it is meaningless to talk about the objective reality [13]. Meanwhile, natural sciences often visualize and conclude an accurate and realistic representation. The difference between constructivism and realism tends to lead to different positions [14,15]. The main interest in these two disciplines is very different. Natural or natural science predicts and manifests all-natural phenomena, whereas social science predicts and explains human behavior and psychology [16]. If scientific construction is different, the authentic assessment should also be different. Therefore, research is needed on the implementation of authentic assessment in these two fields of science, so that the future authentic assessment is no longer considered something universal but having different implementations between natural science and social science. This study aims to explore the facts of authentic assessment implementation in two fields of science: social and natural sciences, to clearly described the authentic assessment.

The subjects of this study were lecturers in eight faculties at the State University of Malang. The classification of social science and natural science is seen through the criteria for new student admissions in Indonesia, several faculties are included in the classification of social sciences and humanities such as the Faculty of Education, Faculty of Letters, Faculty of Economics, Faculty of Social Sciences, Faculty of Sports Science, and Faculty of Psychology Education. Meanwhile, those classified into natural sciences are the Faculty of Engineering and the Faculty of Mathematics and Natural Sciences.

The novelty of this research is that there is a comparison between the implementation of authentic assessments in two major fields of science, namely natural and social science for all Higher Education Institutions that use a life-based learning approach (LBL) which aims to develop student capacity holistically. Previous research has discussed more about

approaches and differences in scientific context, has not discussed the implementation of authentic assessments carried out in these two fields.

2. Research Methods

This research was survey research with quantitative descriptive method. This study focused on the elaboration of lecturers on the performed assessment, assignment techniques, assessment components, and post-assessment of authentic assessment at the State University of Malang as an educational provider institution for prospective educators in two different fields of science, natural science and social science.

The population and samples are adjusted to the scope and objectives of the research. The population is the entire research subject [17,18]. The population in this study were all lecturers of the State University of Malang. The samples were taken randomly and were representatives of each faculty in the State University of Malang. The number of samples was calculated using the Slovin formula [18] with the error tolerance limit of 5%.

$$n = N/(1 + N.(e)2) \quad (1)$$

Note:

n = Total Samples
N = Total Population
e = Error Tolerance Limit

To consider the number of samples at each faculty, as seen in Table 1, the calculation was carried out proportionally using the proportional allocation formula [19,20].

$$ni = (Ni/N)n \quad (2)$$

Note:

ni = Total samples from each faculty
Ni = Total population in the faculty
N = Overall total population
n = Overall total samples

Table 1. Number of samples for each faculty.

Science	Faculty	Population	Proportional Sample
Natural Science	Faculty of Mathematics and Science	201	55
	Faculty of Engineering	187	51
Social Science	Faculty of Letter	197	53
	Faculty of Economics	145	39
	Faculty of Education	149	40
	Faculty of Sport Science	57	15
	Faculty of Social Science	105	9
	Faculty of Psychology Education	28	8
	TOTAL	1069	270

A Processed from 2020 UM Statistics.

Data collection using primary data was collected through a questionnaire in the form close-ended question. The results of the study were analyzed using quantitative descriptive based on the highest number of scores for each of the provided options and chi-square analysis [21–23]. The scientific division in question was adjusted to the new student registration rules which divide the faculties at the State University of Malang into the natural and social sciences.

The pre-research activities carried out were establishing research problems, determining the scope and limitations of the research, compiling research designs or designs, developing research grids and instruments, validating research instruments, conducting instrument trials, refining instruments based on test results.

In the research process, research data were collected through Google Form by distributing the link to various WAg lecturers in each faculty and institution by asking for help from the Unit Leaders (faculty and institutions), openly to all lecturers and students at the State University of Malang. The data were then analyzed descriptively using percentages. Post-research activities are writing research results in the form of research articles for publication.

The variables in this study consisted of: the assessment form, the assignment technique, the assessment component, and the post-assessment which are summarized in 12 validated questions as seen in Table 2. The results of the validation show that all questions were valid and could be used in data collection activities.

Table 2. Item questions and instrument validation.

Indicator	Code	Questions	Validation Result
Assessment Form	Q1	The assessment form that is often carried out by lecturers	1.00
	Q2	If the "test" type of assessment is selected, what you will often use is (can choose more than one answer).	0.90
	Q3	For the written "test", the items difficulty level given to students is	0.90
	Q4	For the "non-test" assessment, an assessment form that is often done is	1.00
Assignment Technique	Q5	The most frequently performed individual assignment is	0.90
	Q6	The most frequently performed group assignment is	1.00
	Q7	Assessment implementation that is relevant to current topics in the student learning environment	0.90
	Q8	Assessment implementation that is currently being carried out helps students gain, change or develop skills, attitudes, ideals (ideals), appreciations (awards) and knowledge (knowledge)	0.80
Assessment Component	Q9	The assessment component that is carried out by the lecturer	0.90
Post-assessment	Q10	As a lecturer, do you provide feedback (discussion) for the assessment results that have been done	0.80
	Q11	If lecturers follow-up the assessment results, the form of follow-up is	0.90

In looking at the differences in authentic assessments carried out in the social sciences and natural sciences, several indicators are needed, namely: the form of assessment, assessment techniques, assessment components, and post-assessment so that a complete picture is obtained where the basic differences between authentic assessments in the social field are located with the natural sciences.

3. Result and Discussion

Authentic assessment can be underlined as a type of assessment which, when integrated into the learning framework, becomes an exciting instruction [11]. As a type of assessment, it is vital to know the assessment form and the assignment technique.

3.1. Assessment Form

The form of assessment that is most often carried out by lecturers in the natural field of science is 92.04% stating that the assessment most often carried out is a combination of test and non-test assessments. Meanwhile 8.64% only carry out the test. Meanwhile, in the field of social science, 84.57% stated that the most frequent assessments were a combination of test and non-test assessments. Meanwhile 8.64% only carried out the test and 6.79% only carried out the Nontest assessment (Figure 1). So, the form of assessment carried out by the majority of lecturers at this time is a mixture of tests and non-tests. But the visible difference between the form of assessment in the social sciences and the natural field lies in the use of the non-test form only just found in the social sciences.

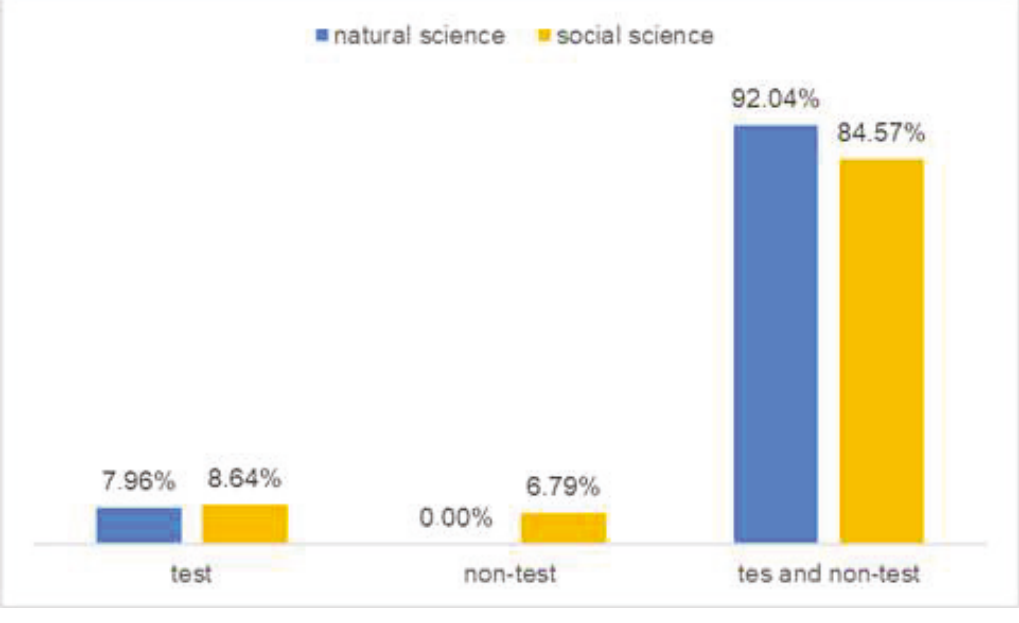

Figure 1. The form of assessment carried out by Lecturers in the field of natural sciences and social science.

Meanwhile, test form in the field of natural sciences, most often carried out according to natural science lecturers are written tests of 45.95%, performance tests (performance) 37.39%, oral tests (interviews) 16.67%. Meanwhile, for social science lecturers, according to the lecturers, the form of tests that are most often carried out are written tests by 21.22%, performance tests (performance) 54.69%, oral tests (interviews) 24.08% (Figure 2). The real difference seen in these two fields of science is that the use of written tests is mostly carried out by lecturers in the natural field, in contrast to lecturers in the social field who do more performance tests. This indicates a good implementation. Performance appraisal, portfolio assessment, and project appraisal are good authentic assessments to carry out. So, the implementation of the performance assessment has been carried out and needs to be supplemented with various other tests such as portfolio assessment and project assessment [24]. The data above were corroborated by the previous research of [25] that stated that to improve the effectiveness of mathematics learning, collaborative problem-

based learning model with authentic assessment model could be considered as one of the learning models in the classroom. The benefit of authentic assessment is that students learn to develop and use self-regulated skills to achieve high learning goals [26].

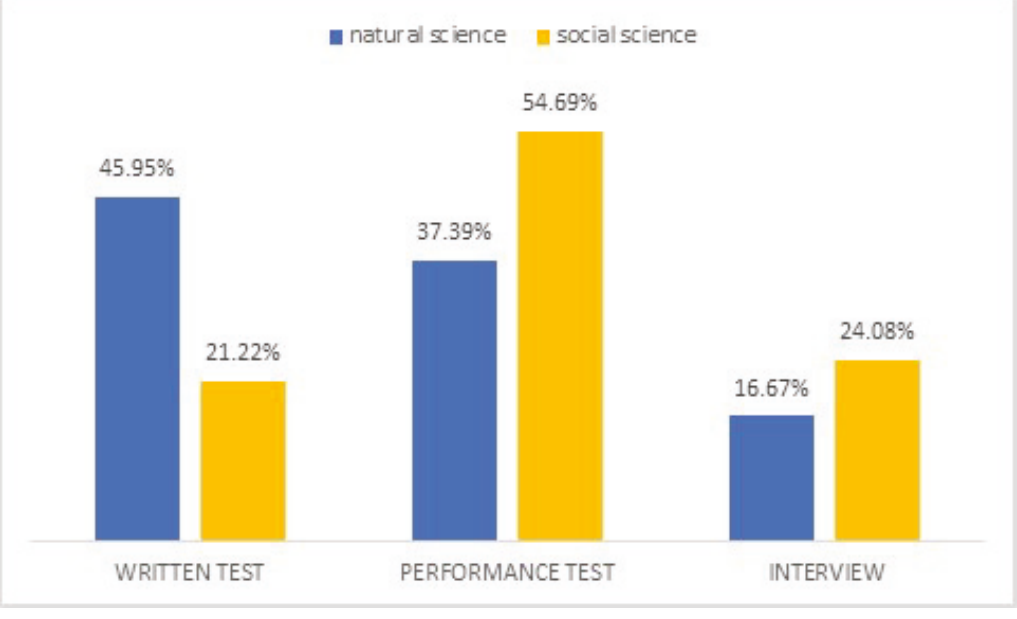

Figure 2. The form of tests conducted by lecturers in the field of natural sciences and social science.

Difficulty level: The tests carried out according to the natural lecturer are 24% higher order thinking, case or analysis questions 29%, application questions 26%, easy questions 11%, reviews 11%. While social lecturers are 21% of difficult questions (Higher Order Thinking), 36% case or analysis questions, 27% application questions, 7% easy questions, 9% reviews/reviews. Some lecturers have different views with high-complexity questions not just HOTS, no written tests, a mixture of difficult and easy questions, multiple choice to not using tests at all.

The forms of non-test assessment that are most often carried out according to natural lecturers are activeness in conducting discussions 24%, mastery of material in discussions 25%, independent tasks 28%, group assignments 23%. While social lecturers are active in conducting discussions 25%, mastery of material in discussions 24%, independent assignments 27%, group assignments 24%, and the rest the level of attendance in class, self-reflection, projects, individual assignments, peer assessment, ethics, discipline, attitude, etc.

Meanwhile, for the written test, the level of implementation of the questions given are 56.6% higher order thinking questions, 82.7% case or analysis questions, 66.1% application questions, questions are classified as easy 24.7%, reviews/reviews 0.3%. Higher order thinking skills (HOTS) are an important aspect in the education system and are very good if they can be applied in authentic assessment [27,28]. Unfortunately, only 56.6% of lecturers have used questions with the HOTS difficulty level, the rest are mostly applied and analyzed. Meanwhile, the types of non-test assessments that are often carried out by lecturers are activeness assessments and discussions. Changing the testing and assessment practices is a long process that requires enormous effort from lecturers. The courage to try something new and to actively change their behavior and beliefs could be done; they are well motivated if there is a clearly defined theoretical basis [29].

The form of the assessment carried out by the Lecturer of the State University of Malang turned out to have a difference between the natural sciences and the social sciences. The implementation of the form of test assessment is mostly carried out by the natural sciences field, while the social science field mostly conducts non-test assessments. As seen in Table 3, there is a significant difference between natural sciences and social sciences in the form of non-test assessments.

Table 3. The difference between natural sciences and social sciences in the form of non-test assessments.

Chi-Square Tests					
	Value	**Df**	**Asymptotic Significance (2-Sided)**	**Natural Sig. (2-Sided)**	**Natural Sig. (1-Sided)**
Pearson Chi-Square	7723 [a]	1	0.005		
Continuity Correction [b]	6129	1	0.013		
Likelihood Ratio	11.852	1	0.001		
Fisher's Natural Test				0.004	0.003
Linear-by-Linear Association	7697	1	0.006		
N of Valid Cases	296				

[a]. 1 cells (25.0%) have expected count less than 5. The minimum expected count is 4.58. [b]. Computed only for a 2 × 2 table.

In addition, the application of the high order thinking question has a difference. Natural science applies more questions to high order thinking than in the field of social science. As seen in Table 4.

Table 4. The difference application of high order thinking question.

Chi-Square Tests					
	Value	**Df**	**Asymptotic Significance (2-Sided)**	**Natural Sig. (2-Sided)**	**Natural Sig. (1-Sided)**
Pearson Chi-Square	10.215 [a]	1	0.001		
Continuity Correction [b]	9459	1	0.002		
Likelihood Ratio	10.378	1	0.001		
Fisher's Natural Test				0.002	0.001
Linear-by-Linear Association	10.181	1	0.001		
N of Valid Cases	296				

[a]. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 49.25. [b]. Computed only for a 2 × 2 table.

3.2. Assessment Technique

The assignments given by the lecturers consist of individual assignments and group assignments. The individual assignments that are most often carried out according to lecturers are making 48% summaries, making 55.4% papers, making 51.7% articles, doing 53% questions, 8.4% making books, and the others are reviews, product design, case analysis, making applied examples, essay projects, making presentations, making modules and learning media, case studies, answering questions, analysis/case studies, analyzing cases, making vlogs or video programs, practicing maps, preparing presentations, analyzing articles, making research proposals, making videos, compiling mind maps related to the material, and reporting reading results.

Table 5 shows that individual assignments in the form of writing articles have a real difference between social sciences and natural sciences. The field of social sciences gives more individual assignments in the form of making articles when compared to the field of natural sciences.

Table 5. The difference of individual assignments in the form of writing articles.

Chi-Square Tests					
	Value	**Df**	**Asymptotic Significance (2-Sided)**	**Natural Sig. (2-Sided)**	**Natural Sig. (1-Sided)**
Pearson Chi-Square	9808 [a]	1	0.002		
Continuity Correction [b]	9072	1	0.003		
Likelihood Ratio	9915	1	0.002		
Fisher's Natural Test				0.002	0.001
Linear-by-Linear Association	9775	1	0.002		
N of Valid Cases	296				

[a]. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 53.06. [b]. Computed only for a 2 × 2 table.

Meanwhile, the group assignments that are most often done according to natural science lecturers make 24% summaries, 76.6% papers, 41.4% articles, 27.8% project books, 11.9% reading reports, and so on. The assignments are group analysis, working on a project, review and presentation, working on a group project, social service projects and case studies, field measurements, discussions and performance observations/field visits, and project analysis/case studies.

From several assignment techniques that have been carried out by lecturers, the forms of assessment that can provide real experience to students according to lecturers are project assignments 78%, portfolio assignments 42%, case study assignments 72.6%. Other assignments that are considered to provide real experience to students are writing reflection textbooks, teaching skills tests, and observations.

In addition, in implementing the assessment, as many as 54.41% of respondents from natural science lecturers have carried out assignments that can assist students in obtaining, changing or developing skills, attitudes, ideals, appreciations, and knowledge. Meanwhile 41.18% rarely, 4.41% sometimes, and 0% never. Meanwhile, social science lecturers have carried out assignments that can assist students in obtaining, changing, or developing skills, attitudes, ideals (aspirations), appreciations (awards), and knowledge (knowledge) by 63.74%. Meanwhile 41.18% rarely, 4.41% sometimes, and 0% never (Figure 3). There is no difference in the application of the assessment in the two fields of science.

3.3. Assesment Component

A good assessment should be able to measure the ability of students as a whole. Authentic assessment is considered to be able to provide information that cognitive, affective, and psychomotor aspects can measure and assess students' abilities holistically. Assessment components that are relevant to current topics in the student learning environment in accordance with the lecturer's implementation are the assessment of cognitive aspects (knowledge) 91.9%, assessing affective aspects (attitudes) 86.6%, assessing psychomotor aspects (skills) 84.7%. Several lecturers gave different responses by mentioning that the assessment should be comprehensive, peer assessment, analytical ability, and social and spiritual competence should be aspects of the assessment. When viewed from the percentage of answers from respondents, all assessment components have an implementation rate of more than 80%, meaning that 80% of lecturers have carried out a holistic assessment related to relevance to the latest topics in the student learning environment.

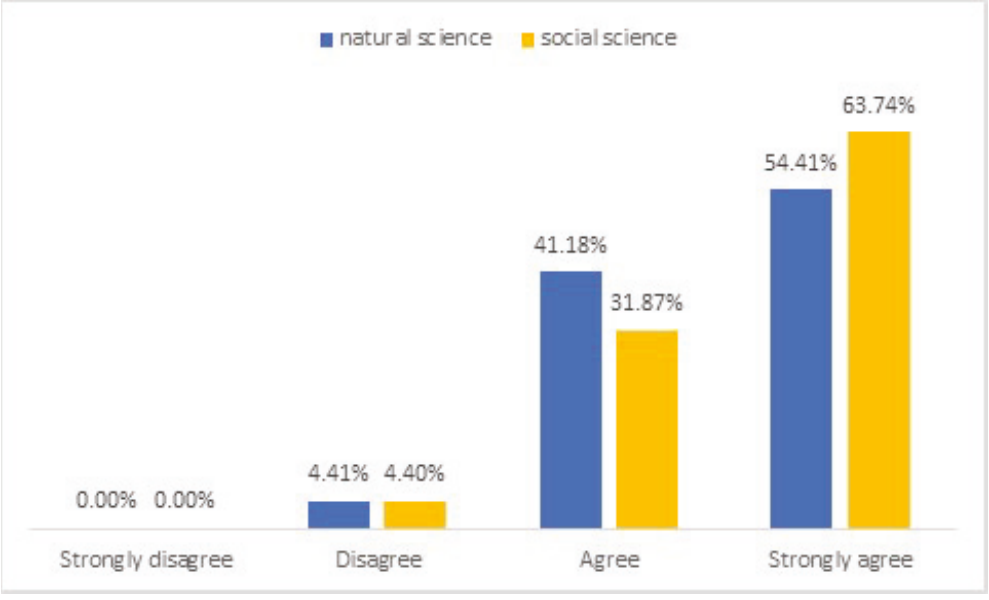

Figure 3. Implementation of assignments that can assist students in obtaining, changing, or developing skills, attitudes, ideals, appreciations, and knowledge.

3.4. Post-Assesment

The post-assessment process that has been carried out shows that as many as 45.90% of respondents lecturers of natural science provide feedback (discuss) on the results of the assessments that have been carried out, while 49.18% rarely, 4.92% sometimes, and 0% never. While the social sciences lecturer gave feedback (discussed) on the results of the assessment that had been done as much as 43.24%, while 50.00% rarely, 5.41% sometimes, and 1.34% never (Figure 4).

The follow-ups that were carried out were discussion by lecturers (one way) for 46.6%, joint discussion (FGD) for 72.6%, 33.4% of enrichment assignments, personal comment, in-class discussion, review, self-assessment, improvement of results using new assignments with higher level, discussion and questions and answers.

In the post-assessment, the majority of lecturers rarely provide feedback to students. If any, the form of follow-up that was performed would be collective discussion, one-way discussion by lecturers, and provision of enrichment. This situation should be a concern and evaluation material following the implementation scale as stipulated in the Regulation of the Minister of National Education No. 35 of 2010, a scale below 40% indicates that implementation is still not maximal [8]. The development of educational material that could be initiated through the students' opinions during the feedback process [29]. In relations to the three factors of problem-based learning, authentic assessment, and meaningful systems that are a combination of powerful tools in online learning, can provide education to students in effective digital learning [26].

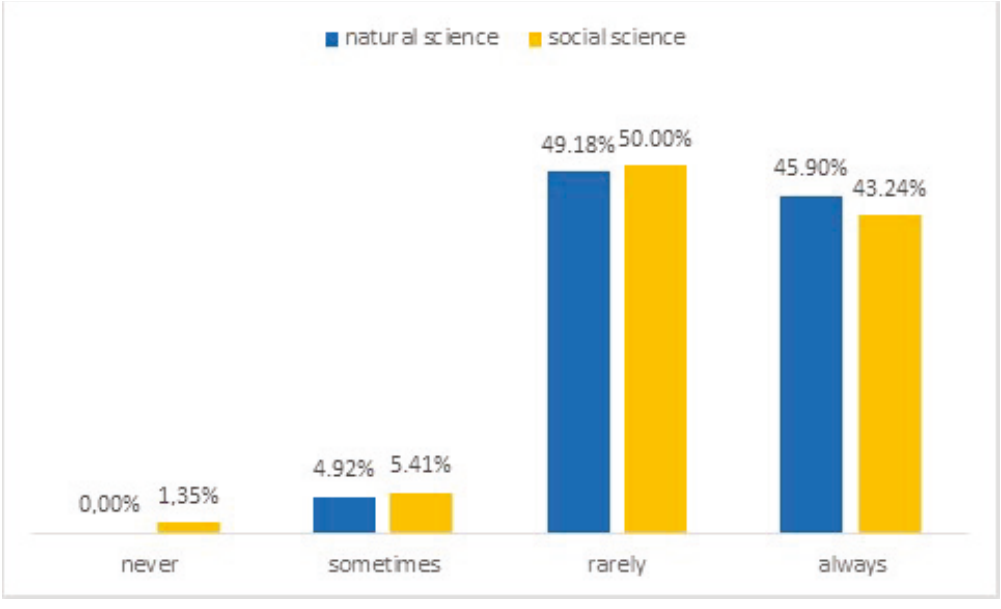

Figure 4. Lecturers provide feedback on the results of the assessments that have been carried out in the field of natural sciences in the field of social sciences.

Through the collected data, it can be seen that almost all aspects of authentic assessment between the natural and social sciences had no differences. In relation with the opinion that stated that social objects are not the same as natural objects; however, it should not leave us with the view that reality interacts with constructs in certain domains and specific sciences [14,30]. The only differences are in the form of assessment and individual assignment techniques, as shown in Table 6. Social science conducts more non-test assessments than natural science. This is understandable because sometimes, social objects are not definite and not measurable things so that a test assessment that contains specific parameters cannot be used in some areas of social science. Humanities and social sciences, in this respect, appear to be more complicated than chemistry and physics [31]. Actually, in both the social and natural sciences, there are seven steps for optimal use of portfolio assessment: (1) planning, (2) preparation for the study, (3) collecting evidence, (4) monitoring progress, (5) improving performance, (6) reflection, and (7) displaying works [32,33].

Measured tests are more widely used in the field of natural science using higher-order thinking skill questions. The use of HOTS questions in several previous studies had a significant effect on students' learning, competence, and science process skills [28,34]. Individual assignments in the form of writing articles are more widely used in social science concerning the learning outcome demands of multiple portfolios [35,36]. Both methods of assessment can be used, although not all problems in the natural and social sciences can be generalized. In practice, it would be better if an interdisciplinary approach is used [37,38], and a mixed assessment form could be used in this model.

Table 6. Types of natural and social science assessment.

Factor	Natural	Social
Assessment Form	• Written assessment (essay) • Performance assessment • Project assignment • Analysis of published research • Observation • Drawing • Multiple assessments	• Written assessment (essay) • Case studies • Class discussion • Conducting critical analysis of film shows • Fieldwork • Analysis of published research • Oral/verbal debate
Post-assessment	• Discussion of assessment results • Material enrichment • A focused discussion (FGD) • Discussion and questions and answers • Product revision • Feedback (written) • Critical analysis review	• Peer-assessment • Self-assessment • A focused discussion (FGD) • Feedback (written)

The contextual nature of critical thinking is a concept, its complex interactions with disciplinary knowledge approaches, and diverse and complex epistemologies, for assessors of critical thinking for pedagogy, curriculum, and assessment [39,40]. All scientific fields can ultimately use the implementation of authentic assessment, but it must be adapted to the epistemology and conform the learning outcomes in its application.

Figure 5 shows the proposed authentic assessment model for natural science. The abbreviation AF refers to authentic form, and PA refers to post-assessment. Based on the model, the assessments are performed sequentially. However, the instructor has an option if the next section is not required, then it can skip to the next one. It aims to make an assessment quickly and efficiently.

Figure 6 illustrates the proposed authentic assessment model for social science. The illustration represents that the assessment in social science tends to be more flexible. In other words, the instructor has free control to perform a first assessment that relates and is suitable to learning conditions. Moreover, 20 lecturers from each field study have validated these two models so that it can be implemented in the learning activities.

The results of this study are useful for academia because they can provide an overview of how lecturers in the social and natural science fields translate learning goals that might be achieved in teaching students using the PBK approach, what authentic tasks they choose to teach, how to use them to improve the quality of the process and subsequent learning outcomes, and what follow-up was provided regarding each of the authentic assessments used. The results of this study are useful for the professional world because they can provide a foundation and inspiration to conduct more in-depth research in their respective fields, according to existing interests and needs.

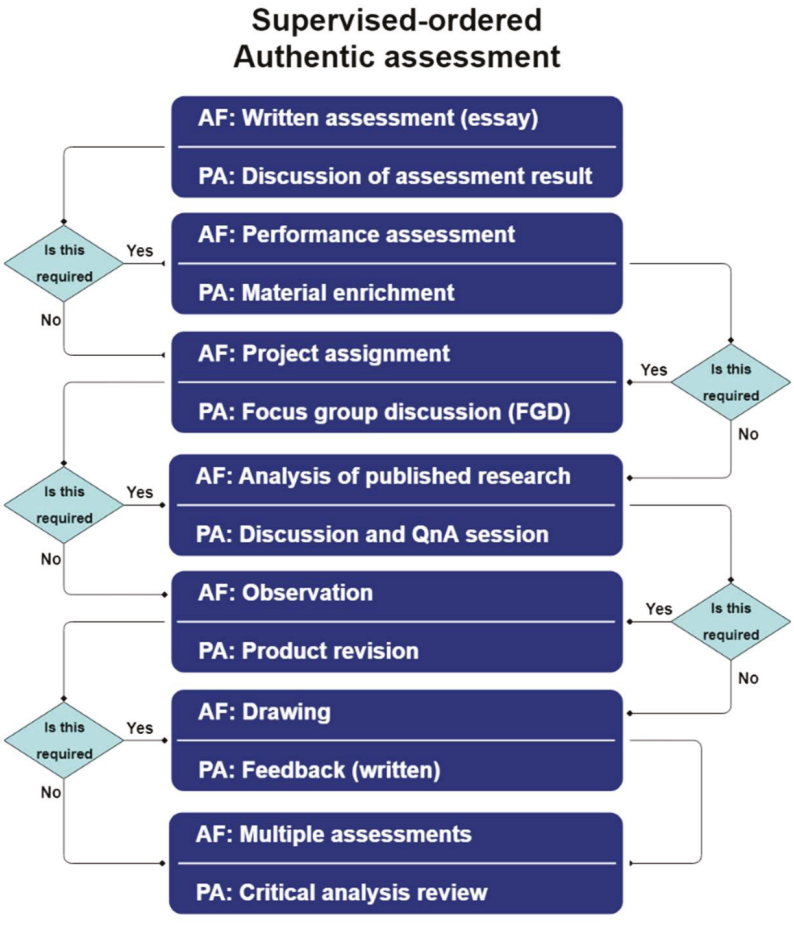

Figure 5. The proposed authentic assessment model for natural science.

The contribution of research results to the academic world related to authentic assessments include: (1) the process of seeking and finding information in the learning process, (2) the basis for making judgments, decisions, conclusions from an evaluation, (3) providing direct meaning in the educational process, for example real in biology learning that uses a process approach, while the contribution of research results to the professional world related to authentic assessments include: (1) authentic assessment known as performance assessment is a form of assessment that emphasizes professional performance related to the actual situation, (2) can know the attitude of professionals as expected, (3) allow to measure professional skills in a complex manner, and (4) enable the professionals being assessed to demonstrate their abilities in a real context.

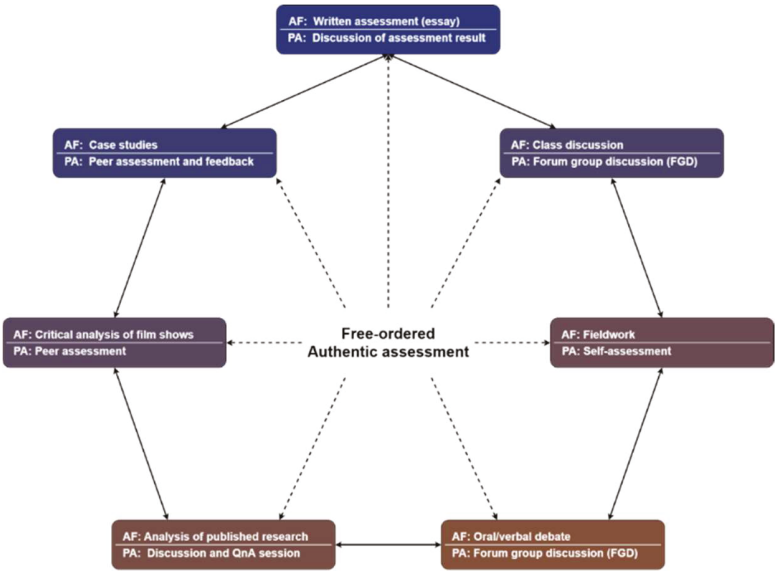

Figure 6. The proposed authentic assessment model for social science.

The comparison between authentic assessments used in the social and natural science fields is useful as a source of information regarding various kinds of authentic assessments that can be used by each field of science. However, the use of this authentic assessment must still be adjusted to the learning objectives and the type of task chosen in achieving the goals to be achieved.

The benefits of authentic assessment research for other researchers are (1) it can be used as an effort to develop assessments that aim to assess individual abilities through certain tasks, determine learning needs, help and encourage students and educators (teachers) to become better at determining learning strategies, institutional accountability, and improving the quality of learning, while further research is required in (1) finding obstacles in conducting comprehensive, holistic, and consistent assessments, and (2) finding solutions in overcoming difficulties in improvising/developing valid and reliable research instruments.

4. Conclusions

In conclusion, almost all aspects of authentic assessment between the natural and social sciences had no difference. The only differences were in the assessment form and individual assignment techniques that were performed. Social science conducted non-test assessment only higher than the natural science. Measured tests were primarily used in the natural science using higher-order thinking skills questions. Performance test was mostly conducted in social science. Further research is needed to identify following assessment model form particularly in higher education sample that could not give response as an accountability form of authentic assessment implementation that is more specific, holistic, comprehensive in a scientific cluster following the expert group.

Author Contributions: Data curation, E.S. and H.S.; investigation, N.A.M.J.; methodology, E.S.; project administration, E.S.; resources, S.N.R.; validation, A.P.W.; writing—original draft, E.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were waived for this study, due to reason that all students had agreed to participate in this study.

Informed Consent Statement: All of the participants stated their informed consent to participate in the study before doing the research.

Data Availability Statement: The data is available.

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

References

1. Pade, H.; Ingram, P.; Wright, A.J. *Teaching and Supervising Assessment beyond COVID-19*; Division 12 (Clinical Psychology) of the American Psychological Association: New York, NY, USA, 2020.
2. *International Baccalaureate Organization, Online Learning, Teaching and Education Continuity Planning for Schools*; International Baccalaureate Organization (UK) Ltd.: Cardiff, UK, 2020.
3. Olfos, R.; Zulantay, H. Reliability and validity of based course. *Educ. Technol. Soc.* **2007**, *10*, 156–173.
4. Kinay, I.; Bagceci, B. The Investigation of the Effects of Authentic Assessment Approach on Prospective Teachers' Problem-Solving Skills. *Int. Educ. Stud.* **2016**, *9*, 51. [CrossRef]
5. Susani, R.G. The Implementation of Authentic Assessment in Extensive Reading. *Int. J. Educ.* **2018**, *11*, 87. [CrossRef]
6. Abdullah, A.G.; Aryanti, T.; Setiawan, A.; Alias, M.B. *Regionalization and Harmonization in TVET*; Taylor & Francis: London, UK, 2017.
7. Wangid, M.N.; Mustadi, A.; Senen, A.; Herianingtyas, N.L.R. The evaluation of authentic assessment implementation of Curriculum 2013 in Elementary School. *J. Penelit. Dan Eval. Pendidik.* **2017**, *21*, 104–115. [CrossRef]
8. Suarimbawa, K.A.; Marhaeni, A.A.I.N.; Suprianti, G.A.P. An Analysis of Authentic Assessment Implementation Based on Curriculum 2013 in SMP Negeri 4 Singaraja. *J. Educ. Res. Eval.* **2017**, *1*, 38–45. [CrossRef]
9. Idham, F.I. The Use of Authentic Assessment in English Writing Skill to The Eleventh Grade Students. *e-J. Engl. Lang. Teach. Soc. (ELTS)* **2015**, *3*.
10. Rizavega, I.H. Authentic Assessment Based on Curriculum 2013 Carried by EFL Teacher. *J. Profesi Kegur.* **2018**, *4*.
11. Ardianti, T.M.; Mauludin, L.A. Students' Responses on the Application of Authentic Assessment in EFL Reading Class. *J. Engl. Lang. Lit. Teach.* **2017**, *1*.
12. Kearney, S.; Perkins, T.; Clark, S.K. Using self- and peer-assessments for summative purposes: Analyzing the relative validity of the AASL (Authentic Assessment for Sustainable Learning) model. *Assess. Eval. High. Educ.* **2016**, *41*, 840–853. [CrossRef]
13. Ingthorsson, R.D. The Natural Vs. The Human Sciences: Myth, Methodology and Ontology. *Discus. Filosóficas* **2013**, *14*.
14. Persson, J.; Hornborg, A.; Olsson, L.; Thorén, H. Toward an alternative dialogue between the social and natural sciences. *Ecol. Soc.* **2018**, *23*. [CrossRef]
15. Milanez, B. Dialogues between social and natural sciences: Contribution to the debate on socio-environmental conflicts. *Acad. Bras. Cienc.* **2015**, *87*, 2335–2348. [CrossRef]
16. Kagan, J. *The Tree Cultures: Natural Sciences, Social Sciences, and the Humanities in the 21st Century*; Cambridge University Press: New York, NY, USA, 2009.
17. Borg, W.R.; Gall, J.P. *Educational Research Introduction*, 6th ed.; Longman Publishers: White Plains, NY, USA, 1996.
18. Creswell, J.W. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 4th ed.; SAGE Publications: New Delhi, India, 2014.
19. Alam, M. Selection of the Samples with Probability Proportional to Size. *Sci. J. Appl. Math. Stat.* **2015**, *3*, 230. [CrossRef]
20. Wright, T. Probability proportional to size (πps) sampling using ranks. *Commun. Stat. Theory Methods* **1990**, *19*, 347–362. [CrossRef]
21. Pandis, N. The chi-square test. *Am. J. Orthod. Dentofacial. Orthop.* **2016**, *150*, 898–899. [CrossRef]
22. Ottenbacher, K.J. The chi-square test: Its use in rehabilitation research. *Arch. Phys. Med. Rehabil.* **1995**, *76*, 678–681. [CrossRef]
23. Franke, T.M.; Ho, T.; Christie, C.A. The Chi-Square Test. *Am. J. Eval.* **2011**, *33*, 448–458. [CrossRef]
24. Fitriani, F. Implementing Authentic Assessment of Curriculum 2013: Teacher's Problems and Solusions. *Getsempena Engl. Educ. J. (GEEJ)* **2017**, *4*, 217656.
25. Darma, I.K.; Candiasa, I.M.; Sadia, I.W.; Dantes, N. The Effect of Problem Based Learning Model and Authentic Assessment on Mathematical Problem Solving Ability by Using Numeric Ability as the Covariable. *J. Phys. Conf. Ser.* **2018**, *1040*, 012035. [CrossRef]
26. Litchfield, C.B.; John, V.D. Authentic Assessment of Knowledge, Skills, and Attitudes. In *New Directions for Teaching and Learning*; Wiley: Hoboken, NJ, USA, 2015.
27. Mohamed, R.; Lebar, O. Authentic Assessment in Assessing Higher Order Thinking Skills. *Int. J. Acad. Res. Bus. Soc. Sci.* **2017**, *7*, 466–476.
28. Haryati, B.M.; Gultom, T. The Effect of Learning Model on Higher Order Thinking and Student Science Process Skills in Ecology. *Int. J. Humanit. Soc. Sci. Educ.* **2017**, *4*, 150–155.
29. Hus, V.; Matjašič, J. Evaluation and Assessment in Early Social Science. *Univers. J. Educ. Res.* **2017**, *5*, 664–670. [CrossRef]
30. Montuschi, E. ShouldWe Still Compare the Social Sciences to the Natural Sciences? *Sociologica* **2008**.
31. Jaffe, K. Social and natural sciences differ in their research strategies, adapted to work for different knowledge landscapes. *PLoS ONE* **2014**, *9*, e0113901. [CrossRef]

32. Srikaew, D.; Tangdhanakanond, K.; Kanjanawasee, S. Development of an English Speaking Skill Assessment Model for Grade 6 Students by Using Portfolio. *Soc. Behav. Sci.* **2015**, *191*, 764–768. [CrossRef]
33. Setiawan, D. Authentic Assessment Model In Social Studies Learning To Improve The Social Skills. *Turk. Online J. Des. Art Commun.* **2017**, *7*, 1393–1403.
34. Mahzum, E. Making Analysis of Higher Order Thinking-Based Objective Test at State Junior High School 5 Banda Aceh. *J. Phys. Conf. Ser.* **2018**, *1116*, 032018. [CrossRef]
35. Alleman, J. Authentic Assessment in Social Studies. *HSSE Online* **2012**, *1*, 7–26.
36. Martin, A.; Arrambide, M.; Holt, C. The Impact of Flipped Instruction on Middle School Mathematics Achievement. *J. Educ. Hum. Dev.* **2016**, *5*. [CrossRef]
37. Lowe, P.; Phillipson, J.; Wilkinson, K. Why social scientists should engage with natural scientists. *Contemp. Soc. Sci.* **2013**, *8*, 207–222. [CrossRef]
38. Barthel, R.; Seidl, R. Interdisciplinary Collaboration between Natural and Social Sciences—Status and Trends Exemplified in Groundwater Research. *PLoS ONE* **2017**, *12*, e0170754. [CrossRef] [PubMed]
39. Nicholas, M.C.; Labig, C.E., Jr. Faculty Approaches to Assessing Critical Thinking in the Humanities and the Natural and Social Sciences Implications for General Education. *J. Gen. Educ. A Curric. Commons Humanit. Sci.* **2013**, *62*. [CrossRef]
40. Morris, R.V. Drama and Authentic Assessment in a Social Studies Classroom. *Soc. Stud.* **2001**, *92*, 41–44. [CrossRef]

education sciences

MDPI

Article

Impact of Governance Factors over Lecturers' Scientific Research Output: An Empirical Evidence

Hue Thi Truong [1], Hung Manh Le [2,*], Duc Anh Do [3], Duc Anh Le [4], Huyen Thi Nguyen [5] and Thanh Kim Nguyen [6]

[1] VNU School of Interdisciplinary Studies, Vietnam National University, Hanoi 11310, Vietnam; tthue@vnu.edu.vn
[2] Board of Management, Trade Union University, Hanoi 11514, Vietnam
[3] School of Trade and International Economics, National Economics University, Hanoi 11616, Vietnam; ducda@neu.edu.vn
[4] Faculty of Mathematical Economics, National Economics University, Hanoi 11616, Vietnam; ducla@neu.edu.vn
[5] HR Department, Trade Union University, Hanoi 11514, Vietnam; huyennguyenthi@dhcd.edu.vn
[6] Department of Research Management, Trade Union University, Hanoi 11514, Vietnam; thanhnk@dhcd.edu.vn
* Correspondence: hunglm@dhcd.edu.vn

Abstract: This study aims to determine the governance factors that influence the scientific research output of lecturers through the application of PLS-SEM, in conjunction with ANOVA and *t*-test. Based on a survey of 398 lecturers in twelve higher education institutions (HEIs) in Vietnam, the psychometric properties of the scales measuring the considered dimensions of scientific research outputs were initially examined through the Exploratory Factor Analysis (EFA) procedure, prior to being input into the PLS-SEM model. The SEM model comprised six constructs for the scientific research outputs: scientific research objectives of HEIs, leadership, decentralization, policies for lecturers, support for scientific research activities, and resources for scientific research. The results reveal that resources for scientific research have the most impact on lecturers' scientific research output, followed by policies for lecturers, support for scientific research activities, scientific research objectives of HEIs, and finally, leadership.

Keywords: lecturer; scientific research; governance; management; higher education institutions

Citation: Truong, H.T.; Le, H.M.; Do, D.A.; Le, D.A.; Nguyen, H.T.; Nguyen, T.K. Impact of Governance Factors over Lecturers' Scientific Research Output: An Empirical Evidence. *Educ. Sci.* **2021**, *11*, 553. https://doi.org/10.3390/educsci11090553

Academic Editors: Sandra Raquel Gonçalves Fernandes, Marta Abelha and Ana Teresa Ferreira-Oliveira

Received: 8 July 2021
Accepted: 10 September 2021
Published: 17 September 2021

1. Introduction

Scientific research achievement is among the key criteria used to evaluate performance and university rankings on a global scale [1]. The future of universities is related to lecturers' achievements and progress in scientific activity [2]. Scientific work is among the most important activities of the faculty of any higher education institution [3]. Lecturers' scientific research output has received increasing attention from many countries, because many governments, including those in developing countries, have found that this is crucial in today's knowledge-based economy [4].

Due to the high importance placed on scientific research activities, several scholars have investigated the factors affecting lecturers' scientific research results [5–11]. Some studies considered this issue in the context of Vietnam [12–16]. However, studies focusing on governance factors that affect scientific research outputs of university lecturers are lacking. To date, the majority of studies in the literature have approached institutional factors [17–19] and individual factors [7,10], or a mix of both [7,9,11,20]. The main research methods applied are regression analysis [1,7,21], analysis of variance (ANOVA) [22,23], structural equation modeling (SEM) [10,19], analytic hierarchy process (AHP) [14,18], interviews [15,16], and Delphi [13]. Therefore, the impact of governance factors (not institutional factors or individual factors) on the scientific research results of lecturers remains under-investigated.

SEM, one of the most commonly applied techniques to analyze cause-and-effect relationship models, has also been widely applied in the context of economics and management studies [10]. There are two popular approaches to estimate linear structural models: Covariance-based (CB-SEM) and Partial least squares (PLS) SEM. In contrast to CB-SEM, PLS-SEM is commonly used to develop theories in exploratory research by focusing on explaining the variance in the dependent variables [24]. Meanwhile, according to [25], there is no theoretical basis to date that can be explicitly used to explain the cause-and-effect relationship, nor is there any common theoretical framework in previous studies on the factors influencing the research output of lecturers. Some studies have applied CB-SEM [10,19], but have not considered PLS-SEM.

This study aims to determine the impact of governance factors on the scientific research output of lecturers. In order to meet the research goals, the PLS-SEM method was used. Furthermore, the ANOVA and *t*-test approaches were also applied to investigate the differences in the scientific research output among lecturer groups stratified by demographics.

2. Literature Review

2.1. Governance Factors

To date, several studies have been conducted on the factors that impact the research output of the faculties within a higher education institution. Some studies focused on collaboration and networking factors. For instance, Ref. [5] showed that organizational factors such as job satisfaction and international collaboration had a positive effect on scientific research productivity among academic staff. The works by [26,27] concluded that the key factors that contribute to the success of publishing research in highly ranked American journals were collaboration with established overseas scholars, English language, and research ideas. The study by [14] demonstrated that the most important factors influencing faculties' research performance were collaboration with domestic and international peers, and receiving support from research assistants and supervisors. The study by [13] revealed that "networking-related factors" were a key determinant in the success of international publishing. However, these studies have not yet distinguished the collaboration factors from governance policies by higher education institutions (HEIs), or lecturers themselves.

Some studies examined the impact of university policies, work environment, and resource-related factors. For instance, Ref. [5] investigated IT funding and university policies; and [13] concluded that policy-related factors have a positive impact on scientific research productivity. The study by [7] examined institutional characteristics, including the number of undergraduate students enrolled, percentage of Ph.D. students enrolled, and funding allocated for research function, and revealed that they are significantly associated with faculty research productivity. The work by [12] revealed that work environment affected the adjusted research productivity of social scientists. However, these studies provide limited understanding on the specific policies (e.g., policies for lecturers) and resources, as well as the specific factors of work environment, that HEIs create to support research activities by lecturers.

Other studies investigated individual factors, particularly motivation factors. For instance, Ref. [7] showed that the research productivity of faculty varied by gender, institution, terminal degree, rank, discipline, and work experience. Notably, Ref. [11] examined the relationship between research productivity and the intrinsic and extrinsic motivators associated with conducting research. According to this study, there was a significant difference in research productivity between faculties associated with doctoral vs. non-doctoral degree granting programs. Receiving or having tenure was the most important reward, while securing a possible administrative position was the least important. There was a significant difference in the importance of these rewards between tenured-untenured and between male-female faculty members. A strong link was found between research productivity and the attainment of a tenure position and of a promotion, but the link between publications and salary increase was not strong. However, Ref. [7,11] have

several prominent limitations. It is difficult to draw a distinction between these two types of motivation. One factor that may be an extrinsic source of motivation for one person, can be considered an intrinsic source of motivation by another. Besides, Ref. [7,11] have not shown whether motivation originates from the governance policies of HEIs to enhance research achievement.

Some authors proposed a set of characteristics for a productive research organization. The work by [9] indicated that creative accomplishments are associated with small group size, organizational contexts with sufficient access to a complementary variety of technical skills, stable research sponsorship, timely access to extramural skills and resources, and facilitating leadership. However, these factors are features for organizations with great creative achievements in general, not necessarily every university where lecturers work. Ref. [20] recommended a synthetic model of individual, institutional, and leadership characteristics predicting individual and department research productivity, based on the combined results from multiple regressions. However, this model was examined at the individual and departmental level (i.e., not the university level), and combined governance factors with several individual or environmental factors.

To date, three approaches were applied to investigate the influential factors on the research performance of lecturers, namely, individual, institutional, and a mix of both. According to [28], institutional factors play an important role in developing a research culture to promote greater interaction and transfer of knowledge to society. Some factors that are related to university governance have been found in prior studies that considered institutional factors. However, it seems that there has been very little discussion about the impact of governance factors in particular (i.e., not institutional factors, in general). Therefore, it is worthwhile to empirically investigate the governance factors that influence the research output of lecturers, as well as the level of impact of those factors.

In Vietnam, only a few studies have attempted to address the factors associated with the research output of lecturers. The studies of [12–16] have been undertaken to investigate the factors influencing internationally-indexed publishing. Among these studies, Ref. [12] is one of the most notable, indicating that collaboration with international researchers resulted in higher productivity among social scholars in Vietnam. However, these studies were conducted with secondary data, and did not examine the governance factors as perceived by researchers in HEIs. The studies by [13–15] overcame the limitations of [12] by utilizing in-depth interviews, Delphi, and AHP respectively, however they only focused on internationally-indexed publishing by all types of academic organizations and academics staff, and not by HEIs and lecturers in particular.

It should also be noted that according to the literature review, the most commonly employed research methods for this topic were regression analysis, CB-SEM, ANOVA variance analysis, Delphi, interviews, case analyses, descriptive statistics, and AHP, whilst the PLS-SEM method has not been adopted. The methods of some prominent previous studies are listed in Table 1.

Table 1. Research methods on governance factors affecting scientific research output of lecturers.

Author	Methods
[29]	CB-SEM (LISREL)
[10]	CB-SEM
[19]	Delphi, CB-SEM
[11]	Regression
[20]	Regression, *t*-test
[1]	Regression
[22]	Regression, ANOVA
[21]	Regression
[30]	Regression, EFA
[31]	Regression

Table 1. *Cont.*

Author	Methods
[32]	Regression, *t*-test
[7]	Regression, ANOVA
[33]	Regression, Bootstrap
[34]	Regression
[12]	Ordinary least squares, secondary data
[14]	AHP
[18]	AHP, ANOVA, and *t*-test
[23]	ANOVA, semi-structured interviews
[35]	*t*-test
[36]	Mann–Whitney U test
[37]	Mann–Whitney U test
[13]	Delphi
[15]	Semi-formal, in-depth interviews
[16]	In-depth interview
[35]	In-depth interviews
[17]	Descriptive statistics, interviews
[9]	Design and analysis of typical situations
[38]	Document analysis, model recommendations

2.2. Scientific Research Outputs

Regarding the evaluation and measurement of scientific research output, Ref. [39] indicated that research evaluation studies employ various instruments and indicators, depending on the particular aims of the study.

In terms of the quantity-based approach, researchers generally agree that scientific research output is measured by the total number of publications by the institution. According to [4], research output relates to creative ideas that, after being studied, are published in magazines, newspapers, patent applications, or academic journals. In addition, some researchers pointed out other indicators such as the amount of research funding [8], membership in a scientific association [40], and the financial budget allocated to research [4].

In terms of quality and the influence-based approach, Ref. [41] emphasized that scientific output must be presented by internationally-indexed publishing. The authors of [42] utilized H index to measure the research performance of scientists. The works by [7,43] argued that H index is a reliable indicator that is recognized worldwide to evaluate the scientific research performance of scientists.

The work by [38] defined the holistic approach in quantity, quality and influence. Table 2 shows a framework that identifies research output measurements and is recommended by [38].

So far, there has been a controversy on the influence-based approach. The authors of [44] mentioned that not all publications are indexed in research databases for citation, and it is difficult to assess the true value of a publication by H index. There are differences between industries and fields; the number of years of publication (first published is more likely to be cited) and the age of scientists significantly affects the citation index. Therefore, the H index, in some cases, is not accurate and fair when comparing the research productivity of scientists and organizations. According to [25], the approach based on the number of publications has been more widely used than that based on the qualitative measures to evaluate the research productivity of academics at most universities around the world. Hence, in this study, the quantity-based approach was adopted.

Table 2. Indicators of measuring scientific research output.

No.	Indicators
1	Number of articles published in refereed or professional journals
2	The impact factor of the journals
3	Number of published books
4	Number of edited books
5	Number of published chapters in refereed books
6	Number of edited chapters in refereed books
7	Number of citations
8	Citations as a measure of impact
9	Patent registration
10	Received research grant
11	Participated in research projects
12	Number of honors and awards
13	Number of papers presented in meetings or conference
14	Number of invitations to present papers
15	Number of supervised dissertations
16	Supervised one or more honors/master students
17	Supervised one or more Ph.D. students
18	Served as an editor of an academic journal
19	Positions held in a professional association
20	Maintained professional contact with colleagues overseas

2.3. Hypotheses

Among the institutional and individual factors in prior studies there are some typical governance factors to be found, including the scientific research objectives of HEIs, leadership, policies for lecturers, support for scientific research activities, and resources for scientific research. Besides these, some studies also mentioned "decentralization" or related content among the institutional factors.

Regarding the scientific research objectives of HEIs, Ref. [6,38] showed that having a research agenda defined as the combination of strategic problem-solving frameworks to achieve research goals is the best technique to enhance research productivity and easily monitor the measurement of academic progress. Other studies also argued that in order to achieve worthy scientific research achievements a university's scientific research objectives need to be clear, feasible, and widely shared [22]. The university's statement of its core mission has a strong effect on its research output [43]. The findings above were developed and supported by [20], who suggested that universities should set clear, visible, shared goals to enhance their overall research performance. Therefore, the first hypothesis was proposed:

Hypothesis 1 (H1). *There is a positive relationship between the scientific research objectives of HEIs and the scientific research output of lecturers.*

To the best of our knowledge, decentralization has not been considered a scale in any quantitative research model. However, some studies have found a positive correlation between decentralized organizations, the autonomy of HEIs, lecturers, and lecturers' scientific research output. The work by [20] showed that an effective research apparatus feature is the "assertive-participative management" mechanism characterized by management decisions with wide participation by stakeholders. The work by [38] also mentioned that this mechanism, which is characterized by decentralization and empowerment, has a positive impact on scientific research output. In addition, Ref. [20,22,30] showed that the autonomy of the organization and lecturers is among the most contributing factors to scientific research achievements. Therefore, this research generalizes and supplements this new scale in the proposed model. Hence, the second hypothesis was proposed:

Hypothesis 2 (H2). *There is a positive relationship between decentralization and the scientific research output of lecturers.*

Concerning leadership, Ref. [30] defined that the leadership factor group has the greatest positive impact on scientific research output among several factors. The authors of [9] argued that both organizational (i.e., university) leaders and group leaders are important to the effectiveness of creation. The work by [18] also demonstrated that the support of leaders for scientific research contributes to improving scientific research output by creating a departmental research atmosphere. Notably, Ref. [20] investigated the leadership characteristics of productive research organizations, namely, that they are highly-regarded, have able scholars, are research-oriented, emphasize an assertive-participative style, fill critical roles in areas such as management and fundraising, and keep goals visible. Basing on these findings, the third hypothesis was proposed:

Hypothesis 3 (H3). *There is a positive relationship between leadership and the scientific research output of lecturers.*

Concerning support for scientific research activities, several authors emphasized the role of creating collaborative opportunities and environments for academic staff. The studies by [13,14,28,45] identified that collaboration was a key determinant of international publishing. Therefore, it was recommended that HEIs should establish research groups and collaboration with international and domestic peers to enhance their research productivity. The work by [29] and [35] showed that the university's explicit support of scientific research activities has a positive effect on lecturers' scientific research results. The authors of [19] found that ongoing supportive activities create a culture that values research, thereby influencing the conscientiousness and scientific research results of scientists. Therefore, the fourth hypothesis was proposed:

Hypothesis 4 (H4). *There is a positive relationship between support for scientific research activities and the scientific research output of lecturers.*

In relation to policies for lecturers, Ref. [13,16] argued that time available for research purposes was a key factor for international scientific publishing. The work by [16] showed that university policies and job satisfaction had a positive effect on scientific research productivity among academic staff. The study by [11] revealed a positive relationship between research productivity and having tenure, rewards, promotion, and salary. The study by [35] revealed that the policy of recruiting lecturers and signing contracts based on lecturers' scientific research capacity can enhance their overall research performance. Similarly, income policy was considered by [17] and [23], while promotion policy was considered by [37] and [38]. Notably, Ref. [20] reported several characteristics of productive research organizations. The first is recruitment and selection, whereby significant effort is expended to recruit and hire members who have the training, goals, commitment, and socialization that match the institution. The second characteristic is rewarding, whereby research is rewarded equitably and in accordance with predefined benchmarks of achievement. Hence, the fifth hypothesis was proposed:

Hypothesis 5 (H5). *There is a positive relationship between policies for lecturers and the scientific research output of lecturers.*

With reference to resources for scientific research, Ref. [38] presented resources-related factors, including technology and equipment, libraries, and research funding. The work by [13] revealed that accessibility of international scientific documents, experimental devices or tools, software used for research purposes, and accessibility of research funding sources are the determinants in the performance of international publishing in Vietnam. The study by [18] demonstrated that research funding is considered to be the most important

factor for research results, followed by human resources, journal and library resources, and facilities. Similarly, several other studies concluded that one of the institutional factors for improving scientific research output is financial support [5,7,20]. For this reason, the final hypothesis was proposed:

Hypothesis 6 (H6). *There is a positive relationship between resources for scientific research and the scientific research output of lecturers.*

According to the literature review and hypotheses proposed, a research model was developed with the following factors: scientific research objectives of HEIs, leadership, decentralization, policies for lecturers, support for scientific research activities, and resources for scientific research. Additionally, to determine the difference between population groups in research output, the research model included the following factors (i.e., control variables): gender, age, degree and academic title, experience, position (managers or non-managerial lecturers), place of graduation (abroad or domestically), and scientific area. Figure 1 depicts the proposed research model.

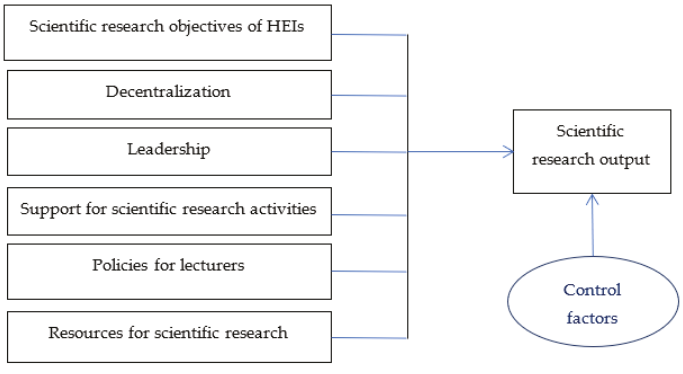

Figure 1. Proposed research model.

3. Methods

3.1. Research Design

A questionnaire was constructed based on the literature review and was adjusted based on the interview results with five experts. The final questionnaire consisted of 49 items, not including the participants' demographics section, which included 41 statements of governance factors and eight items of lecturers' research output. Participants were requested to rate a 5-point Likert scale, whereby 1 = totally disagree, 2 = somewhat disagree, 3 = neither agree nor disagree (neutral), 4 = somewhat agree, and 5 = totally agree.

The data were collected from April to September 2020. Firstly, a pilot study including 82 observations was implemented before distributing the online survey (via Microsoft Forms) and the offline survey. Next, the questionnaire was revised based on the Cronbach's Alpha value from the initial pilot test. In the formal survey stage, a non-probability sampling method was implemented, and a total of 413 responses were collected from 12 HEIs in the North of Vietnam. The 15 bias observations were eliminated. Finally, there were 398 observations valid for further analysis, of which 313 observations (78.6%) were from the online survey via Microsoft Forms (in the period of social distancing due to the COVID-19 epidemic in Hanoi, in April 2020), while 85 observations (21,4%) were from the offline survey (after the period of social distancing; from May to September 2020). Table 3 shows the descriptive statistics of the participants' demographics.

Table 3. Descriptive statistics of participants' demographics.

Variables	N	%
Gender		
Male	169	42.5%
Female	229	57.5%
Age		
22–30	50	12.6%
31–40	194	48.7%
41–50	121	30.4%
>50	33	8.3%
Academic title, degree		
Professor	4	1.0%
Associate Professor	59	14.8%
Doctor	210	52.8%
Master	125	31.4%
Abroad graduation or not		
Abroad graduation	190	47.7%
Domestic graduation	208	52.3%
Experience as a lecturer		
<1 year	30	7.5%
1–3 years	42	10.6%
>3–6 years	46	11.6%
>6 years	280	70.4%
Scientific research area		
Natural Sciences	179	45.0%
Social Sciences	219	55.0%
Position		
Manager	146	36.7%
Non-managerial employee	252	63.3%

According to Table 3, the ratio of males to females is 42.5% and 57.5% respectively. In terms of age, the majority of lecturers (48.7%) fall into the 31–40 age group, 30.4% fall into the 41–50 group, 12.6% fall into the 22–30 group, and only 8.3% of the lecturers are over 50 years old. Qualification in the survey sample is high; up to 52.8%, 31.4%, 14.8%, and 1% of lecturers have the academic title of Doctor, Master's degree-holder, Associate Professor, and Professor, respectively. The number of lecturers who have more than six years of working experience accounts for quite a high proportion (70.4%), while 11.6% and 10.6% of lecturers have 3–6 and 1–3 years of working experience respectively. Only 7.5% of the lecturers have less than one year of experience. Regarding scientific research areas, social science takes the highest proportion at 55%, while natural science accounts for 45%. The ratio of abroad to domestic graduation is similar, with 47.7% and 52.3% respectively. The number of lecturers without managerial positions accounts for a higher percentage than those with managerial positions, with 63.3% and 36.7% respectively.

3.2. Data Analysis Techniques

Firstly, the *t*-test and ANOVA approaches were adopted to determine the differences in scientific research output between lecturer groups, according to their demographic characteristics. Next, Partial Least Squares Structural Equation Modeling (PLS-SEM) was employed to achieve the research objectives. This approach has several advantages. PLS-SEM was recommended in the initial period of theory development to access and verify the exploratory research models [46]. It also has several benefits for cause-and-effect analysis in behavior studies [47]. PLS-SEM has been the best alternative to CB-SEM in cases wherein there is little background theory available [30,48] even though PLS-SEM is not as effective as CB-SEM in model fit evaluation. Furthermore, there have not been any common

theoretical frameworks in previous studies on the factors influencing the research output of lecturers [25]. In addition, a new scale (i.e., decentralization) was added to the research model proposed in this study. For these reasons, it is appropriate to utilize PLS-SEM.

PLS-SEM is based on two main steps, namely, measurement model assessment and structural model assessment [46]. Within this study, the SmartPLS 3.3.3 application of PLS-SEM was used to assess the measurement model, the convergent, discriminant validity, and composite reliability. Finally, the bootstrapping technique analyzed the t-statistics for the path coefficients to assess the importance of the hypothesized connections.

The factors were coded as follows: OBJ = scientific research objectives of HEIs, DEC = decentralization, LEA = leadership, POL = policies for lecturers, SUP = support for scientific research activities, RES = resources for scientific research, and OUT = research outputs (Appendix A).

4. Results

4.1. t*-test and ANOVA*

Table 4 presents a summary of *t*-test and ANOVA results to demonstrate the difference between participants' research output among demographic variables.

Table 4. Summary of and *t*-test and ANOVA results.

Lecturer Group	Sig. of Levene's Test	Sig. of *t*-test/Welch/F Test		N	Mean
Gender	0.477	Sig. of *t*-test	0.000		
Male				169	3.971
Female				229	3.731
Age	0.001	Sig. of Welch Test	0.000		
22–30				50	3.557
31–40				194	3.876
41–50				121	3.935
>50				33	3.625
Academic title, degree	0.087	Sig. of F Test	0.000		
Professor				4	4.437
Associate Professor				59	3.924
Doctor				210	3.868
Master				125	3.712
Abroad graduation or not	0.700	Sig. of *t*-test	0.018		
Abroad graduation				190	3.884
Domestic graduation				208	3.787
Experience as a lecturer	0.000	Sig. of Welch Test	0.003		
<1 year				30	3.500
1–3 years				42	3.732
>3–6 years				46	3.856
>6 years				280	3.880
Scientific research area	0.002	Sig. of *t*-test	0.000		
Natural Sciences				179	3.997
Social Sciences				219	3.699
Position	0.111	Sig. of *t*-test	0.000		
Manager				146	3.684
Non-managerial employee				252	3.919

The difference in the scientific research results of lecturer groups is summarized hereafter. Regarding gender, there is a difference in scientific research output between the sexes, shown by the sig. of the *t*-test being <0.05. More specifically, the mean value shows that male lecturers have higher scientific research output than female lecturers.

Concerning age, there is a difference in scientific research output between age groups, shown by the sig. of the Welch test being <0.05. More specifically, the mean value of age

groups shows that scientific research output decreases in the order of the following age groups: 41–50, 31–40, >50, and 22–30.

In terms of academic title and degree, there is a difference in scientific research output among lecturer groups with different qualifications, shown by the sig. of the F test being <0.05. More specifically, the mean value of the groups shows that scientific research output increases gradually by the level of academic title and type of degree.

In terms of lecturer experience, there is a difference in scientific research output among lecturer groups with different seniorities, shown by the sig. of the Welch test being <0.05. More specifically, the mean value shows that scientific research output increases gradually with seniority.

Regarding the scientific research area, there is a difference in scientific research output among lecturer groups, shown by the sig. of the T-test being <0.05. More specifically, the mean value shows that scientific research output is higher for the field of natural sciences than it is for the social sciences.

Concerning managerial position, there is a difference in scientific research output between lecturer groups, shown by the sig. of the T-test being <0.05. More specifically, the mean value shows that scientific research output is higher for lecturers who do not hold managerial positions compared to those who do. This is likely because lecturers who hold managerial positions have busier schedules and less time allocated for conducting research.

In terms of abroad graduation, there is a difference in scientific research output among lecturer groups, shown by the sig. of the T-test being <0.05. More specifically, the mean value shows that scientific research output is higher for lecturers who graduated abroad.

4.2. Measurement Proposed Research Model Assessment

The psychometric properties of the scales measuring the considered dimensions of scientific research output were first examined through the Exploratory Factor Analysis (EFA) procedure, before being included in the PLS-SEM model. The study procedure was carried out using SmartPLS software version 3.3.3. The SEM model included six constructs, namely, scientific research objectives, decentralization, leadership, support for scientific research activities, policies for lecturers at HEIs, and resources for scientific research. To assess the measurement model, the convergent and discriminant validity and composite reliability were considered.

In Table 5, for this measurement model, all of the quality criteria were met, since all factor loadings, Cronbach's alpha values, composite reliability (CR) values, and average variance extracted (AVE) values were above the recommended thresholds (0.7, 0.7, 0.7, and 0.5, respectively) [24,49–51].

Discriminant validity is the degree to which items distinguish between constructs. Using the Fornell–Larcker criterion, the results indicate that the square root of the average variance extraction is greater than the inter-construct correlations. Regarding the cross-loadings criterion, the factor loadings of each item or indicator must be greater than the rest of its cross-loadings to ensure the discriminant validity of the construct [51]. Seven criteria were used to test discriminant validity, including Fornell–Larcker and cross-loadings. Table 6 illustrates the results of the discriminant validity measurements. In the first column, the square root of the extracted variance that appears in the upper part in parentheses must be greater than the correlations that appear in the following lines of the same column. This criterion was applied for each column. Table 7 shows the fulfillment of this criterion for all the subscales, demonstrating the discriminant validity of the tested instrument.

Table 5. Convergent validity and reliability.

Constructs	Items	Loadings	Cronbach's Alpha	rho_A	Composite Reliability	Average Variance Extracted (AVE)
Scientific research objectives of HEIs (OBJ)	OBJ2	0.829	0.869	0.879	0.905	0.656
	OBJ3	0.759				
	OBJ4	0.836				
	OBJ5	0.846				
	OBJ6	0.775				
Decentralization (DEC)	DEC1	0.731	0.860	0.866	0.895	0.587
	DEC2	0.747				
	DEC3	0.820				
	DEC4	0.790				
	DEC5	0.777				
	DEC6	0.728				
Leadership (LEA)	LEA2	0.801	0.768	0.811	0.845	0.577
	LEA3	0.791				
	LEA4	0.737				
	LEA5	0.704				
Support for scientific research activities (SUP)	SUP1	0.776	0.882	0.888	0.908	0.585
	SUP2	0.754				
	SUP3	0.766				
	SUP4	0.776				
	SUP5	0.737				
	SUP6	0.780				
	SUP7	0.763				
Policies for lecturers (POL)	POL1	0.812	0.894	0.897	0.919	0.654
	POL2	0.821				
	POL3	0.815				
	POL4	0.818				
	POL6	0.775				
	POL7	0.810				
Resources for scientific research (RES)	RES1	0.837	0.881	0.882	0.913	0.679
	RES2	0.870				
	RES3	0.834				
	RES4	0.785				
	RES5	0.791				
Scientific research outputs (OUT)	OUT1	0.784	0.900	0.901	0.920	0.589
	OUT2	0.751				
	OUT3	0.775				
	OUT4	0.767				
	OUT5	0.771				
	OUT6	0.806				
	OUT7	0.748				
	OUT8	0.734				

Table 6. Discriminant validity.

Scales	OBJ	DEC	LEA	SUP	POL	RES	OUT
OBJ	0.810						
DEC	0.271	0.766					
LEA	0.278	0.390	0.760				
SUP	0.227	0.253	0.192	0.765			
POL	−0.003	0.079	0.265	−0.011	0.809		
RES	0.291	0.268	0.186	0.241	0.049	0.824	
OUT	0.391	0.349	0.412	0.440	0.414	0.629	0.767

Table 7. Hypotheses testing.

Relationship	Hypothesis	Std. Beta	T Statistics (|O/STDEV|)	*p* Values	Decision
DEC->OUT	H2	0.039	1.236	0.217	Not supported
LEA->OUT	H3	0.121	3.306	0.001	Supported
POL->OUT	H5	0.359	10.872	0.000	Supported
RES->OUT	H6	0.471	16.025	0.000	Supported
OBJ->OUT	H1	0.151	4.768	0.000	Supported
SUP->OUT	H4	0.263	9.587	0.000	Supported

Note: OBJ: scientific research objectives of HEIs; DEC: decentralization; LEA: leadership; SUP: support for scientific research activities; POL: policies for lecturers; RES: resources for scientific research; OUT: scientific research outputs.

As the goal of SEM-PLS is to explain the endogenous latent variance, the key target is to have a higher R square. The greater the value, the better the explanatory power of the model [52]. The authors of [53] argued that a value of R square greater than 0.26 is considered substantial, as a rule of thumb. The results obtained in this present restudy show that the R square value for scientific research output was 0.684, which was acceptable. The corresponding results are presented in Figure 2.

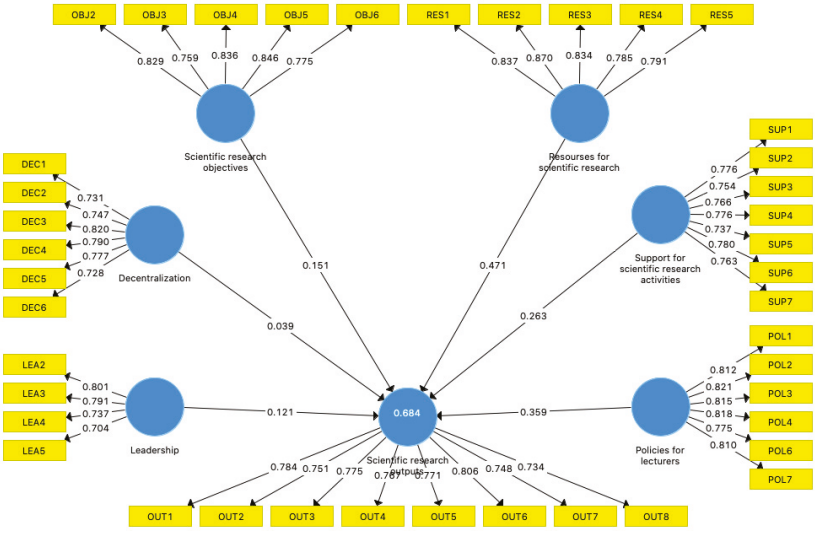

Figure 2. Confirmatory Factor Analysis Result.

4.3. Testing Research Hypotheses

In the structural model, the relevance and significance of all the direct and indirect effects were assessed, examining the path coefficients, associated t-statistics, and their bias-corrected confidence intervals, which were computed through a bootstrapping procedure. The study conducted the test with a sample size of bootstrapping N = 5000 [46,54]. The proposed hypotheses were considered statistically significant at the 99%, 95%, and 90% reliability levels.

The bootstrapping technique was used to analyze the t-statistics for the path coefficients to assess the importance of the hypothesized connections [55,56]. The *p*-value is a constant measure of evidence, but it is usually dichotomized into highly important, marginally important, and not statistically important at conventional levels, with cut-offs at $p \leq 0.01$, $p \leq 0.05$, and $p > 0.10$, respectively [57].

Table 7 shows that scientific research output ($p < 0.01$) shares a significant relationship with scientific research objectives, leadership, support for scientific research activities, policies for lecturers, and resources for scientific research. This means that five hypotheses in the conceptual model were fully supported (Hypothesis 1, Hypothesis 3, Hypothesis 4, Hypothesis 5, and Hypothesis 6). Decentralization had no direct effect on scientific research output ($p = 0.217 > 0.01$), thus Hypothesis 2 was not supported. Among these variables, resources for scientific research (RES) was the most effective factor ($\beta = 0.471$, $t = 16.025$, $p = 0.000$). Policies for lecturers (POL) and support for scientific research activities (SUP) had the second strongest influence on scientific research output ($\beta = 0.359$, $t = 10.872$, $p = 0.000$; $\beta = 0.263$, $t = 9.587$, $p = 0.000$, respectively). Scientific research objectives of HEIs (OBJ) and leadership (LEA) were also significant factors that affected scientific research output ($\beta = 0.151$, $t = 4.768$, $p = 0.000$; $\beta = 0.121$, $t = 3.306$, $p = 0.000$, respectively).

5. Discussion

5.1. Findings and Implications

The results reveal that "resources for scientific research" have the most significant impact on scientific research output by lecturers among the six given governance factors. This reflects the reality of such resources, especially the limited facilities of universities in Vietnam, with narrow university campuses that lack synchronization. Financial resources are also limited in terms of both volume and procedures because most Vietnamese HEIs remain partially dependent on the state's budget; there exist only 23 autonomous universities [58] with better resources.

This finding supports the results of prior studies, from a university governance perspective. By empirical evidence, Ref. [5,7,17,18,37,38] proved that resources for scientific research, including space, equipment, information systems, databases, expenses, funds, and colleagues with good research capacity had a positive influence on lecturers' scientific research output. In Vietnam, Ref. [13] revealed that factors such as the accessibility of international scientific documents, research data, experimental devices or tools, software, and funding sources played an important role in international publishing. The work by [16] demonstrated that the main barriers to publication in Vietnam are funding and time for research.

Therefore, the implication is that HEI managers should pay careful attention to ensuring adequate resources for scientific research. First of all, it is vital to improve research space, equipment, free information systems, databases, and digital libraries for scientific research activities at HEIs. In terms of finance, HEIs should attract more non-state budget revenues and establish funds for internationally-indexed publications, intellectual property applicants, and the commercialization of scientific research outcome. In addition, HEIs should reform the mechanism of budget allocation for scientific research in particular, at both the university and faculty levels.

Secondly, the obtained results demonstrate that "policies for lecturers" had a significant influence on lecturers' scientific research output. This finding provides a comprehensive assessment of how policies for lecturers affect their research outcomes. In fact, the income of lecturers at Vietnamese HEIs that are not yet autonomous is generally low, according to the general regulations of the State [59]. The regulations on workload and rewards for lecturers are all "one-size-fits-all" policies [14]. The finding of this research is consistent with the previous studies. However, provided with a governance perspective, the study has become more relevant for university administrators and management researchers. For instance, the work by [11,20,37,38,60] affirmed that policies related to income, recruitment, remuneration, evaluation, reward, retention, training, and development of teaching staff had a positive impact on the scientific research results of facilities. In Vietnam, Ref. [12,13] pointed out that "time for research purpose" affected the research productivity of Vietnamese social scientists.

Thus, it is recommended that university administrators pay close attention to developing and improving policies for lecturers in order to enhance the overall effectiveness of

scientific research. In Vietnam, the time available for research purposes must be taken into account by HEI managers and policy-makers. Sabbatical leave, which is popular in developed countries as a time to focus on research, should also be considered by Vietnamese HEIs. Furthermore, "tailor-made" incentives for different types of lecturers should be implemented in Vietnamese HEIs. Lecturers with high research output should be rewarded and paid differently compared to others with relatively lower research output.

Thirdly, the results of testing hypotheses show that both "support for scientific research activities" and "scientific research objectives of HEIs" had a positive effect on lecturers' scientific research output. This is also in line with previous research. The authors of [19] and [29] argued that scientific research support had the most significant effect on institutional factors. The work by [20] demonstrated that setting clear common goals in scientific research was one of the factors that had the strongest impact on the results of scientific research by lecturers. The influence of an organization's scientific research objectives on the faculty's scientific research results has also been tested and recommended by [6,19,22,32,61].

Therefore, the implication offered here is that HEIs should actively plan long- and medium-term strategies for scientific research activities, with the positioning of clear scientific research results, along with mechanisms and resources to ensure the validity of the strategies. They should not simply offer annual plans according to the force of administrative procedures. This solution can potentially increase scientific research results in a sustainable and focused manner.

Additionally, HEIs should pay close attention to creating an environment and culture that values and supports scientific research, building and developing research groups, and encouraging collaboration between scientists inside and outside the organization. The managers of Vietnam's HEIs also need to pay attention to administrative reforms, scientific research management mechanisms, and digital transformation to actively support scientific research activities.

Fourthly, the factor "leadership" had a positive influence on lecturers' scientific research output, but it had less of an impact compared to the above-mentioned factors. This finding is consistent with previous studies, but it has more significance when applied to university governance particularly. The authors of [9,20,30] demonstrated that institutions with academic excellence had leaders who could link scientific fields, select and train young faculty members, encourage and develop new scientific ideas, attract funds, build an environment to promote research and creativity, and set and disseminate goals.

Therefore, the implication for university governance is to further improve the awareness of HEIs' leaders of the importance of scientific research in the current context. They need to have a sense of regular and clear communication about the organization's scientific research goals, always recognize and appreciate the results of scientific research, and be fair in resource allocation. The selection and appointment of managers in departments, laboratories, faculties, or universities should have specific criteria that are suitable for the particular job (i.e., high creativity and different from ordinary administrative management). In addition, in the current integration context, HEIs also need to pay attention to fostering and improving management skills towards international standards for university managers.

Fifthly, "decentralization" is a new factor that was added in the research model, and was the only factor that did not have an impact on lecturers' scientific research output. Due to the relevant findings in previous studies, this factor was tested and included in the research model. For example, Ref. [20] demonstrated that participative governance was a characteristic of organizations with high scientific research achievements. The authors of [30] clarified that autonomy in building a research team and staff evaluation policy each had a good influence on the results of scientific research. This result can be explained and supported by several prior studies. The theory of management Y [62] noted that it is impossible to have complete autonomy to achieve both organizational goals and individual needs. Therefore, there is a need to decentralize at an appropriate level. The authors of [63]

argued that decentralization does not mean sharing power to the extent that senior leaders in the organization do not know the important rights of subordinates.

Interestingly, to the best of our knowledge, this may be the first study to quantify "decentralization" in relation to university lecturers' research results. It has experimentally proven whether decentralization affects the research performance of lecturers. Although the results of testing the model and hypothesis H2 show that this factor does not affect the scientific research output of lecturers with a significance level <=0.05, this finding still forms a theoretical and practical contribution as a basis for formulating appropriate university governance policies.

Overall, this study has built a model and verified the influence of governance factors on the scientific research output of lecturers. The results obtained can act as a guideline to aid university managers in improving governance by prioritization of resources for scientific research, policies for lecturers, support for scientific research activities, setting scientific research objectives, and leadership.

Sixthly, the *t*-test and ANOVA results indicate that there is a difference in the scientific research output between groups of lecturers. More specifically, the scientific research output by males is more than that by females. The output of natural sciences is more than that of social sciences. The output by lecturers who do not hold managerial positions is more than that by those who hold such positions. The output by graduates from universities abroad is more than that by graduates from domestic universities. The output by those with advanced degrees is more than that by those with relatively low ones. The output by higher seniority is more than that by lower seniority. Finally, the output by the middle-aged group is more than that by the elderly and younger lecturer groups. The difference between these groups is also mentioned in prior studies [7,11].

Thus, it is recommended that HEI managers formulate suitable policies for different groups of lecturers. For example, the *t*-test and ANOVA results show that lecturers in the natural sciences have higher scientific research output compared to those in the social sciences. Hence, HEIs should have appropriate resource investment policies for each field, and during each period. To rapidly increase scientific research results in the short term, HEIs should focus on attracting and prioritizing more investment for groups of lecturers with higher scientific research output, such as groups with doctoral degrees or higher (especially associate professors and professors); those who graduated abroad; those in the 31–40 and 41–50 age groups; those with at least three years of seniority; and those in the natural sciences. However, to ensure sustainable and balanced development in the long term, HEIs need to pay attention to the scientific research activities of the remaining groups. In addition, the *t*-test results show that the group that does not hold a management position had higher scientific research output than the other. Therefore, the implication is that HEIs should change their mindset about appointing managers: excellent scientists should manage research groups, laboratories, and centers of excellence, without the necessity of holding other administrative positions that limit time and scientific research performance.

Finally, this study has a novel contribution in terms of research methodology on this topic. As the literature showed, it is difficult to find a study that adopted PLS-SEM, even though this approach was considered to be the best alternative to CB-SEM in cases where there is little theoretical background available [30,48]. Moreover, in this study, PLS-SEM was employed in conjunction with the bootstrap technique, *t*-test, and ANOVA to provide detailed insights on this increasingly important topic. In Vietnam, to the best of our knowledge, this is the first study that utilized the SEM approach with such a large population and primary data for this particular scientific issue.

Furthermore, compared with other works in the literature, this work can be considered a rare empirical study that has focused on testing the influence of governance factors (i.e., not institutional factors in general) on the scientific research output of university lecturers. Although the impact of management activities on employee performance is no longer a new research topic, the number of studies focusing exclusively on governance factors

affecting the scientific research productivity of lecturers is lacking. To date, most of the studies only consider the institutional, environmental, and personal factors that affect the results of scientific research.

In summary, this study not only offers theoretical contributions, but practical contributions as well, helping policymakers, managers, and lecturers to take appropriate solutions for groups of lecturers in order to enhance scientific research performance.

5.2. Limitations and Suggestions for Further Studies

All studies have their limitations, and this study is no exception. Firstly, this study uses a non-probability sampling method, so PLS-SEM has not been analyzed by different groups of lecturers (according to demographic characteristics) to understand the influence of governance factors on scientific research output by the groups. Meanwhile, non-parametric measures (Kruskal–Wallis Test, Mann–Whitney U Test, etc.) that are considered more suitable for non-probability sampling have not been utilized.

Second, this study explores the influence of governance factors based solely on lecturers' perceptions of their scientific research output. In other words, this is a cross-sectional study, so it is impossible to compare the change of scientific research output in reality, according to specific time and space milestones.

Third, this study assumes that governance factors have a direct influence on scientific research output. However, in reality there are other mediating factors, such as the behavior, motivation, and attitude of lecturers. Therefore, it is recommended that intermediary factors (both as independent and dependent variables) be added to the research model. Additionally, the review of the literature and practice also showed that there are many other governance and personal factors that affect scientific research output which have not been tested in this study.

Finally, the scope of this research has not focused on research-oriented universities in Vietnam in particular. Therefore, caution should be taken when generalizing the results of this study to all types of HEIs.

In the future, it is more reliable and worthwhile if researchers deploy the probability sampling method, thereby applying PLS-SEM for different sample groups and employing ANOVA more deeply for each sample group. It is possible to use time string data or array data (not cross-sectional data) to conduct empirical research to clarify the relationship between governance factors and scientific research results, which can be expressed after a long time. In addition, it is necessary to explore the relationship between governance factors and the attitude, motivation, and behavior of lecturers with respect to scientific research and scientific research output. It is also possible to combine different approaches (i.e., other individual factors and governance factors) to explain why, in the same environment and with the same governance, lecturers have differing scientific research output. Finally, it is necessary to expand the scope of the research according to groups of HEIs, which can be divided by their research or practice orientation, autonomy, whether they are multidisciplinary or focus on a single field, and whether they are public or private universities.

6. Conclusions

This study has proposed and empirically tested a research model on governance factors affecting the scientific research output of university lecturers. It has adjusted the measurement criteria (observed variables) of the governance factor scales and scientific research output according to the context of Vietnam. The results show that the level of influence of governance factors on lecturers' scientific research output, in order from strongest to weakest, is as follows: resources for scientific research, policies for lecturers, support for scientific research activities, scientific research objectives of HEIs, and leadership. A new scale of "decentralization" has been added to the analytical framework, but it was not statistically significant regarding its impact on lecturers' research output.

In addition, the study examined the difference in the scientific research output of lecturer groups according to their demographics, including gender, age, degree and aca-

demic title, experience, position (manager or not), place of graduation (abroad or domestic), and scientific area. The obtained results reveal that there are differences in the research output between lecturer groups. Hence, they provide empirical evidence for implications in management decision-making and policy-making.

Overall, the obtained results can guide HEIs in evaluating the current status of scientific research activities and the governance factors that influence lecturers' scientific research output, thereby helping these HEIs take appropriate measures to enhance their scientific research achievements in today's knowledge-based economy.

Author Contributions: Truong, H.T. and Le, H.M. have mainly contributed for giving the paper idea, writing, reviewing and editing. Le, D.A.; Nguyen, H.T.; and Nguyen, T.K. contributed for interviewing and collecting the data, questionnaire designing and result calculating. Do, D.A contriuted for reading and giving the comment and conclusion parts of the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare that they do not have any conflict of interest.

Appendix A Description of Scales

	Code	Observed Variables	Sources
		Scientific research objectives of HEIs	
1.	OBJ1	My university always attaches great importance to scientific research objectives.	[4,6,32]
2.	OBJ2	The scientific research objectives have been set with coordination among the faculties and departments.	[1,20,22]
3.	OBJ3	The scientific research objectives are disseminated via many different channels (website, email, software, etc.).	[20,22]
4.	OBJ4	The scientific research objectives are clear and measurable.	[17]
5.	OBJ5	The scientific research objectives are feasible.	[20,22]
6.	OBJ6	The scientific research objectives are in line with the interests and desires of lecturers.	[1,22]
7.	OBJ7	My university has high requirements for the scientific research output of lecturers.	[20]
		Decentralization	
8.	DEC1	My university encourages the participation of lecturers in the process of management decision-making related to scientific research.	Suggested by the authors
9.	DEC2	Faculties and departments are empowered to decide many professional and academic-related issues.	Suggested by the authors
10.	DEC3	Lecturers are facilitated to pursue their research directions in compliance with their professional capacity.	Suggested by the authors
11.	DEC4	Lecturers actively participate in bidding and implementing research projects.	Suggested by the authors
12.	DEC5	The management of the scientific research output of lecturers is based on autonomy.	Suggested by the authors
13.	DEC6	My university has channels for listening to lecturers' feedback.	Suggested by the authors
		Leadership	
14.	LEA1	The university's leaders tend to communicate clearly about the objectives and orientation of scientific research activities.	[22,30]
15.	LEA2	The university's leaders have a high reputation and are respected.	[30]
16.	LEA3	The university's leaders always show recognition and appreciation for the scientific research output of lecturers.	[20,21]
17.	LEA4	The university's leaders are fair in allocating and approving research projects.	[60]
18.	LEA5	The direct leaders (at the Faculty/Department level) have a high reputation for scientific research.	[30]

	Code	Observed Variables	Sources
19.	LEA6	The direct leaders (at the Faculty/Department level) have good management capacity.	[20]
20.	LEA7	The direct leaders (at the Faculty/Department level) always have effective guidance and orientation in scientific research activities.	[20]
21.	LEA8	The direct leaders (Faculty/Department level) actively support and encourage the faculty's scientific research efforts.	[9,61]
22.	LEA9	The direct leaders (at the Faculty/Department level) provide fair and competitive development opportunities for lecturers.	[9,61]
		Support for scientific research activities	
23.	SUP1	My university puts effort to create an environment and culture that values scientific research activities.	[17,18,20,36]
24.	SUP2	My university puts effort to simplify administrative procedures (within its ability) to support scientific research activities.	[17,36,60]
25.	SUP3	My university has attracted many funding sources for research and development activities.	[5]
26.	SUP4	My university supports lecturers effectively in the transfer and commercialization of research output.	[33,36]
27.	SUP5	My university encourages the establishment and development of research groups and collaborations.	[13,14,27,30,45]
28.	SUP6	My university's departments and faculties always support lecturers' scientific research effectively.	[18,32]
29.	SUP7	My university organizes several scientific activities such as conferences, seminars, workshops, and training courses.	[17,60]
		Policies for lecturers	
30.	POL1	Policy on lecturer recruitment focuses on scientific research capacity.	[1,35]
31.	POL2	The reward regime is encouraging and motivating for lecturers.	[11,32]
32.	POL3	The income regime is encouraging and motivating for lecturers.	[11,17,31]
33.	POL 4	Lecturers are facilitated to become main lecturers, senior lecturers, associate professors, and professors.	[11,35,37]
34.	POL 5	Lecturers are facilitated to improve their capacity at home and abroad.	[17,32]
35.	POL 6	Lecturers are facilitated to arrange the time for scientific research.	[13,15,17]
36.	POL 7	Young lecturers are trained, consulted, and oriented on scientific research.	[34,36]
		Resources for scientific research	
37.	RES1	Lecturers are provided with research space (e.g., private room/desk).	[17,18,61]
38.	RES2	Lecturers are provided with the necessary equipment for scientific research (computers, internet, printers, laboratory equipment, etc.).	[7,13,18,30]
39.	RES3	Lecturers can easily access free information systems, databases, and necessary documents for scientific research at the university, faculty library.	[13,18,32,61]
40.	RES4	Lecturers are supported with funding for internationally-indexed publications/intellectual property applicants/commercialization of scientific research products, or funding support to carry out scientific research works.	[5,13,18]
41.	RES5	My university has many lecturers with good scientific research abilities.	[17,18,30]

	Code	Observed Variables	Sources
		Scientific research output	
1.	OUT1	The number of my international scientific articles in ISI/Scopus-indexed journals tends to increase every year.	
2.	OUT2	The number of my other international scientific publications tends to increase every year.	
3.	OUT3	The number of my scientific papers in domestic journals tends to increase every year.	
4.	OUT4	The number of my research papers published in scientific conference proceedings tends to increase every year.	[4,7,20,37]
5.	OUT5	The number of my books/textbook chapters/monographs tends to increase every year.	
6.	OUT6	The number of my patents, inventions, and registration of intellectual property tends to increase every year.	
7.	OUT7	The number of scientific research projects that I lead every year tends to increase.	
8.	OUT8	The number of scientific research awards that I and my learners get tends to increase every year.	

References

1. Jung, J. Faculty Research Productivity in Hong Kong across Academic Discipline. *High. Educ. Stud.* **2012**, *2*, 4. [CrossRef]
2. Vasiljeva, M.; Ponkratov, V.; Volkova, T.; Khairova, S.; Nikitina, N.; Dudnik, O.; Alimova, M.; Kuznetsov, N.; Elyakova, I. The Development of Scientific Activity in Russian Universities. *J. Open Innov. Technol. Mark. Complex.* **2020**, *6*, 110. [CrossRef]
3. Lamanauskas, V.; Augienė, D. Development of Scientific Research Activity in University: A Position of the Experts. *Procedia-Soc. Behav. Sci.* **2015**, *167*, 131–140. [CrossRef]
4. Hedjazi, Y.; Behravan, J. Study of factors influencing research productivity of agriculture faculty members in Iran. *High. Educ.* **2011**, *62*, 635–647. [CrossRef]
5. Ghabban, F.; Selamat, A.; Ibrahim, R.; Krejcar, O.; Maresova, P.; Herrera-Viedma, E. The Influence of Personal and Organizational Factors on Researchers' Attitudes towards Sustainable Research Productivity in Saudi Universities. *Sustainability* **2019**, *11*, 4804. [CrossRef]
6. Santos, J.; Horta, H. The research agenda setting of higher education researchers. *High. Educ.* **2018**, *76*, 649–668. [CrossRef]
7. Nafukho, F.; Wekullo, C.S.; Muyia, M.H. Examining research productivity of faculty in selected leading public universities in Kenya. *Int. J. Educ. Dev.* **2019**, *66*, 44–51. [CrossRef]
8. Altbach, P.G. What Counts for Academic Productivity in Research Universities? *Int. High. Educ.* **2015**, 6–7. [CrossRef]
9. Heinze, T.; Shapira, P.; Rogers, J.D.; Senker, J.M. Organizational and institutional influences on creativity in scientific research. *Res. Policy* **2009**, *38*, 610–623. [CrossRef]
10. Shin, D.; Konrad, A. Causality Between High-Performance Work Systems and Organizational Performance. *J. Manag.* **2016**, *43*, 973–997. [CrossRef]
11. Chen, Y.; Nixon, M.R.; Gupta, A.K.; Hoshower, L. Research Productivity Of Accounting Faculty: An Exploratory Study. *Am. J. Bus. Educ. (AJBE)* **2010**, *3*, 101–115. [CrossRef]
12. Vuong, Q.-H.; La, V.-P.; Vuong, T.-T.; Ho, M.-T.; Nguyen, H.K.; Nguyen, V.-H.; Pham, H.-H.; Ho, M.-T. An open database of productivity in Vietnam's social sciences and humanities for public use. *Sci. Data* **2018**, *5*, 180188. [CrossRef] [PubMed]
13. Tran, T.; Trinh, T.-P.; Le, C.-M.; Hoang, L.-K.; Pham, H.-H. Research as a Base for Sustainable Development of Universities: Using the Delphi Method to Explore Factors Affecting International Publishing among Vietnamese Academic Staff. *Sustainability* **2020**, *12*, 3449. [CrossRef]
14. Trinh, T.; Tran, T.; Le, H.; Nguyen, T.; Pham, H. Factors impacting international-indexed publishing among Vietnamese educational researchers. *Learn. Publ.* **2020**, *33*, 419–429. [CrossRef]
15. Pham, L.T.; Hayden, M. Research in Vietnam: The Experience of The Humanities and Social Sciences. *J. Int. Comp. Educ.* **2019**, *8*, 27–40. [CrossRef]
16. Pho, P.D.; Tran, T.M.P. Obstacles to Scholarly Publishing in the Social Sciences and Humanities: A Case Study of Vietnamese Scholars. *Publications* **2016**, *4*, 19. [CrossRef]
17. Okendo, O.E. Constraints of Research Productivity in Universities in Tanzania: A Case of Mwenge Catholic University. *Int. J. Educ. Res.* **2018**, *6*, 201–210.
18. Yang, J.C.-C. A Study of Factors Affecting University Professors' Research Output: Perspectives Of Taiwanese Professors. *J. Coll. Teach. Learn. (TLC)* **2017**, *14*, 11–20. [CrossRef]
19. Tafreshi, G.H.; Imani, M.N.; Ghashlag, P.M. Designing a model for research productivity evaluation of faculty of district 2 of Islamic Azad University of Iran. *World Appl. Sci. J.* **2013**, *21*, 12. [CrossRef]
20. Bland, C.J.; Center, B.A.; Finstad, D.A.; Risbey, K.R.; Staples, J.G. A Theoretical, Practical, Predictive Model of Faculty and Department Research Productivity. *Acad. Med.* **2005**, *80*, 225–237. [CrossRef]

21. Sheridan, J.; Savoy, J.N.; Kaatz, A.; Lee, Y.-G.; Filut, A.; Carnes, M. Write More Articles, Get More Grants: The Impact of Department Climate on Faculty Research Productivity. *J. Women's Health* **2017**, *26*, 587–596. [CrossRef] [PubMed]
22. Kok, S.K.; McDonald, C. Underpinning excellence in higher education—An investigation into the leadership, governance and management behaviours of high-performing academic departments. *Stud. High. Educ.* **2015**, *42*, 1–22. [CrossRef]
23. Negash, M.; Lemma, T.T.; Samkin, G. Factors impacting accounting research output in developing countries: An exploratory study. *Br. Account. Rev.* **2018**, *51*, 170–192. [CrossRef]
24. Hair, J.; Black, W.; Babin, B.; Anderson, R. Multivariate Data Analysis: A Global Perspective. In *Multivariate Data Analysis: A Global Perspective*; Pearson Education: Upper Saddle River, NJ, USA; London, UK, 2010; Volume 7.
25. Quy, N.H. Factors Influencing the Research Productivity of Academics at the Research-Oriented University in Vietnam. Ph.D. Thesis, Griffith University, Brisbane, QLD, Australia, 2015.
26. Shao, Z.; Li, Y.; Wu, K.; Guo, Y.; Feng, F.; Hui, F.; Niu, Y.; Zheng, Y. How academic librarians involve and contribute in research activities of universities? A systematic demonstration in practice through comparative studies of research productivities and research impacts. *J. Acad. Librariansh.* **2018**, *44*, 805–815. [CrossRef]
27. Li, Y. Seeking entry to the North American market: Chinese management academics publishing internationally. *J. Engl. Acad. Purp.* **2014**, *13*, 41–52. [CrossRef]
28. Salinas-Ávila, J.; Abreu-Ledón, R.; Tamayo-Arias, J. Intellectual capital and knowledge generation: An empirical study from Colombian public universities. *J. Intellect. Cap.* **2020**, *21*, 1053–1084. [CrossRef]
29. Wichian, S.N.; Wongwanich, S.; Bowarnkitiwong, S. Factors affecting research productivity of faculty members in government universities: LISREL and Neural Network analyses. *Kasetsart J. Soc. Sci.* **2009**, *30*, 1.
30. Hwang, H.; Malhotra, N.K.; Kim, Y.; Tomiuk, M.A.; Hong, S. A Comparative Study on Parameter Recovery of Three Approaches to Structural Equation Modeling. *J. Mark. Res.* **2010**, *47*, 699–712. [CrossRef]
31. Salman, S.; Kausar, T.; Furqan, M. Factors Affecting Research Productivity in Private Universities of Lahore: A Discriminant Analysis. *Pak. Bus. Rev.* **2018**, *30*, 1.
32. Jahan, F.; Siddiqui, M.A.; Khabouri, Z.; Riyami, R.; Ahuja, A. Factors influencing quality of academic research: Perception of faculty researchers at Oman Medical College. *Med. Clin. Arch.* **2018**, *2*. [CrossRef]
33. Rubin, A.; Callaghan, C.W. Entrepreneurial orientation, technological propensity and academic research productivity. *Heliyon* **2019**, *5*, e02328. [CrossRef] [PubMed]
34. Yumeen, S.; Ho, E.S.; Wong, K.; Borschel, G.H. What Factors Influence Resident Research Publication in the Division of Plastic Surgery? *J. Surg. Educ.* **2018**, *75*, 409–416. [CrossRef] [PubMed]
35. Aydın, O.T. The Impact of Theory X, Theory Y and Theory Z on Research Performance: An Empirical Study from A Turkish University. *Int. J. Adv. Manag. Econ.* **2012**, *1*, 24–30. [CrossRef]
36. Hosseinifar, A.Z.T.; Isfandyari-Moghaddam, A. Original Article Factors Affecting and Inhibiting Scientific Productivity from the Viewpoint of Faculty Members of Kermanshah University of Medical Sciences. *Educ. Res. Med Sci.* **2016**, *5*, 29–36.
37. Hoffmann, K.; Berg, S.; Koufogiannakis, D. Understanding Factors that Encourage Research Productivity for Academic Librarians. *Évid. Based Libr. Inf. Pract.* **2017**, *12*, 102–128. [CrossRef]
38. Aydin, O.T. Research performance of higher education institutions: A review on the measurements and affecting factors of research performance. *J. High. Educ. Sci.* **2017**, *7*, 312. [CrossRef]
39. Tekneci, P.D. Evaluating Research Performance of Turkish Universities. Ph.D. Thesis, Middle East Technical University, Ankara, Turkey, 2014.
40. White, C.S.; James, K.; Burke, L.A.; Allen, R.S. What makes a "research star"? Factors influencing the research productivity of business faculty. *Int. J. Prod. Perform. Manag.* **2012**, *61*, 584–602. [CrossRef]
41. Costa, C.A.B.E.; Oliveira, M. A multicriteria decision analysis model for faculty evaluation. *Omega* **2012**, *40*, 424–436. [CrossRef]
42. Hirsch, J.E. An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16569–16572. [CrossRef]
43. Abramo, G.; D'Angelo, A.C.; Murgia, G. The relationship among research productivity, research collaboration, and their determinants. *J. Inf.* **2017**, *11*, 1016–1030. [CrossRef]
44. Yamamoto, K.; Ishikawa, M. A Normalized Index to Quantify Publication Productivity across Disciplinary Boundaries. *Int. J. Inst. Res. Manag.* **2017**, *1*, 67–81. [CrossRef]
45. Chang, O.H.; Chow, C.W. Publishing in Internationally Respected Accounting Journals: Insights and Suggestions of Successful Chinese Authors. *Asia-Pac. J. Account.* **1998**, *5*, 150–165. [CrossRef]
46. Henseler, J.; Ringle, C.M.; Sarstedt, M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J. Acad. Mark. Sci.* **2014**, *43*, 115–135. [CrossRef]
47. Lowry, P.; Gaskin, J. Partial Least Squares (PLS) Structural Equation Modeling (SEM) for Building and Testing Behavioral Causal Theory: When to Choose It and How to Use It. *IEEE Trans. Dependable Secur. Comput.* **2014**, *57*, 123–146. [CrossRef]
48. Wong, K.K.-K. Partial Least Squares Structural Equation Modeling (PLS-SEM) Techniques Using SmartPLS. *Mark. Bull.* **2013**, *24*, 1.
49. Fornell, C.; Larcker, D.F. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *J. Mark. Res.* **1981**, *18*, 39. [CrossRef]
50. Hulland, J. Use of partial least squares (PLS) in strategic management research: A review of four recent studies. *Strateg. Manag. J.* **1999**, *20*, 2. [CrossRef]

51. Hair, J.; Hult, G.T.; Ringle, C.; Sarstedt, M. *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*; Sage Publishing: New York, NY, USA, 2016.
52. Huang, C.-C.; Wang, Y.-M.; Wu, T.-W.; Wang, P.-A. An Empirical Analysis of the Antecedents and Performance Consequences of Using the Moodle Platform. *Int. J. Inf. Educ. Technol.* **2013**, 217–221. [CrossRef]
53. Cohen, J. *Statistical Power Analysis for the Behavioural Sciences*; Lawrence Earlbaum Associates Publisher: Mahwah, NJ, USA, 1988.
54. Do, A.D.; Pham, N.T.; Bui, H.P.; Vu, D.T.; Nguyen, T.K.; Nguyen, T.H. Impact of Motivational Factors on the Work Results of Lecturers at Vietnam National University, Hanoi. *J. Asian Financ. Econ. Bus.* **2020**, *7*, 425–433. [CrossRef]
55. Efron, B. Bootstrap Methods: Another Look at the Jackknife BT-Breakthroughs in Statistics: Methodology and Distribution. In *Breakthroughs in Statistics*; Springer: New York, NY, USA, 1992.
56. Yung, Y.-F.; Bentler, P.M. Bootstrap-corrected ADF test statistics in covariance structure analysis. *Br. J. Math. Stat. Psychol.* **1994**, *47*, 63–84. [CrossRef]
57. Rice, W.R. Analyzing Tables of Statistical Tests. *Evolution* **1989**, *43*, 223. [CrossRef]
58. Hue, N. Autonomous University 'on Paper': Is It Time to Ditch the Governing Ministry? 2020. Available online: https://tienphong.vn/dai-hoc-tu-chu-tren-giay-da-den-luc-bo-bo-chu-quan-post1289429.tpo (accessed on 16 April 2021).
59. Long, P. For Successful University Autonomy: The Income of Lecturers Must Be at 50 Million/Month. 2020. Available online: https://giaoduc.net.vn/giao-duc-24h/de-tu-chu-dai-hoc-thanh-cong-thu-nhap-giang-vien-phai-o-muc-50-trieu-thang-post215605.gd (accessed on 16 April 2021).
60. Farzaneh, E.; Amani, F.; Taleghani, Y.; Fathi, A.; Kahnamouei-Aghdam, F.; Fatthzadeh-Ardalani, G. Research barriers from the viewpoint of faculty members and students of Ardabil University of Medical Sciences, Iran, 2014. *Int. J. Res. Med. Sci.* **2016**, 1926–1932. [CrossRef]
61. Whelan, K.; Markless, S. Factors that Influence Research Involvement among Registered Dietitians Working as University Faculty: A Qualitative Interview Study. *J. Acad. Nutr. Diet.* **2012**, *112*, 1021–1028. [CrossRef]
62. McGregor, D. *The Human Side of Enterprise*; McGrawHill: New York, NY, USA, 2006.
63. Le, V.T.; Ngo, K.T. *Textbook of Corporate Governance*; National Economics University Publishing House: Hanoi, Vietnam, 2015.

education sciences

MDPI

Article

Needs Analysis of Psychosocial Module Development Based on Psychoeducation Approach for Public University Students in Malaysia

Salleh Amat [1], Harizah Izyan Samsudin [1], Abu Yazid Abu Bakar [1,*], Mohd Izwan Mahmud [1] and Mohd Hasrul Kamarulzaman [2]

[1] Faculty of Education, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia; sallehba@ukm.edu.my (S.A.); sharizahizyan@gmail.com (H.I.S.); izwan@ukm.edu.my (M.I.M.)
[2] The Language Center, National Defence University of Malaysia, Kem Sungai Besi, Kuala Lumpur 57000, Malaysia; hasrul@upnm.edu.my
* Correspondence: yazid3338@ukm.edu.my

Abstract: This study is a needs analysis that aims to develop a psychosocial module based on a psychoeducation approach in order to strengthen the assertiveness, social skills, and psychological well-being of public university students in Malaysia. The specific objectives of this study were to identify the perceptions of university students on the need to design and build a psychosocial module, identify the constructs that should be included in the phase of designing and building of the psychosocial module as well as the support that has been received and the support that the students desire while on campus. This survey study utilized a questionnaire distributed to a total of 286 first year students studying at three public universities in Malaysia. The data obtained were analyzed by using descriptive statistics. The findings of the study showed a highly positive response from the students toward the development of a psychosocial module and the constructs listed in the questionnaire. Implications from this study can further solidify the constructs that should be included in the design and construction of the psychosocial module as well as identify the availability of the approaches, strategies, and support activities that have been received and desired by the students.

Keywords: needs analysis; psychosocial module development; psychoeducation; assertiveness; social skills; psychological well-being; university students in Malaysia

Citation: Amat, S.; Samsudin, H.I.; Bakar, A.Y.A.; Mahmud, M.I.; Kamarulzaman, M.H. Needs Analysis of Psychosocial Module Development Based on Psychoeducation Approach for Public University Students in Malaysia. *Educ. Sci.* **2021**, *11*, 604. https://doi.org/10.3390/educsci11100604

Academic Editors: Sandra Raquel Gonçalves Fernandes, Marta Abelha and Ana Teresa Ferreira-Oliveira

Received: 29 August 2021
Accepted: 23 September 2021
Published: 1 October 2021

1. Introduction

Although higher learning institutions are considered stable institutions for accommodating a group of high achievers with great intellectual capacity, there are, however, some students who face many problems or difficulties that cause them to be unable to complete their studies. Among the many issues discussed at the national and international levels are mental health and the psychological well-being of students [1]. According to the World Health Organization (WHO), the proportion of students declaring themselves as facing psychological issues against the whole population of students in universities worldwide have increased in recent years [2]. Furthermore, a research conducted by researchers from the WHO World Mental Health International College Student Initiative (WMH-ICS) also showed that one third of first year students experienced at least one psychological disorder such as anxiety, panic disorder, or substance disorder [3–5]. Other research conducted by Lipson [6] and Evans [7] also showed that the issue of the students' mental and psychological well-being was worse and more serious compared to the non-student population of the same age.

In their study, Shamunip and Mageswari [8] evaluated the psychological well-being of 308 Malaysian undergraduates and revealed that the students were operating at a sub-optimal functioning level and experiencing depression. According to Manap et al. [9],

being confined to their hostel rooms or homes without social contact during the recent quarantine made these levels worse. In the field of education, low level of psychological well-being can have a negative impact on the students' achievement and therefore, less or no motivation to continue their studies. In addition, Zahra and Soleimani [10] found the existence of a relationship between psychological well-being and cognitive achievement with academic achievement such as average grade point and exam scores. According to Ryff [11], psychological well-being is determined by a few functionalities in life such as having a positive relationship with others, environmental mastery, autonomy, self-acceptance, purpose in life, and personal growth. However, according to Erpianaand and Fourianalistyawati [12], psychological well-being is seen as the individual's capabilities in facing challenges and striving to function completely. On another account, according to Suvera [13], psychological well-being is connected to the physical and mental well-being of an individual. It also has a connection with a person's feelings toward daily life and psychological pressures such as anxiety, depression, disappointment, emotional fatigue, uneasiness, and dissatisfaction. Generally, a student's psychological well-being is focused on life's experience, feelings of oneself, and good mental health. Therefore, there is a need for the psychological well-being of students to be given more serious attention by all parties.

The effects of academic experience in the first year may differ at every university, but under normal circumstances, it depends on the level of the students' academic and social integration. Therefore, the most important goal for all universities is not only to provide knowledge and focus on student well-being, but also to encourage students to build social skills for their current academic and professional pursuits. This is because first year students not only have to adjust to university life, but they have to adapt to the environment, new social relationships, and different cultures [14,15]. Nonetheless, first year students require different social competencies throughout their academic years such as social skills and assertiveness as mechanisms in communication, adaptability, and well-being as well as strengthening their relationships with their fellow classmates, faculty, lecturers, and community [16,17]. However, the study by Chin [18] found that many university students in Malaysia had low levels of social skills, which were non-assertive [19], low communication efficiency, and as in many countries, on average, many students were also found to have moderate skill level in future orientation, critical thinking, identity creation, and pro social tendencies [20,21].

Before explaining the concept of social skills, it should be known that there are several terms related to social skills, namely interpersonal skills [22], interpersonal abilities, social capabilities [23], and interpersonal and social communication [24]. However, social skills are generally defined as intra and interpersonal skills such as communication skills, teamwork, and cooperation. In addition, social skills are also a class of different social behaviors that help individuals to cope with the interpersonal interaction process. In other words, social skills can contribute to positive interpersonal relationships, which then lead to better health, self-satisfaction, professional achievement, better teamwork, quality life, and respect for others [25,26]. Assertiveness is also one of the components in social skills. Assertiveness is important in helping students to be more open, confident, improve satisfaction, and be able to establish good relationships. This aspect of skills should be emphasized to produce students who are not only knowledgeable in the academic field, but also students who have high competitiveness, are able to communicate effectively, able to think creatively, critically, and to act rationally. There are a few studies that have proven that assertiveness, social skills, and psychological well-being can be shaped and enhanced in someone at school, college, and even at the university through intervention or exercise. For example, Hamidah [27] conducted a study to examine the effectiveness of cognitive behavior therapy (CBT) group counseling, and social skills training on student psychological well-being and social skill. The results showed that the intervention was effective in improving the students' psychological well-being and social skills. Robson et al. [28] conducted a study to determine the effectiveness of psychoeducational interventions in

improving the social skills of 57 medical students studying in bilingual universities. The results showed that there was an average increase in scores related to social skills, assertiveness, and academic experience. This proves that all three aspects can be enhanced and further developed through psychoeducation, counselling, workshops or any given exercises [29–31]. However, a psychoeducation approach is often used in institutions such as colleges and universities to assist the counsellor or facilitator to convey information on a short-term basis, especially with the increase in student numbers on a daily basis. Through this approach, information is conveyed in various methods such as lectures, slide shows, videos, forum discussions, and others. Psychoeducation is an approach that focuses on the client's psychological aspect and also educates clients on the techniques and strategies that are educational. It also plays a similar role to counselling and psychotherapy in molding the individuals toward being more positive [32–34]. Therefore, the whole problem demands an intervention that can play a dominant role in promoting psychological well-being, developing social skills as well as assertiveness and thus increasing the probability of guidance and counselling services being sought and seen by students as high quality, effective, and comfortable resources.

Besides, many researchers have mentioned the need for a regular evaluation of the needs of university students [35,36], highlighting its multiple benefits, and the possibility of designing different interventions according to the specific needs of the students. For example, Joana et al. [36] conducted a study to explore the psychological intervention needs of the students attending Portuguese Catholic University. The results indicate that participants need more information or support in the areas of the management of anxiety and nervousness, fear of failure, job search, career choice, and time management skills. However, it is not necessarily the case that institutions of higher education in this country have invested in the creation of these interventions, and at times, their presence and availability does not guarantee that they are demanded, since students do not always see them as an effective and convincing resource to support their needs [37]. In this sense, the present study aims to explore the student's needs for psychosocial modules based on a psychoeducation approach in strengthening assertiveness, social skills, and psychological well-being. For this purpose, the needs related to constructs that should be included in the phase of designing the module were also explored as well as the preferences regarding the desired support and the support that was previously received.

2. Methodology

The focus of this study was to carry out a needs analysis to (i) explore the student's needs of a psychosocial module; (ii) determine the construct required in the design and construction phase of a psychosocial module aimed to strengthen assertiveness, social skill, and psychological well-being university students; and (iii) examine the support that has been received and the support desired by students. The needs analysis used survey design and data that were analyzed through a quantitative approach [38,39]. The study sample selected was 286 first year university students in Malaysia. These students were randomly selected from three public universities in Malaysia. This study used an online questionnaire in which a Google form acted as the response platform. The general construct of this questionnaire was divided into two sections. Part A consisted of the respondent's demographic information such as gender, race, field of study, socio-economic status (SSE), and place of residence. Meanwhile, Part B consisted of three questions involving the needs for the development of a psychosocial module, the constructs required in the design and construction of a psychosocial module, and the support that has been received and the support desired by students. Part B uses the Guttman Scale, which asks respondents to express their opinions (either agree or disagree) on the items presented.

The questionnaire was developed by the research team based on an adaptation and reconstruction of appropriate items and constructs from three measurement tools, namely, the Rathus Assertiveness Schedule (RAS) [40,41], Ryff's Psychological Well-Being Scales [42,43], and Social Skills Inventory [44,45] as well as past study questionnaires that had been

conducted with an almost similar purpose. Before the actual distribution process of the needs analysis among the respondents, the set of questionnaires were verified and validated by six referral experts in psychology and counselling, module development, and academics. Suggestions for the improvement and modification of the questionnaires from the expert panels were applied before conducting the study on the actual target. Experts also provided written reviews that the module was suitable for an intervention program to achieve the set objectives.

3. Results

Quantitative data were analyzed using descriptive analysis. Descriptive statistical analysis was used to describe the demographic information of the respondents covering gender, race, field of study, socioeconomic status (SSE), and place of residence, the needs of a psychosocial module developed based on demographics, the construct required in the design of a psychosocial module and the support that has been received and desired by the respondents.

3.1. Demographic of Respondents

A total of 286 students enrolled at three public universities in Malaysia participated in this study. Of these participants, 161 were female and 125 male. In terms of ethnicity, most of the respondents were Malay (194), followed by Indians (51), and Chinese (41). The participants comprised students from varying fields of studies from which a total of 97 respondents were from the social sciences while the remaining 189 were from the field of science and technology. Regarding their place of residence, 103 of respondents were from rural areas, and most of the respondents were from urban areas (131). Finally, regarding socio-economic status (SSE), a total of 63 respondents came from a high socio-economic status, 132 were from a medium socio-economic status, and the remaining 91 respondents were from a low socioeconomic status family. The demographic information of the respondents is shown in Table 1.

Table 1. Demographic of the respondents (n = 286).

Demographic	Frequency (f)	Percentage (%)
Gender		
Male	125	44
Female	161	56
Race		
Malay	194	68
Chinese	51	18
Indian	41	14
Field of study		
Science and technology	97	34
Social science	189	66
Place of residence		
Urban	172	60
Rural	114	40
Family socioeconomic status		
High	63	22
Medium	132	46
Low	91	32

3.2. The Need for Psychosocial Module Development

Regarding the evaluation of the needs of the development of a psychosocial module, a total of 234 respondents agreed on the need to develop a psychosocial module. Of the total, 136 were female, while 98 were male. For fields of studies, a total of 172 respondents who agreed on the need for a psychosocial module were respondents from the field of social

sciences while the remaining 62 were students from the field of science and technology. Regarding place of residence, 131 of respondents were urban and 103 respondents were rural. Meanwhile, regarding ethnicity, a total of 171 respondents were Malay, followed by Chinese (32), and the remaining 31 were Indian, indicating their agreement on the development of this module. Table 2 shows the evaluation of the need of psychosocial module development based on the demographics as discussed.

Table 2. The needs of psychosocial module development.

Demographic	Evaluation	
	Agree	**Disagree**
Gender		
Male	98	27
Female	136	25
Race		
Malay	171	23
Chinese	32	19
Indian	31	10
Fields of Studies		
Science and Technology	62	35
Social Science	172	17
Place of Residence		
Urban	131	41
Rural	103	11
Family Socioeconomic Status		
High	52	11
Medium	97	35
Low	85	6

3.3. Constructs Required in Designing the Psychosocial Module

In this study, the constructs used in designing and building the psychosocial module were from three instruments: Rathus Assertiveness Schedule (RAS), Del Prette's Social Skills Inventory; and Olaz and Ryff's Psychological Well-Being Scales. Therefore, the questionnaire for data collection used the constructs from these three tools to identify the constructs required by respondents to be included as the content and elements in the psychosocial module. Based on the descriptive analysis that was undertaken, a total of 76% of respondents agreed that the construct the ability to respect should be included in the psychosocial module. In addition, as many as 82% of respondents agreed and required to construct the ability to say 'no'. Meanwhile, 83%, 80%, and 78% of the respondents agreed to construct the ability to express feeling, probability of making mistakes, and interacting and communication should be included in the psychosocial module. Furthermore, 80% of respondents agreed with the construct of coping and self-assertion with risk, 78% agreed with self-assertion in expressing positive effects, 82% on conversation and social confidence, 78% on self -exposure to unknown people and new situations, and 77% on self-control from aggressiveness. Meanwhile, 74% of respondents agreed with autonomy constructs, 82% with environmental mastery, 77% with personal growth, 75% with purpose in life, 76% with positive relationships with others, and 74% agreed that the self-acceptance construct should be included in the psychosocial module. Table 3 shows the constructs that should be included in the phase of designing the psychosocial module as discussed.

Table 3. Constructs required in designing the psychosocial module.

Original Instruments	Constructs	Evaluation			
		Agree		Disagree	
		f	%	f	%
Rathus Assertiveness Schedule (RAS)	The ability to respect	217	76	69	24
	The ability to say 'no'	235	82	51	18
	The ability to express feeling	237	83	49	17
	Probability of making mistakes	229	80	57	20
	Interacting and communicating	223	78	63	23
Social Skill Inventory (SSI-Del-Prette)	Coping and self-assertion with risk	229	80	57	20
	Self-assertion in the expression of positive effect	223	78	63	22
	Conversation and social confidence	235	82	51	18
	Self-exposure to unknown people and new situations	229	80	57	20
	Self-control of aggressiveness	220	77	66	82
Ryff's Psychological Well-Being Scales	Autonomy	217	74	68	26
	Environmental mastery	235	82	51	18
	Personal growth	220	77	66	23
	Purpose in life.	215	75	71	25
	Positive relationship with others	217	76	69	24
	Self-acceptance	212	74	74	26

When evaluating the types of support desired by the respondents in light of their previously identified personal, career, and learning needs (Table 4), most of the respondents indicated a preference for workshops/brief courses (38%) and psychoeducation programs (24%), followed by group guidance (18%), seminars (6%), individual guidance (5%), individual counselling (4%), and group counselling (3%). About six respondents indicated that they did not want access to any type of support for their needs. For the evaluation of the types of support that were previously received by the participants, most of them indicated that they had not received any type of support (39%). About 33 respondents stated they had access to workshops/brief courses, 29 received psychoeducation training, 18 stated that they attended seminars, 15 participated in group guidance, eight participated in individual counseling, seven participated in individual guidance, and four participated in group counselling sessions.

Table 4. Evaluation of the desired support and the received support.

Types of Support	Desired		Received	
	f	%	f	%
Individual counselling	17	4	8	4
Group counselling	8	3	4	1
Group guidance	65	18	15	8
Individual guidance	21	5	7	4
Seminar	24	6	18	10
Workshop/Brief course	142	38	33	18
Psychoeducation program	88	24	29	16
None	6	2	73	39

4. Discussion

A needs analysis is important in identifying information regarding the aspect of module content that will be developed. Researchers conduct a needs analysis to collect information about the context and situational research. According to Aliza and Zamri [46], a module that is developed should take into account the issues that are being faced and the readily available needs, so that the module that is developed can serve the purpose of the target groups. For this study, students from public universities were selected as the target users to conduct the needs analysis for a psychosocial module whereas information on the content of the module was gathered from past studies as well as analyses on the documents pertaining to the issue and the current problems faced by students that need solutions. In general, the study focused on three main issues that have been studied during research regarding the needs for the development of a psychosocial module, the constructs required in the development of a psychosocial module as well as the form of the support received and the support desired by the students.

First, regarding the student's needs of the psychosocial module, the descriptive analysis showed that female respondents were more likely to agree with the development of a psychosocial module compared to male respondents. This may be because women are more affectionate, prone to crying, and at the same time, have lower assertiveness than male respondents. Apart from that, for the field of study, the findings showed that respondents from the social sciences were more likely to agree with the development of a psychosocial module than the respondents from the field of science and technology. I found that students in technical fields showed lower knowledge of emotions and social competence than students in non-technical fields. This may be related to a stereotypical understanding of the technical field, which requires only knowledge and core skills. On the other hand, according to Agnieszka [47], social skills have also been recognized as skills that should be possessed by all university students and need to be developed in other less formal (social) ways, as the emphasis to develop them in learning is not effective. For this reason, the field of study should also be considered in planning appropriate programs to develop the social skills of students.

However, for SSE, study has shown that most respondents that agreed with the developed psychosocial module were those with low SSE. Based on the research conducted by Namrata and Renu [48], students from low-income families displayed an overall pattern of psychological well-being, which was almost similar to students from middle income families. Besides, students from low-income families also displayed a low level in the autonomous dimension, but a high level in life purpose and self-acceptance. In contrast, for students from high-income families, they controlled all dimensions, which were environmental mastery, positive relationship with others, purpose in life, self-acceptance, and the overall PWB compared to students from middle income families. Therefore, considering all

these effects, it can be concluded that the students' socioeconomic status also contributed to better psychological well-being among the new undergraduates.

As for the constructs required in module development, most respondents agreed that the constructs regarding the ability to respect, the ability to say 'no, the ability to express feeling, probability of making mistakes, and interacting and communication should be included in the psychosocial module to increase their assertiveness. In this context, Malaysian ethics are quite different from Western ethics [49,50]. Previous studies have shown that assertiveness is less commended in Malaysia, which has a collectivist culture. The collectivist culture can be seen as reflected in the Malaysian students' preference for group success, teamwork, and interdependence as opposed to individual success, competition, and independence [51]. Therefore, these differences influence the level of assertiveness among the students in Malaysia, in which it was found that Asian students had a lower level of assertiveness compared to students in the West. In the field of education, the assertiveness skill is important to help students to express their needs and positive or negative feelings truthfully and forcefully without any worry or intimidation. Therefore, recognized constructs are required to develop this skill of assertiveness.

Meanwhile, more than 70% of students agreed that the constructs of autonomy, environmental mastery, personal growth, purpose in life, positive relationships with others, and self-acceptance should be included in the development of psychosocial modules. A study conducted by Cabrera, Daya, and Echague [52] to identify factors that affected the psychological well-being of students found that most university students had low levels of autonomy, environmental mastery and self-acceptance, and younger students had low environmental mastery. Therefore, by understanding the psychological and social constructs that influence PWB, then it can be used as content in the development of student interventions.

Finally, regarding the results of the desired support and the received support for their needs, most respondents indicated having participated in activities or programs in group and educational settings such as workshops/courses, psychoeducation programs, and seminars, which is compatible with other studies. The findings of this study also coincides with the study conducted by Hammad [53]; This is also in line with the views of Rodriguez et al. [54], who claimed that students preferred psychoeducation programs because of the student-centered learning, interactive lessons, and team dynamics. However, only a small number of respondents received and had access to psychoeducation programs. Meanwhile, there are a number of students who state that they do not want any kind of support for their needs. Therefore, it is important to distinguish between the participants who consider that they do not have any needs in particular and those who, even after finding out some issues, deny any type of support due to fear of the stigma or not knowing how or where to find it. Consequently, this finding shows the need to design and build a psychosocial module based on psychoeducation and can further increase the probability that these services will become greater in demand since they are seen by students as a resource of high-quality, is effective, and convenient.

5. Conclusions

In conclusion, it is relevant to investigate and explore the students' needs before developing and implementing interventions so that students that need support have the opportunity to obtain it. With the realization of this study and considering any disparity between the results obtained and the literature mentioned, psychosocial modules based on psychoeducation should be developed and further contribute to responses that are technically and scientifically appropriate to the identified needs. From this juncture, further studies can be conducted by combining both qualitative and quantitative designs. In a nutshell, it is hoped that the findings of this study can play a role as a reference for future research in strengthening assertiveness, social skills, and psychological well-being for university students.

Author Contributions: Conceptualization, S.A.; and H.I.S., investigation, H.I.S.; data curation, H.I.S.; methodology, S.A., H.I.S. and A.Y.A.B.; funding acquisition, S.A.; supervision, S.A. and M.I.M.; project administration, A.Y.A.B.; validation, M.I.M.; software, M.H.K.; writing—original draft, H.I.S. and S.A.; writing—review & editing, A.Y.A.B. and M.H.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received funding from the Faculty of Education, Universiti Kebangsaan Malaysia (Code: GG-2020-022).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The researchers would like to express their gratitude to all participants involved in this study.

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

References

1. Kotera, Y.; Ting, S.H.; Neary, S. Mental health of Malaysian university students: UK comparison, and relationship between negative mental health attitudes, self-compassion, and resilience. *High. Educ.* **2021**, *81*, 403–419. [CrossRef]
2. WHO. World Health Organization: Substantial Investment Needed to Avert Mental Health Crisis. 2020. Available online: https://www.who.int/news-room/detail/14-05-2020-substantial-investment-needed-to-avert-mental-health-crisis (accessed on 9 June 2020).
3. Auerbach, R.P.; Alonso, J.; Axinn, W.G.; Cuijpers, P.; Ebert, D.D.; Green, J.G.; Bruffaerts, R. Mental disorders among college students in the World Health Organization world mental health surveys. *Psychol. Med.* **2016**, *46*, 2955–2970. [CrossRef] [PubMed]
4. Auerbach, R.P.; Mortier, P.; Bruffaerts, R.; Alonso, J.; Benjet, C.; Cuijpers, P.; on behalf of the WHO WMH-ICS Collaborators. Mental disorder risk profiles in the World Health Organization World Mental Health Surveys International College Student Project. *Int. J. Methods Psychiatr. Res.* **2019**, *28*, e1758. [CrossRef] [PubMed]
5. Cuijpers, P.; Auerbach, R.P.; Benjet, C.; Bruffaerts, R.; Ebert, D.D.; Karyotaki, E.; Kessler, R.C. The WHO World Mental Health Surveys International College Student (WMH-ICS) initiative: A systematic support system for mental health problems in college students. *Int. J. Methods Psychiatr. Res.* **2018**, *28*, e1762. [CrossRef]
6. Lipson, S.K.; Lattie, E.G.; Eisenberg, D. Increased rates of mental health service utilization by US college students: 10-year population-level trends (2007–2017). *Psychiatr. Serv.* **2019**, *70*, 60–63. [CrossRef]
7. Evans, T.M.; Bira, L.; Gastelum, J.B.; Weiss, L.T.; Vanderford, N.L. Evidence for a mental health crisis in graduate education. *Nat. Biotechnol.* **2018**, *36*, 282–284. [CrossRef]
8. Kunjiapu, S.; Kunasegaran, M. Psychological wellbeing among undergraduates of Higher Learning Institutions in Malaysia. *Psychol. Educ. J.* **2021**, *58*, 9179–9187.
9. Manap, R.; Hamid, A.S.; Ghani, A.M. Depression, Anxiety and Stress Among Undergraduate Students. *J. Soc. Sci. Humanit.* **2019**, *16*, 1–7.
10. Beigi, Z.; Bahar, S. The relationship between psychological well-being and academic achievement on students. *Traboja Soc. Glob.* **2017**, *12*, 830–839.
11. Ryff, C. Well-being with soul: Science in pursuit of human potential. *Perspect. Psychol. Sci.* **2018**, *13*, 242–248. [CrossRef] [PubMed]
12. Erpiana, A.; Fourianalistyawati, E. Peranan trait mindfulness terhadap psychological well-being pada dewasa awal. *Jurnal Ilmiah Psikologi* **2018**, *5*, 67–82.
13. Suvera, P. Psychological well-being: A comparative study of tribal and non-tribal college students. *Indian J. Health Wellbeing* **2013**, *4*, 1643–1647.
14. MacDonald, H.Z.; Price, J.L. The role of emotion regulation in the relationship between empathy and internalizing symptoms in college students. *Ment. Health Prev.* **2019**, *13*, 43–49. [CrossRef]
15. Shu, T.M.; Lam, S. Is it always good to provide positive feedback to students? The moderating effects of culture and regulatory focus. *Learn. Individ. Differ.* **2016**, *49*, 171–177. [CrossRef]
16. Morales-Rodríguez, F.M.; Espigares-López, I.; Brown, T.; Pérez-Mármol, J.M. The relationship between psychological well-being and psychosocial factors in university students. *Int. J. Environ. Res. Public Health* **2020**, *17*, 4778. [CrossRef]
17. Parmaksiz, I. Assertiveness as the predictor of adjustment to university life amongst university Students. *Int. J. Instr.* **2019**, *12*, 131–148. [CrossRef]
18. Siang-Ang, C. Assessment of Social Competence: Exploring the Dimensions of Young Adult Social Competence. 2016. Available online: http://spaj.ukm.my/ppppm/jpm/article/view/210 (accessed on 22 September 2021).
19. Sieng, L.W.; Yussof, I. Comparative study of Malaysia human capital with selected ASEAN and Developed Countries: A fuzzy TOPSIS method. *Geogr.-Malays. J. Soc. Space* **2017**, *11*, 11–22.

20. Khairani, A.Z.; Ahmad, N.S.; Sutatminingsih, R. Assessing social-psychological functioning between Malaysian and Indonesian undergraduates: A Rasch model analysis. *J. Appl. Res. High. Educ.* **2021**, *13*, 43–50. [CrossRef]
21. Yoke, T.C.; Ngang, T.K. A study on soft skill development among final year Diploma in Business Studies students. *Malays. Online J. Educ. Manag.* **2017**, *3*, 32–50.
22. Ritter, S.M.; Mostert, N. Enhancement of creative thinking skills using a cognitive-based creativity training. *J. Cogn. Enhanc.* **2016**, *1*, 243–253. [CrossRef]
23. Trigueros, R.; Sanchez-Sanchez, E.; Mercader, I.; Aguilar-Parra, J.M.; López-Liria, R.; Morales-Gázquez, M.J.; Fernández-Campoy, J.M.; Rocamora, P. Relationship between Emotional Intelligence, Social Skills and Peer Harassment. A Study with High School Students. *Int. J. Environ. Res. Public Health* **2020**, *17*, 4208. [CrossRef]
24. Batdi, V.; Elaldi, S. Effects of drama method on social communication skills: A comparative analysis. *Int. J. Res. Educ. Sci. (IJRES)* **2020**, *6*, 435–457. [CrossRef]
25. Del Prette, Z.A.P.; Del Prette, A.; Peixoto, E.M. Social Skill Inventory-2 Del-Prette: Expanding and updating psychometric properties. *Estud. Psicol.* **2021**, *38*, e190124. [CrossRef]
26. Del Prette, Z.A.P.; Del Prette, A. Studies on social skills and social competence in Brazil: A history in construction. In *Psychology in Brazil: Scientists Making a Difference*; Koller, S.H., Ed.; Springer: Cham, Germany, 2019. [CrossRef]
27. Hamidah. Keberkesanan Kaunseling Kelompok Tingkah Laku Kognitif (KK-CBT) dan Latihan Kemahiran Sosial (LKS) Untuk Mengurangkan Kemurungan, Meningkatkan Kesejahteraan Psikologi dan Kemahiran Sosial Bagi Remaja Yang Mengalami Kemurungan. Ph.D. Thesis, Universiti Kebangsaan Malaysia, Bangi, Malaysia, 2011.
28. Zazula, R.; Appenzeller, S. Evaluation of a psychoeducational intervention in the development of social and communication skills for incoming medical students at a bilingual university. *Trends Psychol.* **2019**, *27*, 749–762. [CrossRef]
29. Howard, M.C.; Gutworth, M.B. A meta-analysis of virtual reality training programs for social skill development. *Comput. Educ.* **2020**, *144*, 103707. [CrossRef]
30. Cabrera, G.; Daya, H.; Echague, N. Psychological well-being of college students: Validation of the personal-social responsibility and wellness module. *J. Educ. Hum. Resour. Dev.* **2020**, *8*, 59–70.
31. Ishak, Z.A. Kesan Modul Bimbingan Kaunseling Kelompok Penyesuaian Remaja Terhadap Motivasi, Ketegasan, Empati, Harga diri dan Kecerdasan Emosi Dalam Kalangan Pelajar Salah Laku Disiplin. Ph.D. Thesis, Universiti Pendidikan Sultan Idris, Perak, Malaysia, 2016.
32. Güler, M. The effect of psychoeducation program on burnout levels of university students. *OPUS Uluslararası Toplum Araştırmaları Dergisi* **2020**, *16*, 5524–5539. [CrossRef]
33. Hussin, U.R.; Mahmud, Z.; Karim, D.N.F.M. Psychoeducation group counselling for emotional intelligence among secondary school female students. *J. Couns. Educ. Soc.* **2020**, *1*, 53–57.
34. Kho Ke Li, V.; Mahmud, M.I. A case study of psychoeducation group on career choice for secondary school. *Int. J. Arts Humanit. Soc. Sci. Stud.* **2020**, *5*, 23–28.
35. Giovazolias, T.; Leontopoulou, S.; Triliva, S. Assessment of Greek university students' counselling needs and attitudes: An exploratory study. *Int. J. Adv. Couns.* **2010**, *32*, 101–116. [CrossRef]
36. Pinto, J.C.; Faria, L.; Pinto, H.R.; do Céu Taveira, M. Identification of psychological intervention needs: A pilot study in Portuguese higher education. *Psicol. USP* **2016**, *27*, 459–472. [CrossRef]
37. Pheng, T.L.; Marzuki, W.; Noohayati. Stigma and Attitudes toward Seeking Counseling among Undergraduate Students. *Glob. J. Al-Thafaqah* **2019**, *93*, 93–105.
38. Chua, Y.P. *Mastering Research Methods*, 2nd ed.; McGraw-Hill: Kuala Lumpur, Malaysia, 2016.
39. Creswell, J.W. *Research Design: A Qualitative, Quantitative, and Mixed Method Approaches*, 4th ed.; SAGE Publication Inc.: London, UK, 2013.
40. Rathus, R.A. A 30-item schedule for assessing assertive behavior. *Behav. Ther.* **1973**, *4*, 398–406. [CrossRef]
41. Omar, H. Hubungan Antara Tingkah Laku Asertif Dengan Komitmen Organisasi Dalam Kalangan Staf Sokongan Universiti Tun Hussein Onn Malaysia (UTHM). 2015. Available online: https://core.ac.uk/reader/42918800 (accessed on 22 September 2021).
42. Abbott, R.A.; Ploubidis, G.B.; Huppert, F.A.; Kuh, D.; Wadsworth, M.E.J.; Croudace, T.J. Psychometric evaluation and predictive validity of Ryff's psychological well-being items in a UK birth color sample of women. *Health Qual. Life Outcomes* **2006**, *4*, 76. [CrossRef]
43. Ryff, C.D. Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *J. Personal. Soc. Psychol.* **1989**, *57*, 1069–1081. [CrossRef]
44. Del Prette, Z.A.P.; Del Prette, A. Social Skills Inventory (SSI-Del-Prette): Characteristics and studies in Brazil. In *Social Anxiety Disorders: From Research to Practice*; Osório, F.L., Ed.; Nova Science: New York, NY, USA, 2013; pp. 49–62.
45. Olaz, F.O.; Medrano, L.; Greco, M.E.; Del Prette, Z.A.P. Argentinean adaptation of the Social Skills Inventory IHS-Del-Prette. *Span. J. Psychol.* **2009**, *12*, 756–766. [CrossRef]
46. Aliza, A.; Zamri, M. Analisis keperluan terhadap pengguna sasaran modul pendekatan berasaskan bermain bagi pengajaran dan pembelajaran kemahiran bahasa kanak-kanak prasekolah. *Jurnal Kurikulum Pengajaran Asia Pasifik* **2017**, *3*, 1–8.
47. Szewczyk-Zakrzewska, A. Social skills and understanding emotions by students at technical and non-technical universities. *World Trans. Eng. Technol. Educ.* **2019**, *17*, 110–114.
48. Punia, N.; Malaviya, R. Psychological Well Being of First Year College Students. *Indian J. Educ. Stud. Interdiscip. J.* **2015**, *2*, 60–68.

49. Afdal, A.; Wibowo, T.; Alfia, V.; Maysitoh, M. How assertive behavior differences among Minang, Malay, Batak and Kerinci student's culture? *Couns.-Educ. Int. J. Couns. Educ.* **2019**, *4*, 134–139. [CrossRef]
50. Kuswarianandhika, Y.F. *Hubungan Antara Dimensi Budaya Collectivism Dengan Perilaku Asertif Pada Mahasiswi Suku Jawa*; Universitas Mercu Buana Yogyakarta: Yogyakarta, Indonesia, 2017.
51. Imai, T. What helps international students disclose themselves and be assertive to host nationals for their cultural adjustment? Focusing on language ability and length of stay. *Jpn. J. Commun. Stud.* **2020**, *49*, 5–23.
52. Cabrera, G.A.; Daya, H.D.; Echague, N.P. Factors affecting the psychological well-being of college students in a state university. *Guid. J.* **2019**, *46*, 56–72.
53. Hammad, N.A.I.A.A. A suggested program for teaching social life skills and its effect on developing psychological rigidity and affirmative behavior skills among prince sattam bin abdul-aziz university students. *Multicult. Educ.* **2021**, *7*, 334–346.
54. Rodríguez, F.M.M.; Pérez-Mármol, J.M.; García Pintor, B.; Rodríguez, A.M.M. The link between methods preferred by students and other psychoeducational factors. *Eur. J. Child. Dev. Educ. Psychopathol.* **2017**, *5*, 63–72.

education sciences

MDPI

Article

Participation in the Assessment Processes in Problem-Based Learning: Experiences of the Students of Social Sciences in Lithuania

Jurgita Lenkauskaitė, Remigijus Bubnys *, Erika Masiliauskienė and Daiva Malinauskienė

Institute of Education, Vilnius University Šiauliai Academy, 76351 Šiauliai, Lithuania; jurgita.lenkauskaite@sa.vu.lt (J.L.); erika.masiliauskiene@sa.vu.lt (E.M.); daiva.malinauskiene@sa.vu.lt (D.M.)
* Correspondence: remigijus.bubnys@sa.vu.lt

Abstract: The article explores the idea of change in the higher educational process that is implemented via the problem-based learning strategy. Problem-based learning (PBL) is widely understood as an epistemological transformation in higher education. It is emphasized that the transformation should take place throughout the educational process, and assessment is an inseparable and very important part thereof. The study was aimed at revealing the experiences of participation in the assessment processes in PBL of students attending social science programmes in Lithuania. The empirical study, employing a semi-structured interview method, has shown that the students feel empowered when they have the opportunity to assess the entire educational process and (self-)assess the efforts related to the possibilities to become actively engaged in improvement of the assessment strategy. The study has also shown students' critical approach to the previous experience of assessment in the educational process. Difficulties of student participation in the assessment process in PBL were also identified. They were largely due to the change in the assessment system employed by the teacher and the manifestations of student bias when participating in (self-)assessment.

Keywords: learning; problem-based learning; assessment process; student experiences; higher education

Citation: Lenkauskaitė, J.; Bubnys, R.; Masiliauskienė, E.; Malinauskienė, D. Participation in the Assessment Processes in Problem-Based Learning: Experiences of the Students of Social Sciences in Lithuania. *Educ. Sci.* **2021**, *11*, 678. https://doi.org/10.3390/educsci11110678

Academic Editors: Sandra Fernandes, Marta Abelha and Ana Teresa Oliveira

Received: 15 September 2021
Accepted: 22 October 2021
Published: 24 October 2021

1. Introduction

1.1. Problem-based Learning Definitions and Assurance of Conditions for Response to Shifts in the Education Process

The educational sciences explore new educational strategies that help meet students' expectations towards their studies as they aspire to acquire life-long learning skills for successful adaptation to contemporary society and finding a place in it. The purpose is to promote student cooperation, teamwork, teaching and learning from each other by addressing issues in constantly changing situations [1,2]. Problem-based learning (PBL) is one of the strategies enabling students to develop the mindset of future professionals and gain problem analysis and solving skills. Two PBL analysis approaches could be identified. The first one is dedicated to the description of a specific PBL methodology [3,4]. The second is aimed at the exploration of the philosophical grounds of PBL that reflect the epistemological transformation in higher education [5,6]. If viewed as a clearly and comprehensively described teaching and learning method, PBL seeks to enable students to gain practical experience and enhance problem-solving and self-directed learning skills. PBL is emphasized as an educational strategy and philosophy that is important when responding to the contemporary needs of society. This means that, instead of being viewed in narrow terms, as a method, PBL is treated as a set of notions about teaching and learning by highlighting the broad and flexible opportunities offered by it [1,7,8]. The PBL notion that focuses on the philosophical grounds provides that a consistent student-centered and empowering educational environment needs to be designed. Key features of PBL not only

translate into elements of an innovative educational technique, but also create conditions for a new approach towards educational theory and the practice and emergence of new roles of the participants in the study process.

The curriculum launched at McMaster University Medical School in Canada is considered to be the origins of PBL [9]. This strategy has become widely spread across disciplines and varies in terms of its implementation models. Nonetheless, implementation of PBL is addressed by referring to the key principles identified in the early period of PBL application. The principles are as follows: learning should be constructive, self-directed, collaborative and contextual [10].

The researchers who analyze PBL stress that PBL "is an instructional (and curricular) learner-centered approach that empowers learners to conduct research, integrate theory and practice, and apply knowledge and skills to develop a viable solution to a defined problem" [11] (p. 9). When defining PBL, the focus is placed on the learners as the individuals who are actively involved in the educational process. PBL has been noted to positively affect learners' perceptions of their self-directedness in learning [12]. The learners become actively engaged in the entire PBL process and are important participants in PBL assessment, which is also a part of the process [2]. Certain features of empowerment didactics [13] could be traced to the PBL strategy. These features highlight the conditions which enable the students to become the parties that actively construct and assess their own knowledge. The instructor's role of an assistant in PBL enhances learner empowerment [14]. Empowerment builds on the ideas of learner emancipation and changing power relations in education. New power relations also emerge in teamwork, which is also an important notion in PBL. This suggests the possibilities of implementing the ideals of epistemic diversity and democracy. PBL has been noticed to be more attractive to minorities or other marginalized groups than traditional instruction [15]. PBL dynamics empowers and engages the learners who are usually denied the possibilities to express themselves within a traditional curriculum [16]. Implemented in diversified teams, PBL promotes pluralism of opinions, epistemic equality and solidarity.

The understanding of PBL as a philosophy stretches much further and deeper than the limits defined by the curriculum or subject. This transformation of paradigm has been drawing the attention of the experts for a long time [17]. It is only through transformation of the notion of the traditional curriculum that it may be possible to enable the transformation of the paradigm. In view of the differences within higher education, traditional learning is generally referred to as simulation [18]. The learners do not find the learned content to be personally important, and the "juggling" of formulas and rote knowledge does not guarantee that they will be able to deal with real-life problems. Learners often learn the way to meet teachers' expectations and preferences and receive good scores rather than the subject of instruction [18]. These learners aim to memorize and reproduce the information rather than change their own understanding. Memorization and reproduction at the right time (e.g., during the exam session) are also encouraged by the traditional organization of the educational process and assessment is focused on the amount of the material reproduced. The analysis of weaknesses of traditional teaching and learning shows the need to organize teaching/learning in a way so as to encourage conscious activeness in learners and constructive participation in the educational process as provided for in the PBL strategy.

The modern educational process is characterized by change [19]. The new learning often takes place in diverse, unconventional learning environments. The teacher acts as a facilitator who supports learning. It is important to ensure that the relations between the participants of the educational process meet the needs of the democratic society, and education is implemented as a learner empowering practice.

1.2. Assessment in the Problem-Based Learning Process

Assessment is considered to play a special part in the PBL, and it is important that it takes place regularly in order to assess the knowledge and competences gained throughout

the curriculum [20]. The assessment process in PBL is aimed at encouraging its participants to think critically about the ongoing process and the possibilities offered by its evolution. The assessment that includes the specifics of traditional education is obviously unsuitable for assessment of the PBL process and learning outcomes. It should be in line with the PBL philosophy and show the learning achievements and teaching/learning outcomes under the PBL strategy selected [21–25]. Learners gain a number of new skills in the PBL process [1,26]. The assessment should become and be part of the personalized PBL educational strategy that enables critical thinking in learners [27,28].

Learner-empowering assessment is associated with their active participation in the process, requiring the ability to critically reflect on personal knowledge and authority to perform certain actions in order to enhance their own learning practice. When empowered, the learners have the chance to put forward the assessment topic and its format according to their individual learning style. They may also seek agreement with the teacher about the criteria to be used when assessing and scoring the work. Here, the learners are empowered to take direct actions both individually and as team members when performing self-assessment, judging the assessment and negotiating different approaches than those practiced by the teachers. This means that the students are involved in main aspects of the learner experience, such as assessment strategies. Teachers, in turn, enable the students "to challenge and escape the confines of hegemony" [29] (p. 361). Assessment as empowerment is essentially a democratic process. It has an indisputable value, as it helps the learners develop their activity independently by using the self-assessment and experience reflection forms. This method requires systematic and consistent critical thinking and feedback from all participants of the educational process. Involvement of other participants in the educational process in the process is also a significant domain of the process of empowerment [30,31]. "Empowerment evaluation also requires sensitivity and adaptation to the local setting. It is not dependent upon a predetermined set of technologies. Empowerment evaluation is necessarily a collaborative group activity, not an individual pursuit" [32] (p. 9). Learners are the principal consumers of assessment and they should be enabled to benefit from this position. According to the majority of learners, as they were using diverse self-assessment measures, not only did they succeed in focusing all their efforts towards higher work quality, but they also managed to feel "less anxious" when performing various tasks. Moreover, timely assessment/self-assessment of learning also acts as a disciplinary measure and the possibility to realize whether the learner and the teacher are satisfied with the degree of achievement of the expected learning outcomes [33]. An external evaluator—the teacher—undertakes the role of an assistant or facilitator in case of any difficulties or a guide who helps find appropriate direction for the processes. The evaluator is on equal terms with the learners, a critical friend—not a master or servant [32].

Learners can make autonomous decisions regarding their own learning, including the way in which they would be assessed. Learner empowerment encourages the learners to undertake direct actions both individually and in groups in order to assess their own work, criticize the mode of assessment practiced by the academic world and negotiate on a practice other than the one proposed. To give voice to the learners who feel dissatisfied with the existing assessment processes, it is important to provide conditions for them to regularly participate in the dialog by making sure that their ideas and processes turn into the discourse that recognizes their participation [34]. Three fundamental challenges must be overcome in the entire assessment as an empowerment process [35]: (a) student participation in the assessment of their own learning; (b) feedforward, which focuses on the delivery of information about the results of assessment that can be used in a proactive way; (c) producing high quality assessment tasks. These three challenges can be broken down into individual principles that, when operationalized by university tutors, enable them to introduce innovative practice or procedures that influence the way that both staff and students experience assessment. The innovative proposals based on the three challenges mentioned above may be implemented by using appropriate technologies in a way so as to empower the learners' learning process in the academic setting. Simultaneously, the

technology and extension of the respective powers may help the learners to cultivate the skills and competences that determine strategic thinking beyond the academic context (personal and professional) as well.

The theorists and practitioners of PBL have been exploring specific and particular assessment strategies, techniques and instruments for the assessment to reflect the essence of PBL [21,24,25,36,37]. PBL offers a diversity of forms and instruments to be used in assessment. The majority of the authors who deal with the specifics of assessment in PBL emphasize the need to properly blend the summative and formative assessment [38–40]. Formative assessment is a part of developing or ongoing teaching/learning, where the participants in the educational process express their remarks and feedback to each other. It includes feedback to the participants of the educational process with the aim of enhancing teaching/learning and improving the curriculum and process. The main aim of the formative assessment is to develop future learning and encourage the students to reflect, thereby empowering them to review what they have already learned and understood, and receive feedback from the teacher on how to improve their achievements [40,41]. This kind of assessment enables the learners to track their own progress and take note of the competences and abilities that still need to be acquired, establish the link between learning and assessment and adjust the learning style where needed according to the curriculum.

Application of the strategies of assessment by other participants of the educational process, not just by the teacher, is proposed to ensure development of learners' competences and democratic distribution of the powers among the participants of the educational process [20]. It has also been noted that various assessment techniques and instruments [42,43], such as work folders, simulation, essay, reflection journals, reports, etc., could provide substantial input in PBL. Assessment based on specific instruments (e.g., learning journals) may be biased. The bias could be overcome by including groups of different evaluators [27]. A combination of assessment methods may help to compensate for the weaknesses of a single method and provide a better picture of the PBL situation. Various assessment strategies and instruments help the learners assess how well they learn rather than how much they have learned [23]. Assessment in PBL is focused on the entire educational process rather than just an outcome as a final achievement.

In general, assessment could be claimed to be one of the most popular forms of power demonstration and the act of demonstration of (non)compliance with a norm. The traditional assessment practice is often disempowering to the learners, and the learners remain passive objects of the assessment. PBL avoids demonstrating the hegemonic power via exams [44]. To the contrary, formative assessment promotes empowerment of the participants in the educational process. In PBL, assessment as a process involves various participants, assessment forms and methods to avoid the demonstration of hegemonic power relations.

In a number of scientific sources, PBL is construed as an innovative approach to learning medicine, where it has been applied and analyzed the most extensively [9,45–47]. The application of PBL to other science areas, e.g., social sciences, would provide a broader perspective and see the potential for construction of social knowledge in the educational process. In Lithuania, PBL has not yet established strong roots as a novel approach. Despite the potential possibilities for transformation in education offered by it, PBL also challenges the traditional (self-)education theory and practice. To highlight the benefits of PBL to learners and identify the aspects that require improvement, the learners' experiences in PBL need to be analyzed and explored. In view of the above, the research aim was to reveal the experiences of participation of the students of social science programmes in Lithuania in the assessment processes in PBL.

2. Research Methodology

2.1. Context

The research pursued the aim of analyzing the students' experiences of participation in the assessment processes in PBL. The problem question of the research was: how do

students' experiences manifest themselves in terms of participation in the assessment processes in PBL?

The study presented in the article was closely related to the PBL application project conducted at two universities in Lithuania. PBL assumes active student cooperation in teams when analyzing and solving problems. Assessment was considered to be an important part of the entire educational process as provided for by the methodological principles of PBL. The project was mainly aimed at developing the methodological and informational support for problem-based teaching and learning in order to assure specialists' competitiveness on the labour market. The universities were selected for the project (as well as the research presented in the article) according to the degree of their focus on social sciences. The goal behind the application of PBL on the institutional scale was to improve the educational process in social sciences and to encourage the students to become actively engaged in real-life, relevant problem solving, develop their social construction of knowledge and collaboration competences.

The research covered diverse experiences, as the schools of higher education were chosen in a way so as to make sure that there were certain differences between them. One of the purposes was to ensure diversity of the research data. The first university was the state school of higher education promoting change in the country and, in particular, one of Lithuania's regions. The second university was the Liberal Arts and Sciences school of higher education. The outreach of the second university was not specific to any particular region of Lithuania. The University rather emphasized the aspiration to build Lithuania's future and contribute to the development of global culture and science.

Eight study programmes in the social science area, comprising 70 study subjects under the programmes delivered by 60 teachers and attended by more than 800 students, were upgraded by integrating the PBL strategy during the project period. The teachers participated in 6 workshops (duration: 8 h each) dedicated to presentation of the rationale behind PBL and guidelines on its integration into the educational process. Previously, PBL had been employed in Lithuania on the institutional scale in medicine learning only, and project activities were largely centered on learning from the best foreign practice in PBL. Teams of teachers preparing to use PBL would visit Maastricht and Sheffield Universities to learn about the PBL application at those universities. Upon their return, they would share the practice with the colleagues who were preparing to apply PBL in Lithuania.

Following the first semester of PBL application, a questionnaire survey was conducted among the students and interviews with the teachers were held. They were aimed at providing information about the students and teachers' experience in relation to the PBL-upgraded subjects. All teachers were found to have selected a partial, hybrid PBL model for application. This implied the combining of conventional classes and PBL activity. The study subjects, during which the students would become engaged in PBL activities most actively, were identified. PBL was not fragmented, but was rather implemented in these subjects consistently throughout the semester (6 months), covering more than 50% of the educational process. During that period, the students were working in teams to analyze and address real-life, authentic, ill-structured problems, e.g., they conducted a study identifying and assessing the experience of school education and support received by the persons with special educational needs; developed projects on the improvement of the website of University faculties by implementing the knowledge gained during the public communication-related subjects. The educational process was centered on students' self-directed learning, and the teachers were acting as facilitators. Collaborative assessment where both the teacher and the students take active part was emphasized during PBL application. Following the approach towards PBL as a flexible strategy, the assessment process and participants' contribution into it were not subject to general regulations, and could therefore vary from subject to subject.

The students who had attended the study subjects involving the most active use of the PBL activities were invited to participate in a semi-structured interview for the purpose of the research presented in the article. The article presents the part of the research that

addresses the PBL process. Students' social construction of knowledge through cooperative learning in the PBL context was presented in the previous publication by the authors [48].

2.2. Participants

The interviewees were selected by mixed purposive sampling, which implies the use of two or more sampling methods in a single study [49]. The students selected using the criterion sampling method participated in the research. They had been attending the subjects with the most active use of PBL activities. The snowball sampling principles were also employed in the study. This technique refers to the approach where the research participants would recommend holding interviews with specific students who had unique and rich experience in the PBL process. The sampling of the research participants helped reveal the diverse assessment experience. The research involved 31 students from different PBL teams. The participants represented two cities of Lithuania that were homes to the universities which applied PBL in the area of social sciences.

The research participants attended the following Bachelor study programmes: Public Administration, Business Administration, Economics, Public Communication, Special Education, Primary Education and Education. The first university was represented by 15 students, the second—by 16 students. The gender distribution was the following: 3 males and 28 females. The unequal gender distribution of the respondents demonstrated that social science study programmes (in particular, in the area of education) were predominantly attended by females (the students under the Education and Primary Education and Pre-School Education were females only). Twelve respondents were second-year, sixteen—third-year and three—fourth-year students. The majority of the research participants fell under the age group of 20–22, and only two respondents were 24 years old.

2.3. Research Methods

The semi-structured interview method was used to learn the research participants' perceptions about the world, their construction of meanings and their interpretation of their own experience gained in the PBL process. The analysis was focused on the way of construction of the new knowledge as the horizons of meanings were merging. The way of blending of the research participants' previous understanding with new experiences and the meaning it acquired for the subjects of cognition were identified.

With a view towards the research aim, the following main questions of the semi-structured interview were developed: what was your experience of performing the assessment in the PBL process? How was this experience different from your previous assessment experience in the educational process? What is your evaluation of your experience in performing the assessment in the PBL process? Why?

The interpretative phenomenological analysis applied in the investigation of the interview texts was focused not only on the importance of the meanings emphasized by the interviewees, but also on the researchers' interpretations based on their experience and knowledge [50]. Phenomenology-based research largely focuses on the experience [51]. The conducted research, therefore, enabled the researchers to recognize the students' authentic "voices" as they were sharing their experience during the interviews.

Following the idea of "hermeneutic circle", the researcher was moving in the between part of the text and the whole of the text [51]. This showed the key meanings embedded in the text. The following stages of analysis were completed in order to perform the thematic analysis: attaining familiarity with the data through open-minded reading; search for meanings and themes; organization of the themes into a meaningful wholeness [52].

The internal validity [53] of the research results was pursued by the efforts of the researcher to participate in the research process directly and of the team of researchers who are the authors of the present article. The process of data analysis by collaboration helped reflect on the key experience presented by the research participants more accurately and clearly, and to reach common understanding and empirical and conceptual agreement. Confirmability of the qualitative research results was assured by providing the examples

of research participants' experience and ideas for illustrative purposes [54]. The results of the studies conducted by various authors were used for better understanding of the data of the presented research and preparation for the discussion.

2.4. Research Ethics

The research followed the fundamental principles of ethics: voluntary participation, confidentiality, respect, etc. [55,56]. Research participants' verbal consent to participate was obtained during planning for the qualitative research. A convenient time and place for the interviews were arranged with the research participants. All research participants were acquainted with the purpose of the research, its benefits, participants' rights and the possibility to refuse to participate in the research. Prior to the interview, research participants' verbal consent to record the interview was obtained.

The confidentiality of the information obtained was maintained by restricting access to the information provided by the research participants to the researchers only. To maintain the confidentiality of the respondents' identities, the researchers replaced the respondents' names with pseudonyms where the results were presented.

The research followed the principle of respect to the participants. Their opinions and experiences were deemed to be meaningful and respectable and were accepted as such. The benefits of the research were sought to outweigh any potential damage to its participants. Given that the research participants shared a fairly new experience, and this occasionally resulted in tensions and ambiguities, the aim was to listen to the research participants and give them informal feedback.

3. Results

3.1. Student Empowering Assessment Experiences

Students' opinion about and experience in the educational process reflect their authentic knowledge about the problem analyzed. The analysis of the respondents' opinions on the specifics of assessment in PBL has shown change in the power relations among all the participants of the educational process. This included the diminishing role of the teacher and, in particular, students' active participation in the assessment process. The experiences of the teacher–student collaboration and students' active involvement in construction of the assessment process in PBL became a form of student empowerment.

Analysis of assessment as an expression of democratic relations that enables avoiding the power of hegemonic authority primarily focuses on the learners' ability to reflect on their own knowledge, efforts and learning experiences critically and openly. In PBL, teachers usually do not have the possibility to monitor the entire educational process and can see only the final result. This means that the students have positive views towards the opportunity to assess themselves:

> Of course, the teacher can assess objectively, but he doesn't know how much input I made by myself. What if I have the stage fright or something like that, or I might not be able to express myself in the way that I have expressed myself when collecting the materials and so on. So, it's really good that we could perform the assessment ourselves. (Vaiva)

The students have noted that self-assessment empowers them to reflect on their own strengths that are not necessarily evident during the final exam. At the same time, other unnoticed abilities that the students believe may play an important role in the assessment become more prominent. The research results also show a change in the teacher's role and power in the processes. The teacher is no longer and cannot be any longer the only and main evaluator, because his/her participation in the educational process is not active enough to be able to assess the students objectively. In the assessment by the teacher, objectivity is more related to unified assessment criteria applicable to all students. These criteria may not cover students', as different and special individuals, successful activity. Hence, the students gain greater confidence, develop a more positive attitude and are motivated to be more active when they have the possibility to perform the self-assessment.

Another important aspect of the assessment is the opportunity to self-assess their own efforts. The analyzed research results have suggested the popular opinion among the students that the opportunity to perform self-assessment of the efforts was empowering:

> In fact, we gave ourselves the assessment score of 10 just for the efforts <...> There were errors, but it's good. Because the most annoying is when you see that others have not even tried at all. But you try very hard, and you failed, and your assessment score is even lower. (Morta)

The opportunity to be evaluated for the efforts empowers the students to give more effort and assess the errors from a positive perspective, as a path of improvement. *However, the student experienced negative emotions as she was observing the situations where high assessment scores were the result of random success rather than individual's efforts*. The empowering assessment during PBL prevents this kind of manifestation of failure.

The research results have also shown the perspective of adequate assessment. It is associated with each team member's efforts that depend on individual capacities rather than with equal contribution by each team member into the knowledge construction process:

> Someone might have found this a very hard task. Given his capacities, this may have been a rocket science to him, and it would be hard to think of anything he could have done there. For me, it's a different story. In general, I think it's difficult to tell who did more and who did less, in particular, in team work. Because when you are in a team, you are responsible for the entire team. I don't think that it would count as team work, if they started estimating who did more and who did less. I think this is up to the team members on how they are going to coordinate this and what they feel their responsibility is. (Rasa)

The students emphasize that it is necessary to associate assessment with the team members' heterogeneity rather than their equality or benchmarking of individual traits. It also substantiates the opinion that assessment is largely determined or should be determined by the joint team effort and perceived personal responsibility for one's own contribution to the team work.

The students also note that PBL and its assessment specifics empowers and motivates the students to participate throughout the educational process, as it is not a single success or failure during final assessment that secures a positive assessment score:

> I could personally claim that I like this method, because you'll definitely have a passing score if you attend classes and perform the tasks. If this is not PBL, then you may not have success. Because then the student needs to attend class, he is forced to do that, and when not forced, he is then motivated. (Tomas)

The analysis of the research results has revealed that the students have more favorable views towards the PBL processes than the traditional educational process. PBL motivates the learners to give effort throughout the process. In PBL, the students feel that they have the power and control the situation. Positive assessment of students' achievements depends on their actions rather than external circumstances that they have little control over.

The students had favorable views towards the qualitative aspect of PBL, where the quality of their learning path is emphasized. This kind of assessment is set against the exam, an assessment method widely used in traditional educational processes. Although exams follow clearly defined assessment criteria, the students still consider the exam results to also be determined by unexpected circumstances, one of them being either good or bad luck:

> It shows very clearly who and how studied genuinely. Because an exam, well, it sometimes is a lottery. You are other lucky or not. You might know a lot, but no one can see it. (Morta)

The research results show that, due to its specifics, PBL assessment enables the students to express themselves better, be visible and measure each team member's contribution. The student who associated the exam with the lottery shared that she had negative experience during the exam.

Receiving feedback from all participants of the educational process is another important aspect of assessment in PBL. This refers to each student's personal reflections not only on their own activity, but also on other learners' activity and their individual contribution into the joint work during PBL. According to the students, the reflective assessment by other team members served as a formative component that helped each team member improve and become constructively involved in the activity, pursuing the shared goal:

> You are not angry saying this, but rather say it kindly: "You virtually did not bring anything on the table today, and we had to do everything for you. I even cannot be sure whether I should give you any task or not, because you might be incapable of doing it." And I think that this is the push to that person: "I should probably try harder and do it" <...> We communicated a lot, worked together a lot, and the shared approach to the work actually developed. And the assessment and everything else were highly related to this. (Vilma)

The analyzed research results have revealed the students' abilities to assess insufficient efforts by another member of the team and his/her irresponsible approach to the joint work in a polite, but also open and critical manner. According to the students, constructive assessment of other person's efforts pushes the person to try harder and to undertake the assigned tasks with greater responsibility, which gradually leads to unity within the team and promotes team spirit. The students have emphasized that the assessment that empowers improvement is an inseparable part of the PBL process.

The students have acknowledged that self-assessment and assessment of other team members was very hard, but inevitable in order to reflect on their abilities. The pre-established rules and terms related to assessment helped maintain the unbiasedness:

> At each meeting, we had to assess ourselves and another team member by giving a certain score <...> This was the idea, and if you did not prepare or do anything at all, your score would be very low <...> That kind of assessment was very, very hard. Moreover, you need to assess yourself as well, because sometimes you need to distance yourself a lot to be able to see how much and what you can do. (Vilma)

The students' experiences show the importance of objective and responsible (self-) assessment of oneself and other team members. The empowering assessment places particular importance on deep self-cognition, assessment of own powers and capacities that ensure a person's further growth and development perspective.

The research results have revealed students' contemplations that emerge during assessment of the peers and shown their expanding understanding of the assessment:

> A positive score means that there is a certain amount of work that has been done. If none, then why should it be five? (Gvidas).

> ...We came to the point where we decided that giving good scores to the ones who were not doing anything was not an option. (Augustė)

Students' active participation in construction of the assessment principles becomes evident. At the same time, new power relations that the students attempt to justify by reasoning could be observed to emerge in the educational process. To ensure objectivity of the assessment of the contribution, the students are inclined to eliminate the perspective of the closeness of relations. This helps to promote the components that are important for openness and honesty of those who genuinely study rather than those who simply pretend to study.

Empowerment of students as active subjects of assessment is also evident when addressing the manifestation of collaboration with the teachers in the process of assessment. In most of the PBL cases, assessment was performed both by the teacher and by the students in order to avoid any problems that could emerge, if the right of assessment had been put into the hands of a single person in advance. The research results have revealed positive aspects of the collaborative assessment:

> It is also good for the teacher, because he can make more adequate assessment without participating fully in the learning, work process. I think that he looks not only at the work done, but also at the process itself. So, I believe it's the benefit for the teacher as well. (Monika)

The students have noticed that, in order to assure openness, clarity and transparency of the assessment, they were the ones to have initiated collaboration with the teachers in terms of assessment of the team members. The students shared a situation where the teacher gave the highest score for the team work, but the students reduced the score to the minimum for one of the team members after they had reviewed the PBL process and the team members' contribution and provided their reasoning to the teacher:

> The teacher may naturally have the question: why not give the score of ten to everyone? And it turns out that although the joint work was given the maximum score of ten, it does not go to everyone individually. It was probably for this reason that we approached the teacher and told him that we would not be willing to give the maximum score to one person. (Gvidas)

The research results revealed a situation where the student authority in the educational process was legitimate and reasonable as they were reconstructing the assessment mechanism proposed by the teacher in view of the insufficient contribution by the specific team member. This principled but reasonable and weighed decision by the students about another student is important from the perspective of empowering assessment. It emphasizes that the pre-established set of assessment techniques is adapted according to the circumstances by collaboration of the subjects participating in the assessment rather than followed blindly and without any change.

3.2. Difficulties of Student Participation in the Assessment Process

The research has revealed that the students place particular emphasis on assessment as a certain form of manifestation of the student–teacher relationship. This is common in the discussion of the previous experience of the educational process, the unwanted aspects of which are sought to be eliminated by using the PBL strategy. The exam becomes an important tool of assessment and often enables the teacher to take a convenient position by controlling the student's learning process. The students have noted that, where the teacher holds the assessment in his hands, the students learn for the sake of satisfying the teacher and gaining a passing score rather than in order to develop own knowledge:

> The teacher, as he put it, has own little system, and the students master this system. Each teacher has his own system. So, we master that system and then work according to it. (Gvidas)

The excerpt from the interview reflects another problematic situation, where the teacher's power of assessment becomes more imagined rather than real, if the assessment system remains unchanged for some time. The authentic reports by the students provide evidence that students make use of this situation and adapt to the assessment system used by the teacher, leading to partial abandonment of other important parts of the educational process that might have less influence on the final assessment result. Hence, the teacher's position promotes cursory learning.

Nonetheless, there are diverse issues emerging from the situation in PBL where assessment is not only the teacher's prerogative. Construction of individual knowledge of the participants of the educational process is limited by the factor of bias or subjectivity. This means that, in the educational process, the focus is placed on the empowerment of students to assess their own learning achievements or peers' contribution into the joint work. The students have noticed that assessment was complicated in that it was not easy to maintain unbiasedness in assessment of the contribution by other students in the group or by the student who did not participate actively or at all. Focusing on only the work actually performed is almost impossible in the assessment process, as strong orientation towards personal relations could be observed among the students:

> Problems arise due to the friendships. God save me from showing anger or becoming the one frowned upon. Everything is because of that. If the group, for example, is girls-only or boys-only, and the group members do not have much in common in regular life, but have become familiar because of the project and have started working together, then they would actually assess the work itself rather than the friendship. (Goda)

The research participants have emphasized that close personal relations between the group students become a real obstacle to unbiased (objective) assessment of the contribution by the friends in the group. It has been noted that the teams should be formed for group work so as to include diversity of personal relations. In this case, not only the group work, but also the process of assessment of the contribution, would be more constructive. Hence, the analyzed factor of bias becomes the obstacle that limits both the development of PBL strategy and the assurance of real, rather than simulated, formative assessment.

Students' inclination to communicate with their peers in a closer way than with the teachers is another problematic perspective of the assessment which has become evident in the research. Students start negotiating with their own team members to make sure they receive passing scores for the work they have not done:

> There was a student, she did not attend classes. We still had to assess her. When she saw us, she still asked for a passing score, although she had not done anything. I am not a greedy person, so I gave her five, it was not difficult for me. But in this case there may be problems. (Rokas)

> But she asked to give her a passing score just out of solidarity. Of course, we threatened to give her the score of zero at first <....> Then we decided to give her the score of five. (Gvidas)

The research results have revealed that the peer student was asking the team for a passing score by appealing to the principle of solidarity that is important in a democratic society. However, in this case, solidarity is understood not as a precondition for successful social construction of, but rather as an essentialist concept associated with, a certain status (all of us are students, and we can reach an agreement on the score). It may be noted that the assessment was not formative in that team, while PBL principles emphasize the role of formative assessment. The decision to give a positive score to the peer student who had not contributed to the joint work did not originate from the essence of activity in the PBL team, but was more determined by the mere authority of (self-)assessment granted to students. An assumption could be made, namely, that the students were not prepared sufficiently for objective assessment.

The situation where it was attempted to use assessment as a tool of revenge was also observed in the assessment process in PBL:

> This girl who really did not contribute to our work whatsoever, she was assessed respectively. She was very angry at us and said: "Just wait and see. The next time I will be assessing you all like that". (Saulė)

The assessment manifested itself as power used for destructive purposes rather than following the original purpose of the assessment, i.e., to allow the ones who have become closely familiar with the undergone educational process to perform the assessment, as the assessment may be the most adequate if performed by them.

Unbiased (objective) assessment of other team members and oneself is a very difficult and task which comes with responsibilities. An opinion was expressed, namely, that assessment in PBL should be performed by the teacher, as certain students were misusing the assessment:

> It would be better if the teacher performed the assessment herself during the PBL. Because, for example, the leader wanted to get the score of ten, but he had not deserved it, and there were huge discussions on that that even led to conflicts. It was not right that we had to assess ourselves, the leader had to assess himself. (Tomas)

The students have clearly contextualized the perspective of importance of the teacher's role that may help prevent an inadequate assessment. This may happen when a student—a team leader—demonstrates the authority regardless of his/her actual contribution to the final result, thereby seeking favorable assessment for himself/herself. In this case, the students view the assessment performed by the teacher as the possibility to avoid the necessity to assess others and oneself by reflection, by assuming the responsibility, and to avoid unpleasant discussions or conflicts with the group peers and, in particular, leaders.

However, there were the cases of disappointment even in the assessment performed by collaboration between the students and teachers. For example, the participants were assessed using different weights:

> We expressed and assessed everything by reflecting after each class. However, this was not taken into account in the final version. <...> So, the last time that we saw all our scores, it turned out that they were the same for everyone. (Mantė)

Considerable disappointment among the students was caused by inadequacy of the assessment in PBL and insufficient appreciation of the students' contribution:

> I personally believed that I had put a lot of work into this, much more than the final assessment. I think that the teacher did not consider that. I was a bit dissatisfied. <...> No, the teacher knew that it was only me who had conducted the interview. She could have forgotten about it by the time of that class. I did not try to find out. (Skaistė)

The research results have revealed that the dissatisfaction and problems emerging in the learning process under the new mode were not verbalized or discussed, which could otherwise take the PBL to the ultimate goal of a constructive and empowering dialog among all the participants of the educational process. The context of students' passive position has also been revealed. In that situation, inadequate assessment became the tool of the teacher's authority that caused disappointment in the student and, as a result, the student's passive role in the assessment process.

Another case that caused dissatisfaction with the assessment in PBL in a large group of students was related to changes in the pre-established assessment rules initiated by the teacher without consulting the students. The teacher performed the assessment himself on the basis of the students' answers during the defense of the work prepared in line with problem analysis and solution:

> The defense of the work initially had to account for somewhat ten per cent as promised, but then it turned out that the score given for the defense reflected the entire work. So this made one hundred percent. (Sigita)

The student has noted that the assessment method did not necessarily reflect each student's contribution to the work. Assessment by the teacher that became similar to the traditional exam caused stress to the students. It was more related to reflection of the ability to act in a certain situation of defense of work than reflection of the cognitive and social skills that were expected to manifest themselves during the PBL:

> Some of the people might have been displeased with that, because that public defense could have caused them to trip and added the pressure. The person might have worked a lot, but simply did not know how to defend the work at that moment, express his thoughts immediately. And the score was probably lower than it could have been. (Sigita)

This was a situation where, in an attempt to build the educational process on the PBL principles, the teacher still had the dominant role at the stage of assessment of the learning outcomes. Not only did this return full authority to the teacher, but also made the students doubt the possibilities for their active participation in the educational process. They felt as if they had been deceived, because their empowerment in the assessment of the educational process was only superficial.

The teacher's role was not considered to be as that of an empowering facilitator—a feature essential to the PBL process. According to the research findings, the teacher was also

performing the function of an observer that was linked by the students to the assessment; hence, the teacher's role caused anxiety:

> The teacher's role was more related to observing what we were doing. In fact, the teacher would come closer and listen to us discussing. And, suddenly, all the thoughts would freeze, because everyone knew there was someone coming to assess you. (Diana)

The student's experience was related to the teacher's traditional function to observe and assess the students' work. The students shared an experience that was the opposite of an empowering assessment, when the students were observed and any action on the student's part could have a negative effect on the assessment results. Hence, the students had a mental block which prevented them from thinking freely. In that case, collaboration between the teachers and students did not take place, and assessment was viewed as an obstacle rather than an opportunity to improve their own learning outcomes. The contradiction to the PBL practice and, in this case, the assessment principles applied in PBL has become evident.

There were cases where the students strongly lacked the characteristics of formative assessment and feedback upon completion of the PBL activities. The assessment results would often be provided by the teacher without any discussion, even though the student wanted to have the discussion:

> There was no summary that would have enabled us to realize what we had done wrong and what we should be doing the following time<...> You really want the feedback, if you put so much effort. I'm not talking about appraisal, but some healthy criticism. (Mantė)

The importance of feedback as an important component of enhancement of the educational process is revealed.

Manifestations of passive collaboration between the teachers and students in the assessment process have been noticed:

> Then the teacher writes down everything, the final score, scores for the project works, emails it to us. Then the teacher appoints the time for consultation prior to the exam. And then it's up to us whether we want to come or not, discuss on the assessment, result, and so on. But after that the desire goes away, and you just do not come.

The research findings have shown that the participants of the educational process supported the idea of collaborative assessment. This idea, however, was not implemented in practice sufficiently enough to become an important aspect for PBL improvement in the future.

Having experienced dissatisfaction with the fact that the assessment was not formative or empowering, the students made proposals on possible improvements in the assessment procedure for it to become useful for their learning. It is proposed to make assessment more transparent by enabling the students to reflect on their learning process, and to include the peers in the group for them to participate in the reflection. This would enable the learners to formulate their respective conclusions and learn from the experience:

> And, for example, write the reflections down on sticky notes or something like that, so that the reflections are publicly accessible to everyone. You just don't need to write your name: "I believe that everything was fine in this regard because... I didn't like that because..." And it's important that everyone could see them. I think this would be more useful. Because they are not afraid that they might be identified by others who read their notes. There could be an opinion that may be useful for me, for example. It would be more useful than the way it's done currently, when everything is highly concealed. No one knows what, how, or for what purpose. This could be the reason why the ideas are not considered, the conclusions are not made. (Goda)

Publicity of the assessment could level out the dominant position of the teacher, so that it is closer to that of the students. This would enhance the assessment based on the PBL principles in the educational process where learning takes place by observing peers' works and taking into account their constructive, objective remarks.

4. Discussion and Conclusions

Active and diverse discussions on PBL among the researchers [1,3,4,7,8] reflect the relevance of this educational strategy. The article supports the idea that PBL reflects an epistemological transformation in higher education and enables development of the general view of the new understanding about teaching and learning. This kind of approach suggests that a consistent, student-centered and student-empowering system needs to be designed when organizing a PBL curriculum. PBL principle-based assessment is considered to be an inseparable part of PBL as an innovative educational strategy.

The transition from traditional learning to PBL should be gradual and moderate in terms of pace, as the discourse that supports the traditional educational process is deeply rooted and often develops concealed forms, having a subtle, unnoticeable effect. As soon as the PBL was launched, the coexistence of different discourses could be observed: the efforts to attain innovativeness and collaboration were accompanied by the frequently observed tendency to return to the traditional educational practice. When preparing for the application of PBL, it is important to realize that there may be a strong fear of change at the beginning of implementation. The participants in the educational process had new questions about the organization of regular work. It may be difficult for them to understand and accept the new roles and commitments. They were concerned whether the learning outcomes would be clearly measured and properly assessed [57,58]. It would therefore be reasonable to inquire into the experiences of the participants in the educational process that they have had at the beginning of PBL implementation and assessment in line with its principles. The experiences help reflect on the perspective of implementation of PBL.

For the assessment to comply with the PBL principles, the researchers [15,23,42] have proposed taking into account different aspects related to assessment. Students gain a wealth of skills, in particular, transferable skills. The former enables them to be confident in unconventional situations, realize the importance of life-long learning and gain advantages in the post-modern and constantly changing world [59]. For assessment to be focused on the students' skills that are important at present, to be applied to PBL, the very goal of assessment should be put into scrutiny [23]. In PBL, it becomes more distant from the goal of traditional assessment. The goal of assessment in a PBL process is the empowerment of the students to actively study and provide feedback, assess learning in view of the expected learning outcomes and verify the criteria of assessment and standards. The research results have revealed the manifestation of student-empowering assessment. This means that they participate actively in (re)constructing the assessment aspects that do not satisfy them rather than passively accepting the pre-established assessment criteria. Further, to the contrary, the students find themselves in a confusing situation and imply that it would be safer to stay with the traditional assessment in case of the absence of a constructive dialog among the participants of the educational process, obscurity of the assessment criteria or changes to the assessment criteria after the educational process has started.

In PBL, positive assessment is aimed at contribution to the growth of the participants of the educational process and promotion of construction of their knowledge. Hence, it is important to raise the question of whether the assessment is reasonable and meaningful. Student-empowering assessment that promotes their growth is desired in PBL. During summative assessment, the learners try to emphasize what they know and are able to do. They also try to hide what they do not know and cannot do as much as possible [23]. In the traditional assessment, the errors noticed and corrected by the experts bring the learner back into the defined boundaries of what is normal. In PBL, the errors are, to the contrary, viewed as a positive impetus for learning. The students become used to reflecting on their own experience in the learning process and improve professionally by learning from their

mistakes. Hence, the knowledge is expanded as the students learn from other people's and their own mistakes and experience [60]. The traditional assessment system does not provide a sufficient impetus for students' progress. According to the research findings, in their reflections on their own experience of traditional teaching and learning, the students noted that they would learn by taking into account the teacher's system which they had managed to identify. Traditional learning often makes assessment a fearful experience for the students, and a controlling process [61,62]. The research findings have also revealed students' anxiety, as well as their experience of mind blocks when the teacher was watching them or listening to their discussions, as this could influence the assessment. This kind of reaction experienced by the students probably originates from the experience of traditional assessment that is present in students during PBL as well.

Assessment in PBL is characterized by the diversity of shapes and methods, and special and particular assessment strategies, methods and instruments are explored [21,24,25]. The results of studies on the effectiveness of PBL conducted by various researchers have revealed that the students who learn in a traditional way demonstrate better results from standardized tests, and their results are also better in basic knowledge [63–66]. PBL, on the other hand, guarantees learners better problem-solving skills, skills of acting in practical situations, performing self-directed learning, collecting information and performing self-assessment; it encourages them to focus more on the understanding rather than reproduction of information [67–69]. When analyzing the effectiveness of PBL, the researchers [70] have determined that the more appropriate the assessment method which is selected for assessment of the students' skills, rather than knowledge, the more effective is PBL. While traditional assessment seeks to assess students' knowledge by measurement, this strategy is ineffective and loses its sense in PBL. The skills cultivated during PBL are often aimed at students' life-long learning, and it is not easy to measure and assess them. The research participants have revealed that the new forms of assessment, e.g., self-assessment and assessment of other team members, caused a lot of difficulties; however, they were inevitable for reflection on the skills.

Public aspect becomes very important in PBL in relation to the entire educational process, and not just for its completion, i.e., aggregation of the results. Traditional assessment sets the teaching and learning process apart from its result [1]. Contrary to the traditional educational process, the result of PBL cannot be separated from the process. The research results have shown that the students identified a lot of advantages of the PBL assessment as opposed to their traditional assessment. The students were highly discouraged by exam experience, where luck rather than students' efforts was allegedly an important factor of success or failure. For the researchers, these student experiences of assessment in PBL point to certain possible causes behind this situation. One of them is the teachers' insufficient preparedness for assessment in PBL, when theoretical knowledge does not translate into the educational practice in a fully appropriate manner. At the same time, the researchers have new problem questions emerging from the research findings, namely, why the assessment process has become more distant from the principles of the PBL concept in implementation of PBL. One of the possible explanations is a fairly difficult path from the theory of PBL to its practice [71]. The educational process in PBL motivates the learners to exert effort throughout the process. The students feel empowered, as their assessment score usually depends on their efforts rather than the external circumstances that they have little control of. As well, to the contrary, the students become disappointed by the cases in the PBL where not the entire educational process, but, e.g., defense of the work, was assessed, which reminded them of a traditional exam at the end of the course. The researchers also acknowledge that assessment causes certain signs of disappointment both generally and in PBL. On one hand, this could be related to the ambiguity of student's situation when he/she has to assess peer's and their own work. On the other hand, disappointments may also be related to insufficient effectiveness of collaboration in a group [72].

When analyzing assessment in the educational process, it is also important to consider the question of "Who performs the assessment?". The answer is obvious for the traditional

assessment—assessment is the teacher's prerogative. PBL has a strong effect of increasing students' autonomy. When granting greater responsibility for learning, it is reasonable to grant more responsibility for the assessment of achievement of the expected learning outcomes as well. Whereas PBL activities take place in various settings and students work in small teams, assessment where the teacher has the main role becomes inappropriate. The teacher does not have the capacity to observe and assess the entire PBL process. The teacher acts as an assistant in the assessment process [32]. The research results have also revealed the transformation of the teacher's role in the assessment processes of PBL. It has been observed that, due to the PBL specifics, the teacher no longer is and cannot be the only and main evaluator, and this is considered to be a positive aspect.

The research results have revealed that the students often assessed their own work in the PBL process. Assessment of the team members was also common. Self-assessment and activeness of other participants in the educational process are important attributes of the transformation of education that is characteristic of PBL [30,31,73–77]. The students were happy to be able to (self-)assess not only their knowledge, but also the efforts that depended on personal traits. Other studies have also revealed the advantages of peer assessment, as it enabled the development of such skills as active listening, cooperation, tolerance, self-discipline, self-control, collaboration, negotiation skills, openness, empathy, confidence, teamwork skills, skills of sharing the work load, persistence, creativity, civic consciousness and others [20,78]. The research participants could still trace the complexity of assessment of other team members, because it was not easy for them to distance themselves from the bias (subjectivity) when the students would seek to use the assessment as a tool of revenge or ask other students to assess them for the works they had not actually done.

When considering the question of the participants performing the assessment, it would be reasonable to look deeper into the importance of collaboration between them in the PBL process. Empowering assessment emphasizes the importance of collaboration between the teachers and students, as well as students' active engagement by involving them into development of the assessment strategies [34]. The research findings have revealed that the assessment criteria were occasionally adapted to the circumstances as the participants of the assessment were collaborating. However, the study has revealed certain cases of students' disappointment in PBL: insufficient collaboration with the teachers became evident during the assessment, the assessment scores were not discussed and not all assessment scores provided by all the participants were considered. The issues emerging in the assessment process were emphasized by the students as one of the factors pushing them to return to the traditional educational process. This implies hierarchical relations and is clearly considered to be unchanging from the beginning and does not provide the students with unfounded hopes that they might also be holding the authority in assessment.

The research findings have revealed that the participants in the educational process support the idea of collaborative assessment. Nonetheless, it was implemented insufficiently in practice. The students who participated in the study formulated certain recommendations as they recognized the benefits of PBL and reflected on their experiences in the process. The recommendations were focused on the possibility of changing the behavior of the participants in the educational process for the behavior to be more in line with the PBL principles. The students expressed the desire to make the assessment process more open and transparent, and to receive timely feedback that would lead to improvement.

It is acknowledged that various difficulties are encountered in the implementation of PBL in the educational practice. Teacher's directedness and inappropriate organization of the learning process are common issues [79]. On the other hand, the manifestation of teacher and student's roles becomes the challenge or difficulty in PBL [80]. According to the findings of the conducted research, the teacher is inclined to take the dominant position in the assessment process. Meanwhile, the student makes decisions that may be inadequate in certain situations (for example, the student misjudges or makes unreasonable judgements of peers' work) as he/she fails to realize the importance of their own role. Nonetheless, there are other perspectives as well, when assessment or the assessment score reflects the

actual level of competences attained by the students. Moreover, the assessment result is associated with the possibility to continue improving the existing competences, facilitates assessment of the competence level, etc. [81].

The results of the conducted research help us to better understand the conditions that the PBL is implemented in, the roles performed by the teacher and the student and the PBL concepts related to the assessment process that influence PBL from the students' perspective [79]. In terms of novelty, the research findings have demonstrated various assessment practices in PBL. They, in turn, show new possibilities for student engagement in the assessment processes and for improvement of the teachers and students' competences of assessment [82]. This is a fairly new research area in Lithuania.

Given that the study was conducted at two universities in Lithuania and included the students of social science programmes, it would be reasonable to compare their experiences with the experiences of participation in the assessment processes in PBL of students of study programmes that belong to other science areas. Analysis of diverse experiences would be valuable for improvement of PBL processes and dissemination of the results.

When analyzing the problematics of PBL application at schools of higher education, it is important to note that similar studies place greater focus on and analyze the issues related to problem structure and problem-solving algorithms (the way problems are solved, what decisions are made and why, etc.), as well as the way that the teacher and student's roles are changing during problem solving [79,83,84]. The findings of the conducted study have revealed students' experiences of participation in the assessment process that characterize the situation of PBL learning by the students of social sciences.

The conducted study and the discussion of its results showed the necessity to continue analyzing assessment processes in PBL by identifying the difficulties of implementation of PBL in university studies and the possibilities of improvement. It would be reasonable to develop the research and expand its sample by holding interviews with the teachers. This would provide more comprehensive data about the analyzed problem and not only show the student's position and role in the assessment process, but also enable comparison of the students and teachers' understandings about organization of the assessment in the PBL context. It would be reasonable to perform the comparative analysis of the student and teachers' opinion. The collected results could be used for improvement of the assessment process in PBL. Afterwards, it would be reasonable to resume the study of experiences of the participants of the educational process by identifying the lessons learned in the assessment process in PBL and the new perspective of its improvement.

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

Funding: This research was funded by the Fund for Promotion of Internationalisation of Research and Art of Vilnius University Siauliai Academy.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

1. Savin-Baden, M. *Problem-Based Learning in Higher Education: Untold Stories*; The Society for Research into Higher Education and Open University Press: Buckingham, UK, 2000.
2. Youngerman, E.; Culver, K.C. Problem-Based Learning (PBL): Real-World Applications to Foster (Inter)Disciplinary Learning and Integration. *New Dir. High. Educ.* **2019**, *188*, 23–32. [CrossRef]
3. Azer, S.A.; Peterson, R.; Guerrero, A.P.S.; Edgren, G. Twelve Tips for Constructing Problem-based Learning Cases. *Med Teach.* **2012**, *34*, 361–367. [CrossRef] [PubMed]

4. Spinello, E.F.; Fischbach, R. Problem-Based Learning in Public Health Instruction: A Pilot Study of an Online Simulation as a Problem-Based Learning Approach. *Educ. Health* **2004**, *17*, 365–373. [CrossRef] [PubMed]
5. Ravn, O.; Jensen, A.A. PBL and the Postmodern Condition—Knowledge Production in University Education. *J. Probl. Based Learn. High. Educ.* **2016**, *4*, 38–52.
6. Parton, G.; Bailey, R. Problem-based Learning: A Critical Rationalist Perspective. *Lond. Rev. Educ.* **2008**, *6*, 281–291. [CrossRef]
7. Walton, H.J.; Matthews, M.B. Essentials of Problem-based Learning. *Med. Educ.* **1989**, *23*, 542–558. [CrossRef] [PubMed]
8. Engel, C.E. Not just a method but a way of learning. In *The Challenge of Problem-Based Learning*; Boud, D., Feletti, G.E., Eds.; Kogan Page: London, UK, 1997; pp. 44–59.
9. Barrows, H.S.; Tamblyn, R.M. *Problem-Based Learning: An Approach to Medical Education*; Springer: New York, NY, USA, 1980.
10. De Jong, N.; Krumeich, J.S.M.; Verstegen, D.M.L. To What Extent Can PBL Principles Be Applied in Blended Learning: Lessons Learned from Health Master Programs. *Med. Teach.* **2017**, *39*, 203–211. [CrossRef] [PubMed]
11. Savery, J.R. Overview of Problem-based Learning: Definitions and Distinctions. *Interdiscip. J. Probl. Based Learn.* **2006**, *1*, 9–20. [CrossRef]
12. Golightly, A. The Influence of an Integrated PBL Format on Geography Students' Perceptions of Their Self-directedness in Learning. *J. Geogr. High. Educ.* **2018**, *42*, 460–478. [CrossRef]
13. Arnold, R. Systemtheoretische Grundlagen einer Ermöglichungsdidaktik. In *Ermöglichungsdidaktik—Erwachsenenpädagogische Grundlagen und Erfahrungen*; Arnold, R., Schüssler, I., Eds.; Schneider Verlag Hohengehren: Baltmannsweiler, Germany, 2003; pp. 14–36.
14. Hmelo-Silver, C.E. Problem-based learning: What and how do students learn? *Educ. Psychol. Rev.* **2004**, *16*, 235–266. [CrossRef]
15. Savin-Baden, M.; Major, C.H. *Foundations of Problem-Based Learning*; Open University Press: Maidenhead, UK, 2004.
16. Burch, K. PBL, Politics, and Democracy. In *The Power of Problem-Based Learning*; Duch, B.J., Groh, S.E., Allen, D.E., Eds.; Stylus: Sterling, VA, USA, 2001; pp. 193–207.
17. Barr, R.B.; Tagg, J. From Teaching to Learning—A New Paradigm for Undergraduate Education. *Chang. Mag. High. Learn.* **1995**, *27*, 12–26. [CrossRef]
18. Ramsden, P. *Learning to Teach in Higher Education*; Routledge: London, UK, 1992.
19. Kress, G. Thinking About Meaning and Learning in a World of Instability and Multiplicity. *Pedagog. Int. J.* **2007**, *2*, 19–34. [CrossRef]
20. McDonald, B. *Improving Teaching and Learning Through Assessment: A Problem-Based Learning (PBL) Approach*; Common Ground Publishing: Sydney, Australia, 2010.
21. Van der Vleuten, C.P.M.; Schuwirth, L.W.T. Assessment in the context of problem-based learning. *Adv. Health Sci. Educ.* **2019**, *24*, 903–914. [CrossRef] [PubMed]
22. Blumberg, P. Assessing Students During the Problem-Based Learning (PBL) process. *Med. Sci. Educ.* **2005**, *15*, 92–99.
23. Macdonald, R. Assessment Strategies for Enquiry and Problem-Based Learning. In *Handbook of Enquiry and Problem Based Learning*, 1st ed.; Barrett, T., Mac Labhrainn, I., Fallon, H., Eds.; CELT: Galway, Ireland, 2005; pp. 85–93.
24. Tai, G.X.-L.; Yuen, M.C. Authentic Assessment Strategies in Problem Based Learning. In Proceedings of the Ascilite Singapore 2007—ICT: Providing Choices for Learners and Learning, Singapore, 2–5 December 2007. Available online: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.131.3000&rep=rep1&type=pdf (accessed on 18 April 2021).
25. Henderson, S.; Kinahan, M.; Rossiter, E. Problem-Based Learning as an Authentic Assessment Method. Master's Thesis, Technological University Dublin, Dublin, Ireland, 2018. Available online: https://arrow.tudublin.ie/cgi/viewcontent.cgi?article=1018&context=ltcpgdprp (accessed on 12 August 2021).
26. Hutchings, B.; O'Rourke, K. Medical Studies to Literary Studies: Adapting Paradigms of Problem-based Learning Process for New Disciplines. In *Challenging Research in Problem-Based Learning*; Savin-Baden, M., Wilkie, K., Eds.; Open University Press: Berkshire, UK, 2004; pp. 174–190.
27. Chen, S.E. Problem-based Learning—Educational Tool or Philosophy. In *Problem-Based Learning: Educational Innovations Across Disciplines, A Collection of selected papers, Proceedings of the 2nd Asia-Pacific Conference on Problem-Based Learning, Singapore, 4–7 December 2000*; Tan, O.S., Little, P., Hee, S.Y., Conway, J., Eds.; Temasek Centre for Problem Based Learning: Singapore, 2000; pp. 210–219.
28. Amena, M. The Role of Problem Based Learning in Engaging and Empowering Omani EAP Learners: An Exploratory Study. In Proceedings of the 2nd MEC TESOL Conference, 28 October 2020; pp. 184–196. [CrossRef]
29. Massey, D.; Osborne, D. Empowerment and Assessment: A Dichotomy? *Nurse Educ. Today* **2004**, *24*, 357–362. [CrossRef] [PubMed]
30. Francis, R.A. An Investigation into the Receptivity of Undergraduate Students to Assessment Empowerment. *Assess. Eval. High. Educ.* **2008**, *33*, 547–557. [CrossRef]
31. Bain, J. Integrating Student Voice: Assessment for Empowerment. *Pract. Res. High. Educ.* **2010**, *4*, 14–29.
32. Fetterman, D.M. Reflections on Empowerment Evaluation: Learning from Experience. *Can. J. Program Eval.* **1999**, *14*, 5–37.
33. Stiggins, R.J. Student-involved Classroom Assessment. *Educ. Meas. Issues Pract.* **2001**, *20*, 5–15. [CrossRef]
34. Leach, L.; Neutze, G.; Zepke, N. Assessment and Empowerment: Some Critical Questions. *Assess. Eval. High. Educ.* **2001**, *26*, 293–305. [CrossRef]

35. Rodríguez-Gómez, G.; Ibarra-Sáiz, M.S. Assessment as Learning and Empowerment: Towards Sustainable Learning in Higher Education. In *Sustainable Learning in Higher Education: Developing Competencies for the Global Marketplace, Innovation, Technology, and Knowledge Management*; Peris-Ortiz, M., Merigó Lindahl, J., Eds.; Springer: Cham, Switzerland, 2015.
36. Sternberg, R.J. Interdisciplinary Problem-based Learning: An Alternative to Traditional Majors and Minors. *Lib. Educ.* **2008**, *94*, 12–17.
37. Bowe, B. Assessing Problem-Based Learning: A Case Study of a Physics Problem-Based Learning Course. In *Handbook of Enquiry & Problem Based Learning*; Barrett, T., Mac Labhrainn, I., Fallon, H., Eds.; CELT: Galway, Ireland, 2005; pp. 103–111.
38. Elizondo-Montemayor, L. Formative and Summative Assessment of the Problem-Based Learning Tutorial Session Using a Criterion-Referenced System. *J. Int. Assoc. Med Sci. Educ.* **2004**, *14*, 8–14.
39. Kelley, K.W.; Fowlin, J.M.; Tawfik, A.A.; Anderson, M.C. The Role of Using Formative Assessments in Problem-based Learning: A Health Sciences Education Perspective. *Interdiscip. J. Probl. Based Learn.* **2019**, *13*, 1–12. [CrossRef]
40. Houston, D.; Thompson, J.N. Blending Formative and Summative Assessment in a Capstone Subject: 'It's not Your Tools, it's how You Use them'. *J. Univ. Teach. Learn. Pract.* **2017**, *14*, 1–15.
41. Albanese, M.A.; Hinman, G.L. Types and Design of Assessment in PBL. In *The Wiley Handbook of Problem-Based Learning*; Wiley Handbooks in Education; Moallem, M., Hung, W., Dabbagh, N., Eds.; Wiley-Blackwell: Hoboken, NJ, USA, 2019; pp. 389–409.
42. Macdonald, R.; Savin-Baden, M. *A Briefing on Assessment in Problem-Based Learning*; Assessment Series No. 13; LTSN Generic Centre: York, UK, 2004. Available online: https://s3.eu-west-2.amazonaws.com/assets.creode.advancehe-document-manager/documents/hea/private/id349_a_briefing_on_assessment_in_problembased_learning_1568036654.pdf (accessed on 18 April 2021).
43. Painvin, C.; Neufeld, V.; Norman, G.; Walker, I.; Whelan, G. The 'Triple Jump' Exercise: A Structured Measure of Problem-Solving and Self-Directed Learning. In Proceedings of the 18th Annual Conference on Research in Medical Education, Washington, DC, USA, 18–19 November 1979.
44. Foucault, M. *Discipline and Punish*; Penguin: London, UK, 1987.
45. Dring, J.C. Problem-Based Learning—Experiencing and understanding the prominence during Medical School: Perspective. *Ann. Med. Surg.* **2019**, *47*, 27–28. [CrossRef] [PubMed]
46. Garcia, I.; James, R.W.; Bischof, P.; Baroffio, A. Self-Observation and Peer Feedback as a Faculty Development Approach for Problem-Based Learning Tutors: A Program Evaluation. *Teach. Learn. Med.* **2017**, *29*, 313–325. [CrossRef] [PubMed]
47. Moody, K.; McHugh, M.; Baker, R.; Cohen, H.; Pinto, P.; Deutsch, S.; Santizo, R.O.; Schechter, M.; Fausto, J.; Joo, P. Providing Pediatric Palliative Care Education Using Problem-Based Learning. *J. Palliat. Med.* **2018**, *21*, 22–27. [CrossRef] [PubMed]
48. Lenkauskaitė, J.; Colomer, J.; Bubnys, R. Students' Social Construction of Knowledge through Cooperative Learning. *Sustainability* **2020**, *12*, 9606. [CrossRef]
49. Patton, M.Q. *Qualitative Research & Evaluation Methods*; Sage: Thousand Oaks, CA, USA, 2002.
50. Smith, J.; Flowers, P.; Larkin, M. *Interpretative Phenomenological Analysis: Theory, Method and Research*; Sage: London, UK, 2009.
51. Langdridge, D. *Phenomenological Psychology: Theory, Research and Method*; Pearson Education: Harlow, UK, 2007.
52. Sundler, A.J.; Lindberg, E.; Nilsson, C.; Palmér, L. Qualitative Thematic Analysis Based on Descriptive Phenomenology. *Nurs. Open* **2019**, *6*, 733–739. [CrossRef]
53. Urban, J.B.; van Eeden-Moorefield, B.M. Establishing validity for qualitative studies. In *Designing and Proposing Your Research Project*; Urban, J.B., van Eeden-Moorefield, B.M., Eds.; American Psychological Association: Washington, DC, USA, 2018; pp. 119–127.
54. Denzin, N.K.; Lincoln, Y.S. (Eds.) *Handbook of Qualitative Research*; Sage: Thousand Oaks, CA, USA, 2000.
55. Bryman, A. *Social Research Methods*, 4th ed.; Oxford University Press: Oxford, UK, 2004.
56. Cohen, L.; Manion, L.; Morrison, K. *Research Methods in Education*, 6th ed.; Routledge, Taylor and Francis Group: London, UK, 2007.
57. Margetson, D. Why is Problem-based Learning a Challlenge. In *The Challenge of Problem-Based Learning*; Boud, D., Feletti, G.E., Eds.; Kogan Page: London, UK, 1997; pp. 72–85.
58. Hung, W.; Jonassen, D.H.; Liu, R. Problem-based learning. In *Handbook of Research on Educational Communications and Technology*; Spector, J.M., van Merriënboer, J.G., Merrill, M.D., Driscoll, M., Eds.; Erlbaum: Mahwah, NJ, USA, 2008; pp. 485–506.
59. Aczel, J. Does epistemology matter for educational practice? In Proceedings of the Annual Conference of the Philosophy of Education Society of Great Britain, Oxford, UK, 5–7 April 2002; Available online: http://oro.open.ac.uk/7179/ (accessed on 18 April 2021).
60. Popper, K.R. *The Open Society and Its Enemies*, 4th ed.; One Volume Edition; Princeton University Press: Princeton, NJ, USA, 2020.
61. Leary, H.M. Self-Directed Learning in Problem-Based Learning versus Traditional Lecture-Based Learning: A Meta-Analysis. Ph.D. Thesis, Utah State University, Logan, UT, USA, 2012. Available online: https://digitalcommons.usu.edu/etd/1173 (accessed on 18 April 2021).
62. Mergendoller, J.R.; Maxwell, N.L.; Bellisimo, Y. Comparing Problem-Based Learning and Traditional Instruction in High School Economics. *J. Educ. Res.* **2000**, *93*, 374–382. [CrossRef]
63. Albanese, M.A.; Mitchell, S. Problem-based Learning: A Review of Literature on its Outcomes and Implementation Issues. *Acad. Med.* **1993**, *68*, 52–81. [CrossRef] [PubMed]

64. Vernon, D.T.A.; Blake, R.L. Does Problem-based Learning Work? A Meta-analysis of Evaluative Research. *Acad. Med.* **1993**, *68*, 550–563. [CrossRef]
65. Kalaian, H.A.; Mullan, P.B.; Kasim, R.M. What Can Studies of Problem-based Learning Tell Us? Synthesizing and Modeling PBL Effects on National Board of Medical Examination Performance: Hierarchical Linear Modeling Meta-analytic Approach. *Adv. Health Sci. Educ.* **1999**, *4*, 209–221. [CrossRef] [PubMed]
66. Berkson, L. Problem-based Learning: Have the Expectations Been Met? *Acad. Med.* **1993**, *68*, 79–88. [CrossRef] [PubMed]
67. Manisha, M.; Aniruddha, K.; Bajaj, P. Problem Based Learning versus Traditional Lecture Method: A Comparative Study among Second Year Medical Students. *Indian J. Forensic Med. Pathol.* **2012**, *5*, 109–114.
68. Murray, J.; Summerlee, A. The Impact of Problem-based Learning in an Interdisciplinary First-year Program on Student Learning Behaviour. *Can. J. High. Educ.* **2007**, *37*, 87–107.
69. Sendaq, S.; Odabas, H.F. Effect of Problem-based Learning Course on Content Knowledge Acquisition and Critical Thinking Skills. *Comput. Educ.* **2009**, *53*, 132–141. [CrossRef]
70. Dochy, F.; Segers, M.; Van den Bossche, P.; Gijbels, D. Effects of Problem Based Learning: A Meta-analysis. *Learn. Instr.* **2003**, *13*, 533–568. [CrossRef]
71. Segers, M.; Dochy, F. New Assessment Forms in Problem-based Learning: The value-added of the students' perspective. *Stud. High. Educ.* **2001**, *26*, 327–343. [CrossRef]
72. Edwards, S.; Hammer, M. Teacher education and problem based learning: Exploring the issues and identifying the benefits. In Proceedings of the AARE International Education Research Conference, Sydney, Australia, 27 November–1 December 2005; pp. 1–13.
73. Alt, D.; Nirit, R. Problem-based Learning, Self- and Peer Assessment in Higher Education: Towards Advancing Lifelong Learning Skills. *Res. Pap. Educ.* **2020**, 1–26. [CrossRef]
74. Papinczak, T.; Young, L.; Groves, M. Peer Assessment in Problem-based Learning: A Qualitative Study. *Adv. Health Sci. Educ.* **2007**, *12*, 169–186. [CrossRef] [PubMed]
75. Lerchenfeldt, S.; Taylor, T.A.H. Best Practices in Peer Assessment: Training Tomorrow's Physicians to Obtain and Provide Quality Feedback. *Adv. Med Educ. Pract.* **2020**, *11*, 571–578. [CrossRef]
76. Sluijsmans, D.M.A.; Moerkerke, G.; van Merriënboer, J.J.G.; Dochy, F.J.R. Peer Assessment in Problem Based Learning. *Stud. Educ. Eval.* **2001**, *27*, 153–173. [CrossRef]
77. Sluijsmans, D.M.A.; Brand-Gruwel, S.; van Merriënboer, J.J.G. Peer Assessment Training in Teacher Education: Effects on Performance and Perceptions. *Assess. Eval. High. Educ.* **2002**, *27*, 443–454. [CrossRef]
78. Falchikov, N.; Goldfinch, J. Student Peer Assessment in Higher Education: A Meta-analysis Comparing Peer and Teacher Marks. *Rev. Educ. Res.* **2000**, *70*, 287–322. [CrossRef]
79. Dolmans, D.; De Grave, W.; Wolfhagen, I.; van der Vleuten, C. Problem-based learning: Future challenges for educational practice and research. *Med. Educ.* **2005**, *39*, 732–741. [CrossRef] [PubMed]
80. Gao, S.; Wang, Y.; Jiang, B.; Fu, Y. Application of problem-based learning in instrumental analysis teaching at Northeast Agricultural University. *Anal. Bioanal. Chem.* **2018**, *410*, 3621–3627. [CrossRef] [PubMed]
81. Sánchez-Ruiz, L.-M.; Moll-López, S.; Moraño-Fernández, J.-A.; Roselló, M.-D. Dynamical Continuous Discrete Assessment of Competencies Achievement: An Approach to Continuous Assessment. *Mathematics* **2021**, *9*, 2082. [CrossRef]
82. Baret, S. Question choice: Does marker variability make examinations a lottery? In Proceedings of the HERDSA Annual International Conference, Melbourne, Australia, 12–15 July 1999. Available online: https://www.researchgate.net/profile/Steven-Barrett-5/publication/260388928_Question_choice_does_marker_variability_make_examinations_a_lottery/links/553efa440cf210c0bdaac5a7/Question-choice-does-marker-variability-make-examinations-a-lottery.pdf (accessed on 18 April 2021).
83. Chang, C.-S.; Chung, C.-H.; Chang, J.A. Influence of problem-based learning games on effective computer programming learning in higher education. *Educ. Technol. Res. Dev.* **2020**, *68*, 2615–2634. [CrossRef]
84. Dahl, B. What is the problem in problem-based learning in higher education mathematics? *Eur. J. Eng. Educ.* **2018**, *43*, 112–125. [CrossRef]

education sciences

MDPI

Article

Assessment and Evaluation in Active Learning Implementations: Introducing the Engineering Education Active Learning Maturity Model

Humberto Arruda * and Édison Renato Silva

Production Engineering Program, Universidade Federal do Rio de Janeiro, Rio de Janeiro 21941-909, Brazil; edison@poli.ufrj.br

* Correspondence: humberto.arruda@poli.ufrj.br

Abstract: With the technological changes experienced in the world in recent decades, society has changed as a whole, due to the speed and availability of information that exists today. As student attention decreases, critical thinking and Active Learning, which places the student at the center of the learning process, have gained prominence. Considering the growing popularity of these techniques, this article proposes the Engineering Education Active Learning Maturity Model (E^2ALM^2), a framework that allows practitioners to assess the current maturity of Active Learning implementation in a program or a course. E^2ALM^2 was built from a literature review of key success factors (KSF) for Active Learning implementations, which were divided into dimensions. Each KSF is composed of constructs, which are detailed with variables. Each variable has a proposed measurement method and an estimated uncertainty level. The framework can support diagnosis and practical improvements in real settings.

Keywords: Active Learning; maturity model; engineering education

Citation: Arruda, H.; Silva, É.R. Assessment and Evaluation in Active Learning Implementations: Introducing the Engineering Education Active Learning Maturity Model. *Educ. Sci.* **2021**, *11*, 690. https://doi.org/10.3390/educsci11110690

Academic Editors: Sandra Raquel Gonçalves Fernandes, Marta Abelha and Ana Teresa Ferreira-Oliveira

Received: 31 August 2021
Accepted: 24 October 2021
Published: 29 October 2021

1. Introduction

Since the beginning of the second half of the 20th century, the world has gone through technological evolutions that have transformed several areas of knowledge. Since the appearance of the first computers, data processing capacity and speed have increased exponentially, and this has led people and society to new behaviors. Education in general has changed, and so has engineering education [1,2].

With the transformations experienced in recent decades, current students were born surrounded by many technological resources. With almost all the information available on mobile phones, knowing how to make sense of it becomes increasingly important.

Engineering schools are experiencing a global trend of adaptation of their programs to the reality of the 21st century. Several movements are attempting to modernize programs and teaching practices, such as the CDIO initiative [3]. This initiative "provides students with an education stressing engineering fundamentals set in the context of Conceiving—Designing—Implementing—Operating (CDIO) real-world systems and products" [4]. Additionally, accreditation criteria of engineering programs in USA, established by the Accreditation Board for Engineering and Technology—ABET (called EC2000) [5], have changed. Such novel criteria require US engineering departments to demonstrate that, in addition to having a solid knowledge of science, math, and engineering fundamentals, their graduates have communication skills, multidisciplinary teamwork capabilities, lifelong learning skills, and awareness of the social and ethical considerations associated with the engineering profession [6]. Finally, completely novel engineering colleges are being created, with totally different proposals from the traditional 20th century model, such as the Olin College [7] and Aalborg University [8].

A common topic among the engineering modernization movements is the importance of placing the student at the center of the learning process, as highlighted in the learning outcomes of the EC2000 (Criterion 3. i—"a recognition of the need for, and an ability to engage in life-long learning") [9] and Standard 8 of the CDIO ("Active Learning") (p. 153, [3]). Putting the student at the center of the learning process, along with increasing student engagement, is arguably achieved by the use of Active Learning [10–22].

Active Learning still lacks a definitive unique definition, but three stand as the most popular. Prince defines it as "any instructional method [used in the classroom] that engages students in the learning process" [23], Roehl as "an umbrella term for pedagogies focusing on student activity and student engagement in the learning process" [24], and Barkley as "an umbrella term that now refers to several models of instruction, including cooperative and collaborative learning, discovery learning, experiential learning, problem-based learning, and inquiry-based learning" [14]. Hartikainen [10] shows 66 definitions of Active Learning, grouped by three main categories: (1) defined and viewed as an instructional approach; (2) not defined but viewed as an instructional approach; and (3) not defined but viewed as a learning approach.

Among the main Active Learning techniques, the following stand out: Problem-Based Learning (PBL) [8,23,25–29], Cooperative and Collaborative Learning [13,23,30–35], and the Flipped Classroom [20,36–39].

Furthermore, the pedagogical results and effectiveness of Active Learning are also widely documented [19,23,40–45]. Hartikainen [10] related positive effects on the development of subject-related knowledge, professional skills, social skills, communication skills, and meta-competences.

However, there are problems both in research and in the implementation of Active Learning. Prince [23] points out that comprehensive assessment of Active Learning is difficult due to the limited range of learning outcomes and different possible interpretations of these outcomes. Streveler [46] notes that "active learning is not a panacea that is a blanket remedy for all instructional inadequacies. Instead, it is a collective term for a group of instructional strategies that produce different results and require differing degrees of time to design, implement, and assess". Fernandes [47] related that "students identify the heavy workload which the project entails as one of the main constraints of PBL approach". There are also the least researched, but much-mentioned, barriers of resistance to novelty on the part of lecturers and students [43,48–52].

Although Active Learning has already been validated as an effective way to influence student learning and is increasingly being incorporated into the classroom, there is no way to qualify and evaluate the use of Active Learning techniques by faculty members [40]. There are four maturity models in the field of education, but none that specifically allow the assessment of the implementation of Active Learning in a course [53–56]. In addition to the difficulty of measuring Active Learning usage in the classroom, there is no way to assess the maturity level of Active Learning implementations in a course or a program of a Higher Education Institution (HEI), engineering schools included. This gap blurs the diagnostics of the status of a given implementation and consequently leads to less assertiveness in decision making, reducing the effectiveness of changes and Active Learning as a whole.

Maturity models can be a bridge to this gap. They enable practitioners to assess organizational performance, support management, and allow improvements [57]. Maturity modeling is a generic approach that describes the development of an organization over time through ideal levels to a final state [58]. In addition, maturity models are instruments to assess organizational elements and select appropriate actions, which lead to higher levels of maturity and better performance [59].

Therefore, this work will propose a conceptual maturity model that allows evaluating Active Learning implementations at the level of a specific course. This model targets the incremental enhancement of courses and seems logically to be the first step towards a more general and comprehensive framework that can extend its reach to evaluate institutions as a whole.

2. Methodology

Based on the research objectives, the broad keyword "active learning" was used in Scopus and Web of Science databases to search for abstracts of peer-reviewed journal articles. Additional keywords related to "success factors" and "engineering education" were used to refine the search. Ultimately, 31 studies were selected for review. Figure 1 uses the PRISMA model [60,61] to describe the literature review process.

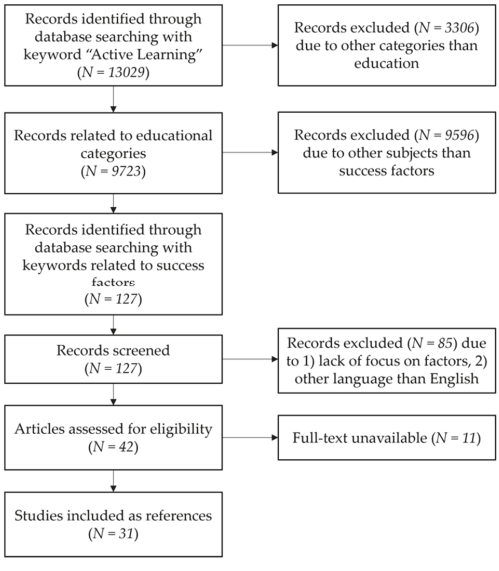

Figure 1. Source selection process (N = 31).

The initial search returned a total number of 13,029 articles. Approximately 25% (3306) of the records were excluded because they belonged to categories other than education. The objective of this criterion was to exclude articles that used "active learning" in different purposes.

With the sample reduced to 75% of the original size (9723), filters were applied in the databases to match keywords related to success factors: "critic* factor*", "key factor*", and "success factor*". This step led to the reduction of the sample to 127 articles.

The abstracts of these 127 articles were judged against the following inclusion criteria: (1) reported research on key factors and (2) written in English. These criteria were intended to eliminate articles that had some keywords related to success factors but that did not actually address them. This step resulted in the reduction of the sample to 42 articles, whose full texts were searched. Of these, 11 full texts were not available for download, which resulted in the selection of 31 references that were included in the literature review. After the literature selection stage, references were read to identify the Key Success Factors (KSF) for the implementation of Active Learning.

The software MaxQDA® was used to extract and accumulate text snippets that represented key success factors. Then, similar snippets were combined into single KSFs to avoid duplication. Next, a definition based on the literature was attributed to each factor. The following step was to define the relevant constructs for each factor and for each construct, the variables that would be used for measurement.

Finally, each variable had a measurement method proposed and an uncertainty degree estimated.

The research method is presented in Figure 2.

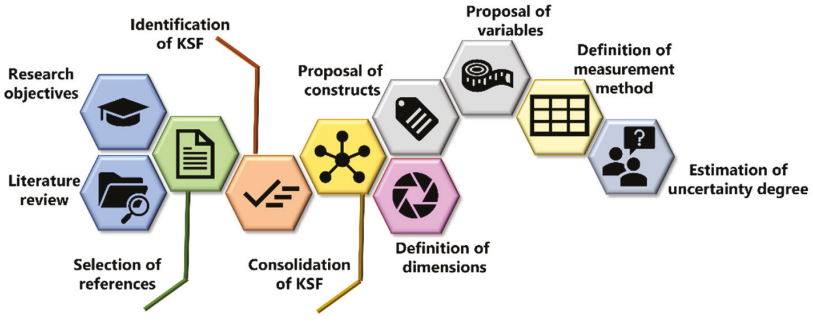

Figure 2. Research procedures.

3. Results

The 31 sources included for the literature review provided 14 key success factors, grouped into five dimensions according to their similarity and relatedness to a specific aspect of the educational environment. Table 1 shows the dimensions and their related KSF.

Table 1. Dimensions, KSF, and references.

Dimension	KSF	References
Content quality	Course artifacts	[16–18,42,49,53,62–66]
	Student assessment	[18,34,42,62,67–70]
	Learning facilitation	[17,18,42,53,67,70,71]
Organizational environment	Culture	[72]
	Policy	[19,72]
	Student feedback	[25,49,73]
Organizational infrastructure	Instructional design	[74,75]
	Classrooms	[76–81]
	Technology	[16,19,42,70,82]
Lecturer	Knowledge	[49,72]
	Skills	[72]
	Attitude	[42,72]
Interactions	Between students	[42,62]
	With lecturers	[42,75]

Following up on the creation of dimensions, each of the 14 KSF was detailed into 41 constructs. The constructs were detailed into 90 variables that could operationalize objective measurements to assess the maturity of a given implementation. Then, a measurement method was proposed for each variable, as well as an uncertainty degree estimated based on each measurement method. Three measurement methods were proposed:

- A questionnaire faculty in charge of a course should answer (Lecturer Questionnaire, LQ),
- Another questionnaire directed to students (Student Questionnaire, SQ), and
- An external evaluation from a third party not directly involved in the course (External Evaluation, EE).

As a result, Figure 3 shows the Engineering Education Active Learning Maturity Model (E^2ALM^2) with four levels.:

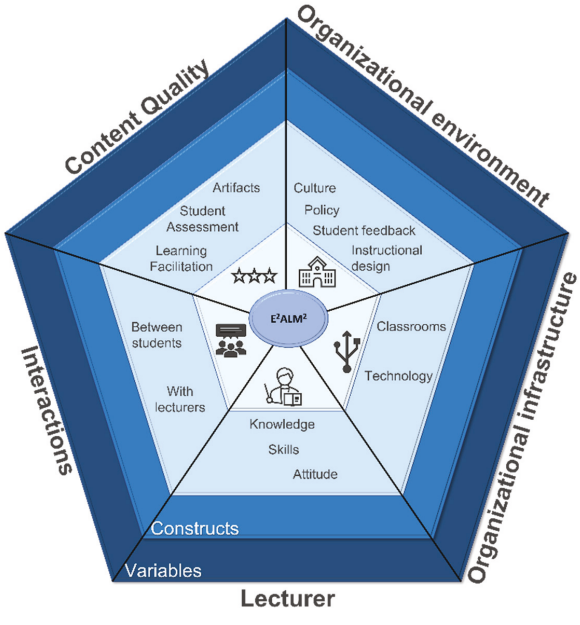

Figure 3. E^2ALM^2.

All dimensions and their KSF are defined in the following sections. Each KSF is detailed with its constructs and variables. Each variable has a measurement method (MM) and uncertainty degree (UD) suggested.

3.1. Content Quality

This dimension concentrates the factors related to the core of the learning process, such as the quality of the problems, projects, or cases studied (artifacts); the level of difficulty required from the students; whether the activities facilitate learning; and whether the evaluation criteria are clear and consistent. The three KSF are detailed below.

3.1.1. Course Artifacts

Course artifacts (problems, projects, or cases studied) should:

- Engage students with real-life problems and active experiences [62];
- Provide students with a variety of additional instructional resources, such as simulations, case studies, videos, and demonstrations [62];
- Be suitable to achieve different targets including the support of the students' learning process and establishing learning outcomes requirements [53];
- Be clearly written, in the right length, useful, flexible, and provide an appropriate degree of breath [63];
- Have suitable intellectual challenge [16–18,42,64]; and
- Begin with an explanation of its purpose [49,65,66].

Table 2 describes the KSF "Course Artifacts" with more detail. Its constructs were derived from the list of requisites presented above. Variables were proposed to measure each construct, as well as the most suitable measurement method (MM) and the uncertainty degree in each measurement.

Table 2. KSF "Course Artifacts".

Construct	Variable	MM	UD
Use of real-life problems	% of course content based on real-life problems	SQ and LQ	Medium
Application of active experiments	% of classes using active methods	SQ and LQ	Medium
	Students' perception of hands-on activities	SQ	Low
Variety of instructional resources	Quantity of instructional resources used	SQ and LQ	Low
	% of classes using resources other than the board or projector	SQ and LQ	Low
	Students' perception of the use of various resources	SQ	Low
Adequacy to learning outcomes (LO)	% of classes linked directly to an LO	SQ and LQ	Medium
	Students'perception of reaching an LO	SQ	Medium
Suitability of intellectual challenge	Students' perception of the level of difficulty presented	SQ	Low
Clarity in writing of course activities	Students' perception of the clarity used	SQ	Low
Size of course activities	Students' perception of size	SQ	Low
Explanation of purpose of course activities	Students' perception of clarity in the purpose of the activities	SQ	Low
	% of activities in which the purpose is explained to students	SQ	Low

3.1.2. Student Assessment

Student assessment needs to be clear, concise, and consistent. This involves instructions, assignments, assessments, due dates, course pages, and office hours [62]. Furthermore, criteria for success must be communicated clearly and monitored [18,34,42,67–70]. Table 3 details the KSF "Student Assessment".

Table 3. KSF "Student Assessment".

Construct	Variable	MM	UD
Clearness of assessment methods	Perception of students on the clarity of assessment methods	SQ	Low
	Are the assessment methods defined in advance?	SQ and LQ	Low
	% of activities that have defined what is expected of the student	SQ and LQ	Low
Clearness of criteria for success	Perception of students on the clarity of success criteria	SQ	Low
	Are the success criteria defined in advance?	SQ and LQ	Low
Communications with students	Is information about assessment methods and success criteria made available before (or at the beginning of) the course?	SQ and LQ	Low
	Students' perception of communication of assessment methods and success criteria	SQ	Low

3.1.3. Learning Facilitation

Learning facilitation includes the preparation of students to conduct activities and tasks required in addition to activities related to the facilitator guiding the learning process of the students [53]. It also involves providing students with regular opportunities for formative feedback from the lecturer [17,18,42,67,70,71]. Table 4 details the following levels of this KSF.

Table 4. KSF "Learning Facilitation".

Construct	Variable	MM	UD
Preparation of students to conduct activities required	% of activities flagged as supporting another activity	SQ and LQ	Medium
	Students' perception of the existing preparation for conducting activities	SQ	Medium
	Students' perception of the teacher's performance as a facilitator	SQ	High
	Intensity of the participation of monitors or auxiliary teachers during the course	SQ and LQ	Medium
Formative feedback from teacher	% of activities where there is formative feedback from the teacher	SQ	Medium
	Students' perception of the intensity of support received via formative feedback	SQ	Medium

3.2. Organizational Environment

The factors of this dimension represent abstract aspects of the institution, such as culture, policy, and the practice of collecting feedback from students.

3.2.1. Culture

Organizational culture is a set of values systems followed by members of an organization as guidelines for behavior and solving the problems that occur in the organization [72]. This way, an organization and its members should have behavior alignment, and an organization should have guidelines to solve problems. Table 5 details the following levels of this KSF.

Table 5. KSF "Culture".

Construct	Variable	MM	UD
Acceptance of changes by the organization	Ease of approval of pedagogical changes	LQ	Medium
	Ease of approval of administrative changes	LQ	Medium
Behavior alignment	Clarity of expected behaviors	LQ	Medium
	Existence (or maturity) of behavioral guidelines	LQ	Low
Ability to solve problems	Perception of the speed with which problems are solved	LQ	Low
	Perception of transparency in problem solving	LQ	Low
Defining rules	Existence (or maturity) of a code of ethics	LQ	Low
Adequacy to the rules	Perception of the existence of punishments for those who violate certain rules	LQ	Medium

3.2.2. Policy

Organization policy is a set of program plans, activities, and actions that allows the prediction of how the organization works and how a problem would be solved [72]. Once time is needed to prepare the activities, teachers must have it for implementing something new in their classes [19]. Table 6 describes more details of this KSF.

Table 6. KSF "Policy".

Construct	Variable	MM	UD
Organizational support for the preparation of activities	Perception of the existence of time available for planning new activities	LQ	Low
	% average of teachers' time in classroom activities	LQ	Low
	% average of teachers' time in administrative activities	LQ	Low
	Average amount of administrative functions performed by teachers	LQ	Low
	Perception of the availability of auxiliary resources for the preparation of activities	LQ	Low
Adequacy of pedagogical plans	Perception of the adequacy of existing teaching plans to the use of AL	LQ	Medium

3.2.3. Student Feedback

Organizations are expected to collect feedback from students [25,49,73] and provide the support needed to successfully complete the activity [49].

Thus, it is possible to identify three different requirements for organizations carry on successfully this process: having a suitable process of feedback collection, using suitable feedback, and having an adequate student feedback process.

The following levels of the KSF "Student Feedback" are shown in Table 7.

Table 7. KSF "Student Feedback".

Construct	Variable	MM	UD
Collecting student feedback	Existence (or maturity) of the process of receiving feedback from students	SQ	Low
Using student feedback	Perception of students on the fulfilment of their placements in feedbacks	SQ	Medium
	Number of objective actions resulting from student feedback in the last years	EE	Low
Quality of the student feedback	Is feedback anonymous?	SQ	Low
	Is the collection in person or remote?	SQ	Low
	Perception of students about the ease of the process of giving feedback	SQ	Low

3.2.4. Instructional Design

Brophy [74] and Paechter et al. [75] highlight the importance of the structure and coherence of the curriculum and the learning materials. Thus, it is possible to identify two requirements to this KSF:

- Curriculum should be suitable to the course needs; and
- Curriculum and learning material should have coherence with each other.

Table 8 shows the following levels of this KSF.

Table 8. KSF "Instructional Design".

Construct	Variable	MM	UD
Structure of the curriculum	Perception about the adequacy of the curriculum to the needs of the course	SQ	High
Coherence of the curriculum and the learning material	Student perception of the alignment of the curriculum with the course material	SQ	Medium

3.3. Organizational Infrastructure

This dimension contains factors that represent the infrastructure available for course activities.

3.3.1. Classrooms

Classrooms designed for improved Active Learning experience [76] and equipped with technologies can enhance student learning and support teaching innovation [77–81]. Thus, two different requirements emerge for this KSF:

- Organizations should have appropriate classrooms for Active Learning; and
- Organizations should provide classrooms with technological support.

Table 9 describes more details of this KSF.

Table 9. KSF "Classrooms".

Construct	Variable	MM	UD
Classrooms designed for improve Active Learning experience	Existence of classrooms for Active Learning	SQ and LQ	Low
	Classroom availability for Active Learning	SQ and LQ	Low
	% of activities performed in an environment suitable for Active Learning	SQ and LQ	Medium
Classrooms equipped with technologies to enhance student learning and support teaching innovation	Existence of classrooms equipped with multimedia devices and/or laboratories	SQ and LQ	Low
	Availability of classrooms equipped with multimedia devices and/or laboratories	SQ and LQ	Low
	% of activities performed in a technologically appropriate environment	SQ and LQ	Low

3.3.2. Technology

The school should provide equipment and technological structure [19,42]. This involves availability, reliability, accessibility, usability of devices, internet (Wi-Fi), learning support, and inclusive learning environment [16,42,70,82].

Table 10 shows the details of KSF "Technology".

Table 10. KSF "Technology".

Construct	Variable	MM	UD
Availability of technology	Availability of multimedia devices	SQ and LQ	Low
	Internet availability on campus	SQ and LQ	Low
	Availability of e-learning system	SQ and LQ	Low
Reliability of technology	Reliability of multimedia devices	SQ and LQ	Medium
	On-campus internet reliability	SQ and LQ	Low
	Reliability of e-learning system	SQ and LQ	Medium
Accessibility of technology	Accessibility of multimedia devices	SQ and LQ	Medium
	On-campus internet accessibility	SQ and LQ	Low
	Accessibility of e-learning system	SQ and LQ	Low
Usability of technology	Usability of multimedia devices	SQ and LQ	Medium
	Campus internet usability	SQ and LQ	Low
	Usability of e-learning system	SQ and LQ	Medium

3.4. Lecturer

The lecturer is single most important actor in a successfully implementation of Active Learning. This dimension groups factors that represents their knowledge, skills, and attitude to carry out education innovation.

3.4.1. Knowledge

Knowledge is a combination of framed experience, values, and contextual information that provides an environment for evaluating and incorporating new experiences [72]. De-Monbrun et al. highlighted the relevance of experience to lecturer [49]. Therefore, lecturer should have suitable experience as faculty member and information about Active Learning.

Table 11 details the KSF "Knowledge".

Table 11. KSF "Knowledge".

Construct	Variable	MM	UD
Experience	Activity time as a lecturer	LQ	Low
	Highest academic title	LQ	Low
	Time since the highest titration	LQ	Low
Contextual information	Level of knowledge about Active Learning	LQ	High

3.4.2. Skills

Skills are the ability to use reason, thoughts, ideas, and creativity in doing, changing, or making things more meaningful so as to produce a value from the results of the work [72]. The lecturer should have skills about educational innovations in general and about Active Learning specifically. Table 12 shows this KSF in detail.

Table 12. KSF "Skills".

Construct	Variable	MM	UD
Skills about Active Learning	Amount of participation in Active Learning events	LQ	Low
	Number of books read on Active Learning	LQ	Low
	Amount of Active Learning techniques over which you have mastery	LQ	Low
Skills about educational innovations	Amount of participation in events on educational innovations	LQ	Low
	Number of books read on educational innovations	LQ	Low

3.4.3. Attitude

Attitude encompasses a very broad range of activities, including how people walk, talk, act, think, perceive, and feel [72]. Hegarty and Thompson [42] highlight the relevance of lecturer attributes and teaching methods, such as approachable, supportive, enthusiastic, and interesting delivery. Table 13 shows the following levels of this KSF.

Table 13. KSF "Attitude".

Construct	Variable	MM	UD
Willingness to adopt Active Learning techniques	Qualitative perception of disposition	EE	High
	Number of periods in which adoption was attempted	LQ	Low
	Number of subjects in which adoption was attempted	LQ	Low
	Time since last adoption attempt	LQ	Low
Demographics	Age	LQ	Low
	Current position	LQ	Low
	Study area	LQ	Low

3.5. Interactions

Placing students at the center of the learning process requires them to step out of the role of recipients of information and become active agents. The interaction between students and between them and teachers allows this transition to happen.

3.5.1. Between Students

Opportunities for students to work together and obtain peer feedback included in the learning design [42]. Chen, Bastedo, and Howard [62] emphasize that the course should provide online and face-to-face opportunities for students to collaborate with others.

Table 14 shows this KSF in detail.

Table 14. KSF "Interactions between Students".

Construct	Variable	MM	UD
Interactions in general	Quantity of work/projects carried out in group in the course	SQ	Medium
	% of the grade of the discipline from group work	SQ	Medium
Online collaboration	Number of remote meetings with other students throughout the course	SQ	Low
	Number of online presentations made by the student with assistance from other students	SQ	Low
Face-to-face collaboration	Number of face-to-face meetings with other students throughout the course	SQ	Low
	Number of face-to-face presentations made by the student with the assistance of other students	SQ	Low

3.5.2. With Lecturers

Interaction between students and lecturer supports knowledge construction, motivation, and the establishment of a social relationship [75]. Furthermore, constructive and enriching feedbacks from the lecturer lead to increasing academic success and feelings of support [42]. Table 15 details this KSF.

Table 15. KSF "Interactions with Lecturers".

Construct	Variable	MM	UD
Interactions students/professors	Number of orientation meetings throughout the course	SQ and LQ	Low
	Number of meetings to monitor projects throughout the course	SQ and LQ	Low

3.6. Measurement Scales

Most of E^2ALM^2 variables are related to the perception of students and teachers. They can be measured on a five-point Likert scale [83], coded as 5: strongly agree; 4: agree; 3: neither agree nor disagree; 2: disagree; and 1: strongly disagree.

The model also involves numerical variables, such as the percentage of activities that define clearly what is expected of the student or the percentage of activities in which the purpose is explained to students. For these variables, it is also possible to use a five-point scale, however with coding based on frequency or ranges, such as 5: always, 4: often, 3: occasionally, 2: rarely, and 1: never.

Finally, there are binary variables, e.g., whether assessment methods are defined in advance.

3.7. KSF Weights

In the proposed model, each dimension has a score independent of the others. Thus, there is no need to define weights for the dimensions. However, it is necessary to define the weight that each KSF has in the composition of the score within its dimension. Two approaches are possible: (i) a uniform distribution inside the dimension and (ii) a distribution according to the relative relevance, based on number of references that support each KSF. Table 16 presents KSF weights under two criteria.

Table 16. KSF Weights.

Dimension	KSF	(i) Uniform Distribution	Number of References	(ii) Relative Relevance
Content quality (references = 26)	Artifacts	0.33	11	0.42
	Student Assessment	0.33	8	0.31
	Learning Facilitation	0.33	7	0.27
Organizational environment (references = 8)	Culture	0.25	1	0.13
	Policy	0.25	2	0.25
	Student Feedback	0.25	3	0.38
	Instructional Design	0.25	2	0.25
Organizational infrastructure (references = 11)	Classrooms	0.50	6	0.55
	Technology	0.50	5	0.45
Lecturer (references =5)	Knowledge	0.33	2	0.40
	Skills	0.33	1	0.20
	Attitude	0.33	2	0.40
Interactions (references = 4)	Between students	0.50	2	0.50
	With lecturers	0.50	2	0.50

4. Discussion

As explained in the introduction, there is a lack of instruments that can help engineering schools and lecturers assess Active Learning implementations. The use of maturity models can support them in this task.

According to Bruin et al. [84], the maturity assessment can be descriptive, prescriptive, or comparative in nature. A purely descriptive model can be applied for an as-is diagnosis, with no provision for improving maturity or providing relationships with performance. A prescriptive model emphasizes the relationships between variables for final performance and indicates how to approach maturity improvement to positively affect the outcome. Therefore, it allows the development of a roadmap for improvement. A comparative model allows benchmarking across sectors or regions. Thus, it would be possible to compare similar practices between organizations to assess maturity in different sectors.

The E^2ALM^2 is a descriptive maturity model (according to Bruin et al.'s classification), which can be understood as the first step in a life cycle that will allow the evolution to a prescriptive model. This evolution requires more knowledge about the impact of actions

and the identification of replicable actions that support the advance in the maturity level. This difficulty is especially important due to the difference in results obtained in education when different contexts and conditions are compared [43].

Although there are four other maturity models in the field of education, they have a different focus from E^2ALM^2. These models are focused on: Project-Based Learning (PBLCMM) [53], Student Engagement (SESR-MM) [54], Curriculum Design (CDMM) [55], and e-Learning [56]. In addition to the difference in focus, none of these models provide an assessment of the same requirements and with the scope of E^2ALM^2. In addition to these four models, there is an extremely simple scale, which is neither a theoretical model with scientific references nor peer-reviewed, but which has a similar objective to assess the use of Active Learning [85].

The E^2ALM^2 model allows the diagnosis of the current stage of Active Learning implementation with a focus on a course, from the objective measurement of 90 variables. For most variables, the suggested measurement method is a questionnaire for the lecturer, for the student, or for both. This choice aims to facilitate the application of the model in real cases, reducing the need for an external evaluator to observe the activities throughout the entire period to issue its report.

Obviously, collecting impressions through questionnaires introduces the possibility of bias, both for the teacher and the student. Therefore, it will be necessary to use response validation techniques when creating the questionnaires. Because of this possibility of bias, all variables had an estimated uncertainty degree. In cases where the uncertainty degree is high, the statistical validation of answers will need to be stricter. As a way to avoid possible contamination in the results due to bias, some variables are measured by questions asked to both the lecturer and the students.

The use of Active Learning has several positive effects, as explained in the introduction, but there are also some difficulties and limitations. Streveler states that Active Learning is not a solution for all instructional inadequacies [46]. The increasing workload for lecturers [52,86] and students [47], the resistance to changes [43,48–51], and the need to align curriculum and course activities [86,87] are challenges that need to be overcome in Active Learning implementations.

Furthermore, it is important to emphasize that the E^2ALM^2 model does not aim to assess the overall quality of an engineering program, but the maturity level of Active Learning implementation, which is a recommendation of the main modernization movements in the Engineering Education field around the world. Courses and schools can still be of a high quality even though they follow a more traditional approach to engineering education. The point here is that whoever wants to modernize their engineering education approach will struggle with the implementation of Active Learning as a pedagogical and cultural element, and the E^2ALM^2 can shed light for managers and lecturers during the messy times of changes, infrastructural adaptations, and resistance from students and faculty members.

As future work, we recommend: (i) defining further studies to test the scale of each variable; (ii) determining empirical testing of the weights of each KSF in their respective dimensions; (iii) testing the questionnaires to measure all variables; (iv) validation of the framework in different cultural settings, for instance with an international panel of experts; and (v) application of the framework to evaluate the maturity of real cases, which will allow qualitative and quantitative analyses.

5. Conclusions

This study proposed a framework to evaluate the maturity of adoption of Active Learning by a specific course. The variables described here can serve as a checklist to lecturers adopting Active Learning and as a metric to evaluate the comprehensiveness and quality of existing initiatives.

The proposed model is descriptive, because it allows evaluating the current situation, but it can be understood as a first step towards the construction of a prescriptive model, which can indicate good practices and replicable actions to increase the level of maturity.

E^2ALM^2 was designed so that its application is easy, centered on questionnaires for lecturers and students, without the need for long periods of external observation, which would lead to greater expenses and prevent scalability.

E^2ALM^2 allows faculty members to assess the current state of Active Learning implementations and therefore compare states before and after planned interventions with specific objectives.

Despite having the focus on a course, the diagnosis of a program or an engineering school can be made as a composition of the evaluations of the courses that comprise it, which also favors managerial actions.

Author Contributions: Writing—original draft, H.A.; Writing—review & editing, É.R.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Beanland, D.; Hadgraft, R. *Engineering Education: Transformation and Innovation*; RMIT Publishing: Melbourne, Australia, 2013.
2. Graham, R. *The Global State of the Art Engineering Education: March 2018*; Massachusetts Institute of Technology (MIT): Cambridge, MA, USA, 2018.
3. Crawley, E.F.; Malmqvist, J.; Östlund, S.; Brodeur, D.R.; Edström, K. *Rethinking Engineering Education: The CDIO Approach*, 2nd ed.; Springer: New York, NY, USA, 2014.
4. CDIO | Worldwide Initiative. Available online: http://cdio.org/about (accessed on 12 October 2021).
5. Lattuca, L.R.; Terenzini, P.T.; Volkwein, J.F. *Engineering Change: A Study of the Impact of EC2000*; ABET, Inc.: Baltimore, MD, USA, 2006.
6. Rugarcia, A.; Felder, R.M.; Woods, D.R.; Stice, J.E. The future of engineering education I. A vision of a new century. *Chem. Eng. Educ.* **2000**, *34*, 16–25.
7. Goldberg, D.E.; Somerville, M. *A Whole New Engineer*; ThreeJoy Associates, Inc.: Douglas, MI, USA, 2014.
8. Mohd-yusof, K.; Arsat, D.; Borhan, M.T.B.; de Graaff, E.; Kolmos, A. *PBL Across Cultures*; Aalborg Universitet: Aalborg, Denmark, 2013.
9. Ramaswamy, S.; Harris, I.; Tschirner, U. Student peer teaching: An innovative approach to instruction in science and engineering education. *J. Sci. Educ. Technol.* **2001**, *10*, 165–171.
10. Hartikainen, S.; Rintala, H.; Pylväs, L.; Nokelainen, P. The Concept of Active Learning and the Measurement of Learning Outcomes: A Review of Research in Engineering Higher Education. *Educ. Sci.* **2019**, *9*, 276. [CrossRef]
11. Thomas, I. Critical Thinking, Transformative Learning, Sustainable Education, and Problem-Based Learning in Universities. *J. Transform. Educ.* **2009**, *7*, 245–264. [CrossRef]
12. Albert, M.; Beatty, B.J. Flipping the Classroom Applications to Curriculum Redesign for an Introduction to Management Course: Impact on Grades. *J. Educ. Bus.* **2014**, *89*, 419–424. [CrossRef]
13. Bonwell, C.; Eison, J. *Active Learning: Creating Excitement in the Classroom. 1991 ASHE-ERIC Higher Education Reports. 1991*; ERIC Publications: Washington, DC, USA, 1991.
14. Barkley, E.F. *Student Engagement Techniques: A Handbook for College Faculty*, 1st ed.; The Jossey-Bass: San Francisco, CA, USA, 2010.
15. Carini, R.M.; Kuh, G.D.; Klein, S.P. Student engagement and student learning: Testing the linkages. *Res. High. Educ.* **2006**, *47*, 1–32. [CrossRef]
16. Kuh, G.D. *The National Survey of Student Engagement: Conceptual Framework and Overview of Psychometric Properties*; Indiana University Center for Postsecondary Research: Bloomington, IN, USA, 2001.
17. Zepke, N.; Leach, L.; Butler, P. Student engagement: What is it and what influences it. *Wellingt. Teach. Learn. Res. Initiat.* **2010**, *1*, 1–22.
18. Zepke, N.; Leach, L. Improving student engagement: Ten proposals for action. *Act. Learn. High. Educ.* **2010**, *11*, 167–177. [CrossRef]
19. Hernández-de-Menéndez, M.; Guevara, A.V.; Martínez, J.C.T.; Alcántara, D.H.; Morales-Menendez, R. Active learning in engineering education. A review of fundamentals, best practices and experiences. *Int. J. Interact. Des. Manuf.* **2019**, *13*, 909–922. [CrossRef]

20. McLaughlin, J.E.; Roth, M.T.; Glatt, G.M.; Gharkholonarehe, N.; Davidson, C.A.; Griffin, L.M.; Esserman, D.A.; Mumper, R.J. The flipped classroom: A course redesign to foster learning and engagement in a health professions school. *Acad. Med.* **2014**, *89*, 236–243. [CrossRef] [PubMed]
21. Slavich, G.M.; Zimbardo, P.G. Transformational Teaching: Theoretical Underpinnings, Basic Principles, and Core Methods. *Educ. Psychol. Rev.* **2012**, *24*, 569–608. [CrossRef] [PubMed]
22. Fernandes, S.; Mesquita, D.; Flores, M.A.; Lima, R.M. Engaging students in learning: Findings from a study of project-led education. *Eur. J. Eng. Educ.* **2013**, *39*, 55–67. [CrossRef]
23. Prince, M. Does Active Learning Work? A Review of the Research. *J. Eng. Educ.* **2004**, *93*, 223–231. [CrossRef]
24. Roehl, A.; Reddy, S.L.; Shannon, G.J. The Flipped Classroom: An Opportunity To Engage Millennial Students Through Active Learning Strategies. *J. Fam. Consum. Sci.* **2013**, *105*, 44–49. [CrossRef]
25. Yadav, A.; Subedi, D.; Lundeberg, M.A.; Bunting, C.F. Problem-based learning: Influence on students' learning in an elec-trical engineering course. *J. Eng. Educ.* **2011**, *100*, 253–280. [CrossRef]
26. Edström, K.; Kolmos, A. PBL and CDIO: Complementary models for engineering education development. *Eur. J. Eng. Educ.* **2014**, *39*, 539–555. [CrossRef]
27. Felder, R.M.; Silverman, L.K. Learning and Teaching Styles in Engineering Education. *Eng. Educ.* **1988**, *78*, 674–681.
28. Hoidn, S.; Kärkkäinen, K. *Promoting Skills for Innovation in Higher Education. A Literature-Review on the Effectiveness of Problem-based Learning and of Teaching Behaviours*; OECD Education Working Papers, No. 100; OECD Publishing: Paris, France, 2014.
29. Chong, J.L.; Benza, R. Teaching Innovation Skills. *Bus. Educ. Innov. J.* **2015**, *7*, 43–50.
30. Rodriguez-Triana, M.J.; Prieto, L.P.; Holzer, A.; Gillet, D. Instruction, Student Engagement, and Learning Outcomes: A Case Study Using Anonymous Social Media in a Face-to-Face Classroom. *IEEE Trans. Learn. Technol.* **2020**, *13*, 718–733. [CrossRef]
31. Holbert, K.; Karady, G.G. Strategies, Challenges and Prospects for Active Learning in the Computer-Based Classroom. *IEEE Trans. Educ.* **2008**, *52*, 31–38. [CrossRef]
32. Carr, R.; Palmer, S.; Hagel, P. Active learning: The importance of developing a comprehensive measure. *Act. Learn. High. Educ.* **2015**, *16*, 173–186. [CrossRef]
33. Salaber, J. Facilitating student engagement and collaboration in a large postgraduate course using wiki-based activities. *Int. J. Manag. Educ.* **2014**, *12*, 115–126. [CrossRef]
34. Chapman, E. Alternative approaches to assessing student engagement rates. *Pract. Assess. Res. Eval.* **2003**, *8*, 2002–2003.
35. Bolton, K.; Saalman, E.; Christie, M.; Ingerman, Å.; Linder, C. SimChemistry as an active learning tool in chemical educa-tion. *Chem. Educ. Res. Pract.* **2008**, *9*, 277–284. [CrossRef]
36. Burke, A.S.; Fedorek, B. Does 'flipping' promote engagement?: A comparison of a traditional, online, and flipped class. *Act. Learn. High. Educ.* **2017**, *18*, 11–24. [CrossRef]
37. Van Alten, D.C.; Phielix, C.; Janssen, J.; Kester, L. Effects of flipping the classroom on learning outcomes and satisfaction: A meta-analysis. *Educ. Res. Rev.* **2019**, *28*, 100281. [CrossRef]
38. Mori, T. The Flipped Classroom: An Instructional Framework for Promotion of Active Learning. In *Deep Active Learning*; Matsushita, K., Ed.; Springer: Singapore, 2017; pp. 95–109. [CrossRef]
39. Howell, R.A. Engaging students in Education for Sustainable Development: The benefits of active learning, reflective prac-tices and flipped classroom pedagogies. *J. Clean. Prod.* **2021**, *325*, 129318. [CrossRef]
40. Van Amburgh, J.A.; Devlin, J.W.; Kirwin, J.L.; Qualters, D.M. A Tool for Measuring Active Learning in the Classroom. *Am. J. Pharm. Educ.* **2007**, *71*, 85. [CrossRef]
41. Freeman, S.; Eddy, S.L.; McDonough, M.; Smith, M.K.; Okoroafor, N.; Jordt, H.; Wenderoth, M.P. Active learning increases student performance in science, engineering, and mathematics. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 8410–8415. [CrossRef]
42. Hegarty, B.; Thompson, M. A teacher's influence on student engagement: Using smartphones for creating vocational as-sessment ePortfolios. *J. Inf. Technol. Educ. Res.* **2019**, *18*, 113–159.
43. Ito, H.; Kawazoe, N. Active Learning for Creating Innovators: Employability Skills beyond Industrial Needs. *Int. J. High. Educ.* **2015**, *4*, 81. [CrossRef]
44. Lizzio, A.; Wilson, K. Action learning in higher education: An investigation of its potential to develop professional capa-bility. *Stud. High. Educ.* **2004**, *29*, 469–488. [CrossRef]
45. Baepler, P.; Walker, J.D.; Driessen, M. It's not about seat time: Blending, flipping, and efficiency in active learning class-rooms. *Comput. Educ.* **2014**, *78*, 227–236. [CrossRef]
46. Streveler, R.A.; Menekse, M. Taking a Closer Look at Active Learning. *J. Eng. Educ.* **2017**, *106*, 186–190. [CrossRef]
47. Fernandes, S.R.G. Preparing Graduates for Professional Practice: Findings from a Case Study of Project-based Learning (PBL). *Proc.—Soc. Behav. Sci.* **2014**, *139*, 219–226. [CrossRef]
48. Borrego, M.; Nguyen, K.A.; Crockett, C.; DeMonbrun, M.; Shekhar, P.; Tharayil, S.; Finelli, C.J.; Rosenberg, R.S.; Waters, C. Systematic Literature Review of Students' Affective Responses to Active Learning: Overview of Results. In Proceedings of the IEEE Frontiers in Education Conference (FIE), San Jose, CA, USA, 3–6 October 2018; pp. 1–7. [CrossRef]
49. DeMonbrun, M.; Finelli, C.J.; Prince, M.; Borrego, M.; Shekhar, P.; Henderson, C.; Waters, C. Creating an Instrument to Measure Student Response to Instructional Practices. *J. Eng. Educ.* **2017**, *106*, 273–298. [CrossRef]
50. Bicknell-Holmes, T.; Hoffman, P.S. Elicit, engage, experience, explore: Discovery learning in library instruction. *Ref. Serv. Rev.* **2000**, *28*, 313–322. [CrossRef]

51. Andrews, M.; Prince, M.; Finelli, C.; Graham, M.; Borrego, M.; Husman, J. Explanation and Facilitation Strategies Reduce Student Resistance to Active Learning. *Coll. Teach.* **2021**, 1–11. [CrossRef]
52. Alves, A.C.; Sousa, R.; Moreira, F.; Carvalho, M.A.; Cardoso, E.; Pimenta, P.; Malheiro, T.; Brito, I.; Fernandes, S.R.G.; Mesquita, D. Managing PBL difficulties in an industrial engineering and management program. *J. Ind. Eng. Manag.* **2016**, *9*, 586–611. [CrossRef]
53. Al Mughrabi, A.; Jaeger, M. Using a Capability Maturity Model in Project Based Learning. *Eur. J. Eng. Educ.* **2016**, 94–107.
54. Nelson, K.J.; Clarke, J.A.; Stoodley, I.; Creagh, T. Using a Capability Maturity Model to build on the generational approach to student engagement practices. *High. Educ. Res. Dev.* **2014**, *34*, 351–367. [CrossRef]
55. Thong, C.L.; Yusmadi, Y.J.; Rusli, A.; Hayati, A.N. Applying capability maturity model to curriculum design: A case study at private institution of higher learning in Malaysia. *Lect. Notes Eng. Comput. Sci.* **2012**, *2198*, 1070–1075.
56. Marshall, S. *New Zealand Tertiary Institution e-Learning Capability: Informing and Guiding e-Learning Architectural Change and Development*; New Zealand Ministry of Education: Wellington, New Zealand, 2006; Volume 19, 118p.
57. Maier, A.; Moultrie, J.; Clarkson, J. Assessing organizational capabilities: Reviewing and guiding the development of maturity grids. *IEEE Trans. Eng. Manag.* **2020**, *59*, 138–159. [CrossRef]
58. Klimko, G. Knowledge management and maturity models: Building common understanding. In Proceedings of the 2nd European Conference on Knowledge Management, Bled, Slovenia, 8–9 November 2001; Volume 2, pp. 269–278.
59. Kohlegger, M.; Maier, R.; Thalmann, S. Understanding maturity models results of a structured content analysis. In Proceedings of the 9th International Conference on Knowledge Management and Knowledge Technologies (I-KNOW '09 and I-SEMANTICS '09), Graz, Austria, 2–4 September 2009; pp. 51–61.
60. Shamseer, L.; Moher, D.; Clarke, M.; Ghersi, D.; Liberati, A.; Petticrew, M.; Shekelle, P.; Stewart, L.A.; The PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (prisma-p) 2015: Elabora-tion and explanation. *BMJ* **2015**, *349*, 1–25. [CrossRef] [PubMed]
61. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Int. J. Surg.* **2010**, *8*, 336–341. [CrossRef] [PubMed]
62. Chen, B.; Bastedo, K.; Howard, W. Exploring design elements for online STEM courses: Active learning, engagement & assessment design. *Online Learn. J.* **2018**, *22*, 59–76.
63. Shee, D.Y.; Wang, Y.-S. Multi-criteria evaluation of the web-based e-learning system: A methodology based on learner satisfaction and its applications. *Comput. Educ.* **2008**, *50*, 894–905. [CrossRef]
64. Evans, C.; Mujis, D.; Tomlinson, D. 2015. Available online: https://www.advance-he.ac.uk/knowledge-hub/engaged-student-learning-high-impact-strategies-enhance-student-achievement (accessed on 24 October 2021).
65. Bacon, D.R.; Stewart, K.A. Lessons From the Best and Worst Team Experiences: How a Teacher Can Make the Difference: Reflections and Recommendations for Student Teams Researchers. *J. Manag. Educ.* **2019**, *43*, 543–549. [CrossRef]
66. Donohue, S.K.; Richards, L.G. Factors affecting student attitudes toward active learning activities in a graduate engineer-ing statistics course. In Proceedings of the 39th IEEE International Conference on Frontiers in Education Conference, San Antonio, TX, USA, 18 October 2009; pp. 1–6.
67. Cochrane, T.; Antonczak, L. Connecting the theory and practice of mobile learning: A framework for creative pedagogies using mobile social media. *Media Educ.* **2015**, *6*, 248–269.
68. Hattie, J.A.C.; Donoghue, G. Learning strategies: A synthesis and conceptual model. *NPJ Sci. Learn.* **2016**, *1*, 16013. [CrossRef]
69. Rutherford, C. Using Online Social Media to Support Preservice Student Engagement. *MERLOT J. Online Learn. Teach.* **2010**, *6*, 703–711.
70. Tarantino, K.; Mcdonough, J.; Hua, M. Effects of Student Engagement with Social Media on Student Learning: A Review of Literature. *J. Technol. Stud. Aff.* **2013**, *1*, 1–13.
71. Mesquita, I.; Coutinho, P.; De Martin-Silva, L.; Parente, B.; Faria, M.; Afonso, J. The Value of Indirect Teaching Strategies in Enhancing Student-Coaches' Learning Engagement. *J. Sports Sci. Med.* **2015**, *14*, 657–668.
72. Priatna, T.; Maylawati, D.S.; Sugilar, H.; Ramdhani, M.A. Key Success Factors of e-Learning Implementation in Higher Education. *Int. J. Emerg. Technol. Learn. (iJET)* **2020**, *15*, 101–114. [CrossRef]
73. Francoise, B.; K Semsar, K.; Kennedy, S. How Not to Lose Your Students with Concept Maps. *J. Coll. Sci. Teach.* **2011**, *41*, 61–68.
74. Brophy, J.E. *Teaching*; International Academy of Education: Geneva, Switzerland; International Bureau of Education: Geneva, Switzerland, 1999.
75. Paechter, M.; Maier, B.; Macher, D. Students' expectations of, and experiences in e-learning: Their relation to learning achievements and course satisfaction. *Comput. Educ.* **2010**, *54*, 222–229. [CrossRef]
76. Park, E.L.; Choi, B.K. Transformation of classroom spaces: Traditional versus active learning classroom in colleges. *High. Educ.* **2014**, *68*, 749–771. [CrossRef]
77. Charles, E.S.; Whittaker, C.; Dugdale, M.; Guillemette, J. College level active learning classrooms: Challenges of using the heterogeneous ecology. In Proceedings of the Orchestrated Collaborative Classroom Workshop, Gothenburg, Sweden, 7 June 2015; Volume 1411, pp. 39–44.
78. Chiu, P.H.P.; Cheng, S.H. Effects of active learning classrooms on student learning: A two-year empirical investigation on student perceptions and academic performance. *High. Educ. Res. Dev.* **2016**, *36*, 269–279. [CrossRef]

79. Chiu, P.H.P.; Lai, K.W.C.; Fan, T.K.F.; Cheng, S.H. A pedagogical model for introducing 3D printing technology in a freshman level course based on a classic instructional design theory. In Proceedings of the IEEE Frontiers in Education Conference (FIE), El Paso, TX, USA, 21–24 October 2015; pp. 1–6. [CrossRef]
80. Dori, Y.J.; Belcher, J. How does technology-enabled active learning affect undergraduate students' understanding of electromagnetism concepts? *J. Learn. Sci.* **2005**, *14*, 243–279. [CrossRef]
81. Soderdahl, P. Library classroom renovated as an active learning classroom. *Libr. Hi Tech* **2011**, *29*, 83–90. [CrossRef]
82. AUSSE. *Australasian Survey of Student Engagement*; Australasian Survey of Student Engagement: Camberwell, VIC, Australia, 2010.
83. Likert, R. A technique for the measurement of attitudes. *Arch. Psychol.* **1932**, *22*, 55.
84. De Bruin, T.; Rosemann, M.; Freeze, R.; Kulkarni, U. Understanding the main phases of developing a maturity assessment model. In Proceedings of the 16th Australasian Conference on Information Systems, Sydney, Australia, 29 November–2 December 2005.
85. Shaping EDU (Arizona State University). Active Learning in Digital Realms: Capability Maturity Model. 2019. Available online: https://shapingedu.asu.edu/active-learning-digital-realms (accessed on 18 October 2021).
86. Lima, R.M. et al. A project management framework for planning and executing interdisciplinary learning projects in engi-neering education. In *Project Approaches to Learning in Engineering Education*; Brill Sense: Leiden, The Netherlands, 2012; pp. 53–76.
87. Fernandes, S.; Abelha, M.; Albuquerque, A.S.; Sousa, E. Curricular and pedagogic innovation in a social education pro-gramme: Findings from the implementation of PBL. *Int. Symp. Proj. Approaches Eng. Educ.* **2011**, *10*, 375–384.

education sciences

MDPI

Article

Can Pedagogical Innovations Be Sustainable? One Evaluation Outlook for Research Developed in Portuguese Higher Education

Cecília Guerra * and Nilza Costa

Research Centre on Didactics and Technology in the Education of Trainers, Department of Education and Psychology, University of Aveiro, 3810-193 Aveiro, Portugal; nilzacosta@ua.pt
* Correspondence: cguerra@ua.pt

Abstract: Across higher education, teachers have been inspiring themselves (and others) to apply pedagogical innovations grounded in a unifying intention: to prepare students for labour markets and societal challenges. Research-based education has been funded to promote pedagogical innovations with valuable impact on the students' academic success and/or the teachers' academic growth. However, there is still few documented research-based evidence that highlight how long-lasting and/or how embedded are the pedagogical innovations in academic practices, particularly when the funding period comes to an end. The purpose of this article is to discuss the extent to which funded national research-based education projects, developed in public Portuguese higher education institutions (universities and polytechnic institutes), have considered the sustainability of research results (e.g., pedagogical innovations), after funding ends. Based on a qualitative research approach, data collection included: document analysis of 39 projects' materials and conceptions collected with 9 coordinators and 17 key participants from selected projects, through interviews and questionnaires, respectively. Content analysis of data collected showed that there are key factors that contribute to enhance and/or limit the sustainability of research results (e.g., funding and institutional support to maintain innovation). Results and recommendations are put forward to inform educational policies, funding agencies and involved actors (e.g., institution leaders, researchers, and teaching staff) to increase the sustainability of research results.

Keywords: sustainability of research; institutional support; political and research agendas; projects dynamics

Citation: Guerra, C.; Costa, N. Can Pedagogical Innovations Be Sustainable? One Evaluation Outlook for Research Developed in Portuguese Higher Education. *Educ. Sci.* **2021**, *11*, 725. https://doi.org/10.3390/educsci11110725

Academic Editors: Sandra Raquel Gonçalves Fernandes, Marta Abelha, Ana Teresa Ferreira-Oliveira and Han Reichgelt

Received: 30 July 2021
Accepted: 4 November 2021
Published: 11 November 2021

1. Introduction

Around the academic world, the development of pedagogical innovations is increasingly a focus of institutional leaders (e.g., Rectors) and teachers [1–4], as a response to specific problems (e.g., lack of students' learning) [5]. A pedagogical innovation can be the introduction of a resource and/or strategy that, when implemented and evaluated by teachers, leads to student learning [6–12].

Teachers should be knowledgeable regarding innovative strategies and resources to promote a student-centred approach [13,14] and create a supportive intellectual and emotional environment that can encourage students to learn actively [15,16]. Some innovative strategies are collaborative learning, mentoring and tutoring, debates, role-playing, peer teaching, problem-solving [17] and formative feedback [18,19].

However, some studies [18,19] highlighted teachers' difficulties in developing innovative strategies in academic practices, such as implementing peer-written feedback in large classes (lectures), or lacking needed resources (e.g., materials and equipment). Therefore, it is also important to reflect upon what works, what does not, and options to integrate a pedagogical innovation into academic practices [20].

One possible way to innovate pedagogically is linking two of the important academic dimensions—research and teaching—in teachers' daily professional routines [21–23]. This challenge could be better achieved, as some authors [24,25] point out, through teachers' participation in research-based education projects in a scenario that also includes financial support. According to [25], the existence of financial incentives and regulations are identified as factors that promote and sustain pedagogical innovations developed during funded re-search-based education projects. There are different ways of financially supporting re-search, including loans, equity investments, award schemes and grants [26], although a grant is the known best source of funding among researchers.

Nevertheless, according to different studies [27,28], obtaining a grant to develop a research-based project, particularly in higher education, requires a considerable amount of effort and time on the part of researchers and teachers and is, in most cases, difficult to obtain [29], particularly in research-based education [27].

We find the preceding observation particularly relevant when exploring the context of Portuguese higher education. Despite the European crisis that affected Portugal during the first decades of the 21st century [30], higher education "key actors" (e.g., teachers) had several funding opportunities, launched by different Portuguese governmental entities (e.g., the Ministry of Science and Higher Education) and national sponsoring research agencies (e.g., the Calouste Gulbenkian Foundation) to apply for grants to develop research-based education projects [31]. Some funded research-based education projects have focused on the development of pedagogical innovations in universities and polytechnic institutes, revealing research evidence regarding the promotion of the teachers' academic development and/or on the students' learning development [19,31–33].

However, when the funding period comes to an end, favourable conditions are necessary to continue to implement the research-based education project and ensure the impact and sustainability of results (e.g., pedagogical innovations). The call for evidence regarding the impact and sustainability of funded research comes from sponsoring agencies (e.g., European Commission) and has been studied by several authors in different areas such as [34–36]. According to Luukkonen [37], impact is related to the effect that the activity has and its results for people, practices, organisations and systems; sustainability is the capacity of the project results and products to persist and remain in use beyond the duration of a funding period. This is what is referred to as the sustainability of pedagogical innovations developed through funded research-based projects [28].

The study of sustainability of research is a significant implementation challenge, particularly in educational research. Some authors, such as [26,38–40], have reflected on the importance of designing methodologies to assess the impact of research grant funding. Additionally, there are already several studies about the sustainability of research (e.g., [41–43]), which demonstrates the need to better understand it, namely, the importance of pursuing research-based education projects focused on the development of pedagogical innovations in higher education (HE) [44]. Nonetheless, much of the literature on sustainability remains theoretical, with little practical guidance on how to sustain research-based education project delivery, implementation strategies, and research outcomes, particularly after the end of a funding period [28,44].

The purpose of this article is to present and discuss a study focused on the understanding of the extent to which research-based education projects, sponsored by Portuguese research agencies between 2004 and 2013 and implemented in Portuguese public higher education institutions (universities and polytechnic institutes), considered the sustainability of research throughout their proposals and development process, and, if so, how.

The following research aims were defined: (1) to characterise which "sustainability actions" were included in the design and development process in the national research-based education projects selected for the study; (2) to determine which "actions" could hinder or promote the sustainability of pedagogical innovations developed through funded projects; (3) to propose recommendations for the future at the political (for politicians and research sponsors), institutional (for leaders) and individual (for teachers) levels, such that

the sustainability of the research, throughout their proposals and development process, is considered.

One considers that the answer to this question will enable the comprehension of the "sustainability actions" that could influence the design of research-based education in an HE context, and therefore give recommendations for the future at political (to politicians and sponsors), institutional (to leaders) and individual levels (to teachers).

2. Rationale

Regarding the creation process of an innovation, there is no single "true" path to develop an innovation in academic practice [45–48]. For instance, the Centre for Educational Research and Innovation proposed, some years ago, that an innovation should have two main subcomponents: what has been introduced is original to a particular individual or group, and the successful change that results from its adoption [47].

Independent of the delivery context, Serdyukov [20] explains that an innovation is understood as the successful introduction of a method, strategy or resource into someone's practice. The author also explains that it requires three major steps: an idea, its implementation, and the outcome that results from the execution of the idea and produces a change [20]. In accordance with this, [48] explained that an innovation is an idea, practice or object perceived as new by an individual, and if the idea seems new to the individual, it is an innovation. It is, then, the perceived novelty of the idea introduced, for the individual, that determines their reaction to it.

However, some obstacles could emerge during the development of pedagogical innovations in HE, such as: the powerful influence of the teachers' discomfort and/or apprehension regarding change from traditional practices to more student-centred practices [49]; an overload of time spent in preparing the new/innovative lessons [50]; the teachers' difficulties innovating in the context of large classes [19]; and the lack of resources (e.g., curricular materials, equipment, human resources) [31]. Therefore, the development of pedagogical innovations could involve a twofold challenge: on the one hand, that students will not participate in activities and/or will not learn sufficient content; on the other hand, that teachers will feel a loss of control, that they lack necessary competencies, and/or that they will be criticised for teaching in unconventional ways.

Innovations, and innovative processes, are intended to increase productivity and learning efficiency, and/or to improve learning quality [24]. Fishman et al. [45] described the process of innovation as viable, by considering the combination of three elements: adaptation to institutional context; successful enactment; sustaining innovation. According to this, one can say that, in order to innovate pedagogically in HE, teachers should look beyond what they are currently doing in their academic practice and create something different, be it in quality, quantity or both [51–54].

The maintenance, adaptation, and/or innovation of an academic practice, emerging from research-based education projects, may involve 'new forms' of work/practices/strategies that should become normal practices for the project participants (e.g., teachers, students, and others). As Costa et al. [24] points out, teachers also need to guarantee the sustainability of the innovations produced. In this context, it is important to reflect on challenges related to the development and sustainability of pedagogical innovations developed through funded research-based education projects (e.g., after the end of financial support).

Studying the sustainability of research has been a persistent challenge for the scientific community, and across a range of settings [55–63]. The term 'sustainability' is still commonly associated with environmental situations addressing subjects that are, for instance, related to the preservation of the planet (e.g., the use of natural resources by human beings). For instances, in the Brundtland Report [57], sustainability can be seen as both an effort to preserve or improve the natural environment, while also providing the means to improve the conditions of the socially and economically disadvantaged people in the world. As De Castell et al. [58] points out, a relevant assumption within the term

sustainability is the preservation or renewal of a set of resources on which social and economic development rests.

A literature review conducted by [59] categorised existing definitions of "sustainability of research" into three categories: maintaining the benefits of a project (e.g., the ability to continue the outcomes that emerge from research); project institutionalisation (e.g., ensuring that activities designed during research are formally incorporated into existing organisational practices); capacity building (e.g., to build on the capacity of the organisation's community to continue implementing the research results).

Shediac-Rizkallah et al. [59] divided the sustainability of research into three categories: maintenance of the research benefits achieved through the project; continuation of the project within an organisation; and the ability to continue to develop a project. Later, [55] considered research sustainability to be related to a process of change with the purpose of strengthening and/or sustaining a given innovation in a specific context. The author also explains that sustainable innovation can be integrated into ongoing operations to benefit diverse stakeholders.

In this vein, Southwell et al. [60] identified three interrelated tendencies of the sustainability of research, namely: research outcomes becoming standard practice of individuals and/or the organisation (e.g., an innovation is sustained); local policies, procedures, and/or structures accommodate and/or provide incentives (e.g., financial support) to adopt the practices (e.g., an innovation is embedded); and normal practice is influenced beyond its initial site of introduction in a beneficial way (e.g., scaling up an innovation into other contexts).

Sustainability requires that 'new forms of work' and the 'improvement of research results' become normal practice for the participants and/or host institutions. According to Fixsen et al. [39] (p. 17), sustainability means that "after the intensity of establishing a fully implemented evidence-based implementation program in a new community (often requiring 2 to 4 years), the implementation site needs to be sustained in subsequent years".

The existence of "program champions" in host institutions in which projects have been implemented could play a key role in the promotion of sustainable actions to promote and sustain results [59,61], even after the end of the funding period. [39] (p. 17) notes that "through it all, implementation site leaders and staff, together with the community, must be aware of the shifting ecology of influence factors and adjust without losing the functional components of the evidence-based program or the program dying due to a lack of essential financial and political support".

Stirman et al. [62] highlighted a relationship between project design, organisational systems, connection with the community, and funding support. Based on a reviewing article identified through the four knowledge syntheses of sustainability, Moore et al. [56] (p. 7) developed a comprehensive definition of sustainability that included five constructs: "(1) after a defined period of time, (2) the program, clinical intervention, and/or implementation strategies continue to be delivered and/or (3) individual behaviour change (i.e., clinician, patient) is maintained; (4) the program and individual behaviour change may evolve or adapt, while (5) continuing to produce benefits for individuals/systems".

Upon critically analysing different authors' perspectives of the concept of research sustainability, and subsequent reflection, Costa et al. [24] (p. 117) explained that sustainability could be related to "the continuation of projects' components (e.g., the innovative practice developed), capacity building (e.g., human and financial resources), and continued benefits or outcomes in the same and/or new settings (e.g., new educational scenarios and participants)".

Notably, the concept of sustainability of research aids understanding how a particular instance of pedagogical innovation developed though a funded research-based education project could continue and/or evolve, over a certain period, after the end of the funding period. Thus, one can see the potential of these definitions to reflect on the sustainability of research-based education projects. However, much of the scientific literature remains theoretical, with little practical guidance on how to sustain research-based education

project delivery, implementation strategies, and outcomes (e.g., pedagogical innovations in academic practice).

In this context, some authors [24,28,44,63] have borrowed the idea of sustainability to understand how pedagogical innovation is developed and sustained after the end of the funded period of research-based education projects. This option could be justified mainly given its focus on the relational aspects of the development process of pedagogical innovations (application, implementation, and the post-funding period).

In this context, Guerra [28] (p. 10) puts forth a definition of the sustainability of pedagogical innovations developed through funded, research-based education projects in HE: "The sustainable use of pedagogical innovations [. . .] is considered to be achieved when after a defined period of time, they continue to be delivered in higher education, with or without adaptations, in the same academic scenery (e.g., course or curricular unit), with the aim of producing benefits for individuals (e.g., students' academic success and/or teachers' academic growth) and/or systems (e.g., institutions, courses, curricular units)".

In research-based education projects, sustainability can mean, for example, the use of a particular infrastructure, continuation of services, mobilisation of participant competencies, development and/or continuation of partnerships and collaborations, and the integration of ideas and/or perspectives by the host institution (e.g., universities). The literature highlights different levels of effort by diverse key actors in the HE context, which may or may not guarantee the sustainability of research results.

Research-based education should be developed, disseminated, and exploited in such a way that: results can be tailored to the needs of others (e.g., teachers); transferred to new areas (e.g., curricular units); sustained after the funding period has finished; and/or used to influence future policy and practice. Therefore, the sustainability of funded research-based education projects require the recognition of different levels of key actors' actions in HE spheres, namely: at the political level (macro level, e.g., the politicians that define the funding research agendas and/or regulations to empower innovation through funded research); at the institutional level (meso level, e.g., the institutional support of HE leaders to embed research output in collective dynamics); and at the individual level (micro level, e.g., personal and academic teachers' motivations and the resilience to sustain, embed and/or upscale outcomes in practices).

3. Methods

A qualitative research paradigm [64–68], adapting principles of a case study approach [67], was adopted during two research phases: the first aimed to characterise the national funded research-based education projects developed in Portuguese public HE institutions (PPHEIs), from 2004 to 2013; the second aimed to determine the macro, meso and micro actions (e.g., institutional structures, personal dynamics) that hinder or promote the sustainability of effective pedagogical innovations (before, during, and after the end of the funding period) in science and engineering fields. Table 1 presents the data collection procedures, purposes, and aims applied in the two research phases of the study.

Table 1. Phases, data collection, and aims of the study.

	Data Collection	Aims
1st phase	Conducting a document analysis of 39 projects	To characterise the national funded research-based education projects developed in Portuguese public HE institutions (PPHEIs), from 2004 to 2013
2nd phase	Performing individual interviews with 9 coordinators of the projects	To determine the macro, meso and micro actions (e.g., institutional structures, personal dynamics) that hinder or promote the sustainability of effective pedagogical innovations (before, during and after the end of the funding period) in science and engineering fields
	Inquiring 17 key participants involved in projects through an online questionnaire	

3.1. Document Analysis of National Funded Research-Based Education Projects (First Phase)

In the first phase, a qualitative document analysis [69] was adopted to characterise the national funded research-based education projects developed in Portuguese public

HE institutions (PPHEIs) from 2004 to 2013. This technique enabled a rigorous, impartial, and systematic analysis of the contents of collected written documents. This process of document analysis occurred in the following steps:

1. Collection of documents: the definition of the criteria for establishing the corpus of analysis—the funded research projects—considered: the research setting (studies should be implemented in Portuguese public higher education institutions (PPHEIs), i.e., universities and/or polytechnic institutes); the research object (studies should be focused on the development of pedagogical innovations in HE courses, e.g., science and engineering courses); the research period (national projects funded between 2004 and 2013); financial support (three main research sponsors in Portugal, namely, the Foundation for Science and Technology (FCT), the Ministry of Science, Technology and Higher Education (MSTHE), and the Calouste Gulbenkian Foundation (CGF)).

Regarding the identification of studies found on each research sponsor's website, it was important to clarify that each national sponsor has a "funded projects database of all scientific domains". All collected documents were in the public domain, although not all appropriate documents were available online. The team requested such documents from its contacts within the organisations.

Only 70 documents that were disclosed in the public domain were collected: 33 projects funded by FCT, specifically in the field of "Education and Educational Policies"; 21 projects funded by the MCTES (the Science and Innovation Operational Programme (POCI 2010), under the Axis IV-Science, higher education, As IV. 1-higher education Qualification, action IV. 1.2-innovative projects in higher education); 16 projects funded by MSTHE in the framework of "Innovative Educational Projects (2009/2015)". Table 2 presents a summary of search query techniques adopted on each research sponsor's websites (FCT, CGF, and MSTHE), as well as the identified funded projects.

2. Eligibility of documents: after the selection of 70 projects, 64 coordinators were contacted by email (November 2015), asking for their collaboration in this study and requesting their authorisation for data collection, such as reports of activities and scientific publications of the projects.

Table 2. Search queries applied in the three national research sponsors databases.

Queries	Sponsors	Studies
"higher education" OR "university" OR "polytechnics" AND year > 2004 < 2013	FCT	33
	MSTHE	21
	CGF	16
Total		70

A total of 36 coordinators agreed to collaborate in this study and sent a total of 39 project documents (e.g., funding applications, scientific reports and/or books and articles). In this way, 39 studies were eligible for data collection in the first phase.

Researchers preserved anonymity by referring to the projects and project coordinators as follows: [Coordinator POSITION and Proj. NUMBER]. For example, [Proj. 1.A] refers to Coordinator A and Project 1.

Each project's document was screened (by reading titles, research team, host institution, keywords, and abstracts) and analysed through a content analysis technique [69–71], using qualitative analysis software (WebQDA: https://www.webqda.net/?lang=en, accessed on 1 July 2021). Five queries (who, what, where, who, when) were defined to characterise the national funded research-based education projects developed in Portuguese public HE institutions (PPHEIs), from 2004 to 2013 (Table 3).

Table 3. Inquiries applied to data collected in the 1st phase.

Document Analysis	
Queries	**Indicators of Analysis**
WHO has afforded the development ofpedagogical innovations?	(political and research sponsors' agendas)
WHAT has developed with this research?	(e.g., activities, the resources)
WHERE did the research take place?	(e.g., the host institution support)
WHO has conducted these projects?	(e.g., team elements)
WHEN did the project start and finish?	(e.g., the length of the financial support)

Each document was analysed to determine the extent to which the project it described addressed or considered each of the identified indicators for characterising the projects (Table 3). Text relevant to each indicator was highlighted and coded using qualitative data analysis software (WebQDA), based on the meaning, relevance and context for each indicator, as 'yes', 'no', or 'unclear'. Scores of 'yes', 'no', or 'unclear' were assigned numerical values (1 to 3) to assist in aggregation and data presentation. The documents served as the individual source of information for the scoring, which facilitated research objectivity.

Each coordinator of the selected projects (n = 39) was involved in order to verify our coding of each project's document. In addition, a third person (researcher in education) provided ad hoc verification and also served as an external mediator for any inconsistencies between the two primary coders. This ensured robust interpretative analysis and conclusions [64].

This first phase served as a baseline, reflecting on the theory and practice of funded research-based education projects in PPHEI, rather than an evaluation of a project. Results are presented in the following section.

3.2. Perceptions of Coordinators and Key Participants of Selected Projects (Second Phase)

In the second phase, the focus was to determine the macro, meso and micro actions (e.g., institutional structures and personal dynamics) that hinder or promote the sustainability of effective pedagogical innovations (before, during, and after the end of the funding period) in science and engineering fields.

In this way, conceptions of 9 coordinators of 12 funded projects were collected, through individual interviews (from January 2017 to February 2018). The group of coordinators was organised considering the following criteria: include coordinators with more than one funded project (e.g., different research sponsoring agencies); incorporate coordinators of projects that finished in different periods, specifically, 1–5 years after the end of the project (medium term); more than 5 years after the end of the project (long term); involve coordinators with projects implemented in the scientific areas of science (biology) and engineering; include coordinators with management responsibilities in the host institutions of the projects (e.g., vice-chancellors of universities).

Two open questions were presented to coordinators (interviews): To what extent were the pedagogical innovations developed during the funded research-based education projects adopted, adapted and/or reinvented, after the end of the funded period?; and What were the main factors that have contributed to the sustainability of pedagogical innovations, mainly at the macro (political and research), meso (institutional) and micro levels (projects' dynamics)?

During the interviews with the 9 coordinators, we asked them to provide us with the contacts of some of their team members—the key participants—to deepen our understanding of the "sustainability actions" applied at individual level (MICRO level). An online questionnaire was designed with a set of questions that all key participants were asked to complete. The aims of the questionnaire were to: check the type of involvement of the key participant in the project; characterise the pedagogical innovation(s) developed and its

operationalisation; determine the type of participation, the level of autonomy, the importance and effectiveness (impact) of participation in the context of the selected project(s); and understand the dynamics that occur after the end of the projects, with regard to the implementation, adaptation and/or reinvention of pedagogical innovations, as well as the reasons underlying them (e.g., where were they implemented, what were barriers, lessons learned and facilitators). A total of 40 key participants were contacted (40 key participants), but only 17 key participants responded to the online questionnaire (Table 4).

Table 4. Corpus collected in the 2nd phase (interviews and questionnaires).

Host	**U_1**		**U_2**		**U_3**		**U_4**		**U_5**		**I_1**	**I_2**	**Total**
Project	**Proj. 1A**	**Proj. 2A**	**Proj. 3B**	**Proj. 4C**	**Proj. 5C**	**Proj. 6D**	**Proj. 7F**	**Proj. 8F**	**Proj. 9G**	**Proj. 10G**	**Proj. 11H**	**Proj. 12I**	**12 P**
Year	2004	2010	2009	2010	2013	2012	2010	2012	2008	2010	2010	2011	
Sponsor	FCT	FCT	FCT	MSTHE	CFG	CFG	CFG	CFG	FCT	FCT	MSTHE	CFG	
Coorddinator (expertise)	A (Edu)		B (Edu)	C (Eng)	D (Eng)	E (Eng)	F (Eng)		G (Eng)		H (Psy)	I (Eng)	9 I
Key particpants	3		2		7		1		2		2	0	17 Q

Legend: University 1 (U_1); Institute 1 (I_1); Education (Edu); Engineering (Eng.); Psychology (Psy); Interviews (I); Online questionnaire (Q)

Notably, seven key participants participated in more than one funded research-based education project [Q1, Q2, Q3, Q7, Q10, Q13, Q15] (Table 5).

Table 5. Corpus collected in the 2nd phase (questionnaires).

Host	**U_1**		**U_2**		**U_3**		**U_4**		**U_5**		**I_1**	**Total**
Q	**Proj. 1A**	**Proj. 2A**	**Proj. 3B**	**Proj. 4C**	**Proj. 5C**	**Proj. 6D**	**Proj. 7F**	**Proj. 8F**	**Proj. 9G**	**Proj. 10G**	**Proj. 11H**	
Q1	x	x										2
Q2	x	x										2
Q3	x	x										2
Q4			x									1
Q5			x									1
Q6				x								1
Q7				x	x							2
Q8					x							1
Q9					x							1
Q10					x	x						2
Q11					x							1
Q12					x							1
Q13							x	x				2
Q14										x		1
Q15									x	x		2
Q16											x	1
Q17											x	1

The analysis of the perceptions of coordinators and other key participants (e.g., teaching staff and students involved in these innovations) helped to explore specific "sustainability actions" applied before the implementation of studies (e.g., in proposals), during the implementation of studies (e.g., during funding), and at the end of studies (e.g., after the financial support ends).

A content analysis technique [69–71], using WebQDA software, was applied to the data collected. Qualitative analysis involved specifying characteristics of a participant's statement (coordinators and key participants), coding them, counting occurrences of the coded categories, and subsequently using descriptive techniques to analyse the data. In effect, the coded categories were treated as variables. The content analysis procedure was similar to the first phase, the distinction being whether the data were used to attempt to interpret meanings or focus on identifying dominant tendencies of "research sustainability actions". Indicators of analysis have emerged from the authors' conceptualisations of pedagogical innovations [24], the sustainability of research [28,42], and from the content analysis process [69–71].

Three main dimensions were considered in this second phase: the Portuguese political and research agendas to produce pedagogical innovations in HE and their sustainability (MACRO level); the host institution's contribution to support pedagogical innovations in HE and their sustainability (MESO level); the project dynamics, through individuals promoting and sustaining pedagogical innovations in HE (MICRO level).

4. Results

The goal for the first phase was to characterise the national funded research-based education projects developed in Portuguese public HE institutions (PPHEIs), from 2004 to 2013; the goal of the second phase was to determine the macro, meso and micro actions (e.g., institutional structures and personal dynamics) that hinder or promote the sustainability of effective pedagogical innovations (before, during, and after the end of the funding period) in science and engineering fields.

Information from the three main sources of data (projects documents, interviews with project coordinators, and online questionnaires from key participants in projects) is given in the following sections, and thus triangulated in order to enhance the credibility of findings [65]. Results are presented and discussed below, considering each research phase.

4.1. Document Analysis of the Research-Based Education Projects

Table 6 presents the funded research-based education projects promoted in the Portuguese public higher education context by year and funding source.

Table 6. Projects categorised by year and funding source.

	Research Sponsor			
Funding Year	**FCT**	**MSTHE**	**CGF**	**Total**
[2004- ...]	3	0	0	3
[2005- ...]	0	0	0	0
[2006- ...]	9	0	0	9
[2007- ...]	0	0	0	0
[2008- ...]	3	0	0	3
[2009- ...]	3	0	0	3
[2010- ...]	3	10	3	16
[2011- ...]	0	0	1	1
[2012- ...]	0	0	2	2
[2013- ...]	0	0	2	2
Included studies	21	10	8	39

In the defined period of analysis (2004 to 2013), HE institutions and academics (e.g., teachers) had several opportunities, launched by different Portuguese Governments, to apply for grants to develop research-based education projects in the Portuguese context.

However, each research sponsoring agency has its specific research purposes in its "open call applications".

Both the MSTHE and the CGF had specific calls to grant research-based education focused on the promotion of pedagogical innovations in Portuguese HE. For example, the Portuguese Ministry of Higher Education (MSTHE) has funded ten research-based education projects, particularly in 2010. From 2010 onward, the FCG has been the main research sponsor of promoting pedagogical innovations, through eight research-based education projects in PPHEIs.

If we look to the kind of calls promoted by each research sponsor, we can highlight a more assertive concern, especially from 2010 onward, of supporting research-based education projects, intentionally focused on the promotion of students' academic success and/or teachers' academic growth. Both the FCG and the MSTHE included "Innovation in Higher Education" as keywords in their open research calls. We believe that this certainly affected the coordinators' options regarding the kind of projects for which they applied for financial support during the selected period.

One reason for the Portuguese political and research agenda's alignment to fund research-based education projects in HE can be explained, for example, by the political scenario of HE during the analysed period (e.g., during the Bologna Process' implementation). In fact, Tavares et al. [72] explained that Portuguese HE academics have combined to change their pedagogical practices to a more student-centred approach, and this was a huge academic challenge, despite the European crisis that affected Portugal, especially from 2008 to 2010 and then going forward [30].

Results also reveal that 2005 and 2007 were weak years, in terms of the financial support of research-based education projects in HE. The "Bologna Process influence" emerged, in the document analysis, as a contextualisation reason for academics to apply for research funds to develop pedagogical innovations in HE, which uncovers an alignment between political and research agendas (in the FCT, CGF, and MSTHE).

MSTHE-funded projects (10) and CFG-funded projects (8) were considered short empirical studies (one-year period of funding), whereas the FCT-funded projects (21) were considered long empirical studies (two- or three-year periods of funding). However, research-based education projects need time to identify a specific educational problem, to look for solutions, to implement and evaluate them, and assure their sustainability, after the end of the funding period. Consequently, funded research-based education projects aiming to develop pedagogical innovation in academia will require more than one year.

Ref. [31] explains that one possible explanation for difficulties in promoting the "sustainability of research" could be related to the length of funding projects. Certainly, to open calls for short periods of time does not allow much more than to understand a particular phenomenon in HE (e.g., to identify factors for the students' academic difficulties).

Results show that the 39 projects were developed in 15 PPHEIs, of two types: universities (9) and polytechnic institutes (7). Despite the tendency of growing convergence, partially due to the Bologna implementation process, polytechnic institutes and universities have specific missions and distinct purposes. Universities have a strong conceptual and research-oriented focus, and polytechnics are known for offering a more practical and vocational-oriented education [73].

Figure 1 systematises the "scientific areas" of curricular units and courses in which the included funded research-based education projects were implemented. This categorisation considered the "year of open call for research".

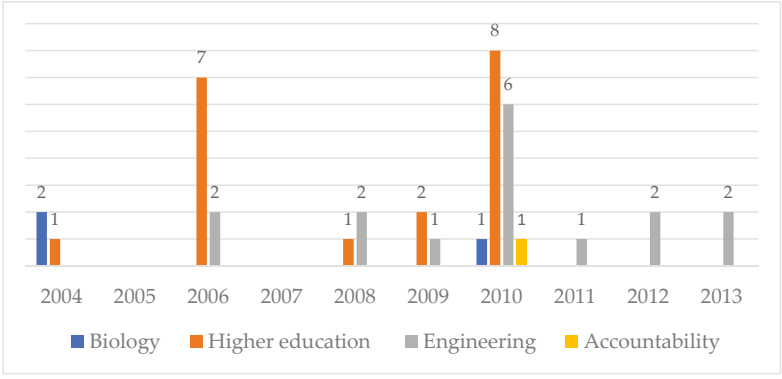

Figure 1. "Scientific areas" of curricular units and courses.

A total of 19 projects were categorised in "HE", with the majority funded by the FCT, and with most funded in 2006 and 2010. These funded projects mainly focused on the comprehension of a broad spectrum of PPHE subjects, such as the study of: professional fulfilment and motivation of teachers (Proj. 20); factors for the students' academic success (Proj. 21); and women's roles in science and engineering courses (Proj. 22).

The remaining projects (20) highlighted pedagogical innovations implemented in engineering courses (16), biology courses (3) and an accounting course (1). Some funded projects were, for instance, focused on the promotion of teachers' academic growth (e.g., Proj. 2.A) and/or on students' active learning (e.g., Proj. 11.H).

To sum up, document analysis of the funded projects suggested awareness in collating key participants of the host institutions (e.g., teachers and students) in order to assure the success of the project, namely, the promotion of an innovative strategy, even after the end of funding. This can also be a consequence of the kind of funding call open at the time (e.g., FCG and MSTHE, as already mentioned), and/or the overall conditions for the host institution's support (e.g., involving "program champions").

4.2. Coordinators and Key Participants of Selected Projects

Concerning the Portuguese political and research agendas to afford pedagogical innovations in HE and their sustainability (MACRO level), results show that all coordinators explained the importance of having funding to follow up projects.

Additionally, all coordinators mentioned the "length and extent of the financial support" as a "sustainability action", which could compromise the dissemination and exploitation of the results, particularly the pedagogical innovations' applicability, after the ending of the funded period.

For instance, Coordinator A had developed two sequential projects (Proj. 1.A and Proj. 2.A), specifically focused on the academic development process of biology teachers. This coordinator applied for new research funds to promote the continuation of the previously developed pedagogical innovations (e.g., questioning strategy in genetic curricular units) (Proj. 1.A) but involving other biology teachers (Proj. 2.A), as expressed by her: "In general, we are more successful [in sustaining pedagogical innovation] when the projects are interconnected [. . .] where there is an articulation between approved proposals [. . .] for instance, my project's financial period ends, I win another one but with different goals" (Proj. 1.A and Proj. 2.A). Another example was presented by Coordinator 7, who had two funded follow-up projects (Proj. 7.F and Proj. 8.F), which allowed the research team to use and sustain different virtual laboratories in the engineering courses.

Probably, the coordinators had accomplished this articulation to be successful in their application for research funds, although they did not assume it in the interviews. This inconsistency may suggest a culture of disarticulation in the alignment between

political and research agendas. [74] (p. 427) infers that teachers tend to assume "technically unarguable, and socially and politically neutral" positions. This reflection aligns with those coordinators' statements, which could lead to the development of pedagogical interventions in HE that do not actually include the challenges and principles that emerge from policies, namely, the Bologna Process.

In what concerns "the host institutions' contribution to supporting pedagogical innovations in HE and their sustainability" (MESO level), all the interviewed coordinators highlighted the importance of coordination between the research design (e.g., aims and methods), and the "institutional strategy of the host institution" to promote pedagogical innovations. The coordinator's leadership in benefiting the host institution leaders contributed to the promotion, and even sustainability, of research, as expressed by this coordinator: "The first contact we did was with the Head of the Departments the President of the Pedagogic Assembly/PA [. . .] therefore they have been always involved [. . .] this dynamic causes decisions made by the PA President to nominate, for example, a person responsible to collaborate with us in the Departments [. . .]" (Proj. 3.C).

Despite the HE organisational institutions' differences (universities and polytechnic institutes), already discussed, all coordinators agreed on the existence of "special elements" in the host institution, because they had the authority and the capacity to meet the commitments in order to achieve the project goals. Those elements had different roles in the host institutions (e.g., directors of HE courses), and were able to motivate their peers (e.g., directors, teachers, etc.) of the courses to be involved in the development process for innovation. Authors call these elements the "program champions" [42,61], highlighting their relevant role in project goals' achievement and sustainability.

The results revealed that two host institutions, both universities, have created "staff development offices" to support the development of pedagogical innovations, after the end of relevant funding periods. One coordinator explained that, when the funding period of a project finished, "the development of tutorial sessions with teachers and students of the host institution" was maintained. Another coordinator explained that the staff development office, installed in an engineering faculty, has enabled: (a) the maintenance of educational resources; and (b) retaining human resources (e.g., research fellows) to support teachers during impact evaluation of the implemented pedagogical innovations. Thus, outcomes emerged from the funded research-based education projects that have contributed to sustainability for pedagogical innovations.

Regarding "the influence of project dynamics to promote and sustain pedagogical innovations" (MICRO level), with respect to the choice of the research team, partnership between different profiles of team members also seemed to be important. For example, all coordinators have highlighted the importance of involving multidisciplinary teams with educational researchers and engineering teachers at the same university.

According to the opinion of the six coordinators (Proj. A.1, Proj. A.2; Proj. B.3; Proj. C.4; Proj. D.5; Proj. E.6; Proj. H.11), during this interdisciplinary collaboration (a) to prepare and/or reflect on the implementation (e.g., the impact of pedagogical innovations in teachers' academic growth and/or students' learning success); (b) the teachers were more focused on their academic practices (e.g., to design and implement pedagogical innovation in curricular units); (c) the educational researchers were responsible to collect and analyse data to monitor and evaluate the impact of the implemented pedagogical innovations.

As Pedrosa-De-Jesus et al. [19] highlighted, it is important to create institutional opportunities to enhance collaboration between teachers and educational researchers, particularly through involvement in research-based education projects. For that to happen, these "key actors" should become "peers" in conceiving and evaluating pedagogical innovations [31].

All coordinators clarified that they published the research results in journals, mostly in English (the common language of the scientific community). This certainly could constitute a constraint on that sustainable channel—science-related communication. Knowledge mobilisation in the social sciences, such as in educational sciences, could be an action

that opens results to a wider audience and contributes to the sustainability of research outputs [75].

As presented in our rationale, currently, the research sponsors (e.g., European Commission) have been asking for the inclusion of "dissemination and exploitation plans for funded research" in order to guarantee: the impact of research on people, practices, organisations, and systems (research impact); and the use of research results beyond the term of the funding period (sustainability of research).

However, all coordinators emphasised that they are usually evaluated concerning scientific production in their areas of expertise (i.e., in biology or engineering), and not so much by the pedagogical innovations developed throughout research-based education projects.

The 17 key participants explicitly confirmed, in the online questionnaire, that the pedagogical innovations had been sustained when the initial funded research came to an end. For instance, key participants 1 and 2 participated in both Proj. 1.A and Proj. 2.A.

As we noted above, Coordinator A's leadership in promoting two funded research-based education projects has contributed to the promotion and the sustainability of research results. Thus, this group of biology teachers was engaged in the development of pedagogical innovations in their academic practice.

In terms of the benefits, the key participants said that the main gains of their participation in the two funded projects were at two levels. First, at the research level, the key participants from biology [Q1, Q2] highlighted two important benefits: implementing pedagogical innovations in their classes as a result of research-based education (e.g., written formative feedback); contact with other scientific cultures, namely, educational research; and reflecting on their academic growth.

Second, at the academic level, other key participants from engineering [Q7, Q10, Q13, Q15] highlighted other potentialities, such as: designing and integrating various web-related technologies to teach and learn in engineering courses (e.g., virtual laboratories in practical lessons); creating supporting pedagogical documentation (e.g., online tutorial videos in theoretical lessons); and involvement in communities of practice of education of engineering.

Beyond the funding of those projects, one key participant from the engineering area [Q15] said that belonging to a community of practice of education in engineering allowed them to sustain the pedagogical innovations developed through previous funded research (e.g., remote laboratories) in other higher education institutions (e.g., in Brazilian universities).

Nevertheless, the key participants identified constraints to sustain the innovations after the end of the funding period, which emerged within the borders of the host institution where those projects were implemented, such as: the existence of leaders' misconceptions of the conceptual matrix of the project [all Q]; the difficulty of reaching and convincing more colleagues to (re)use pedagogical innovations in their academic practices [Q4]; and the difficulty of maintaining technical and operational remote laboratories, without human resources capacitated for that [Q15 and Q16].

5. Conclusions

This study helped to characterise which "sustainability actions" have been included in the design and development process in national research-based education projects selected in the study and to identify which actions could hinder or promote the sustainability of pedagogical innovations developed through funded projects. The development process of sustaining pedagogical innovations emerging from funded research-based education projects in HE is complex, due to the multiple factors involved (political, institutional, and individual).

Results also demonstrated that, in some cases, the pedagogical innovations developed during funded research-based education projects were sustained by the key participants after the end of the funding period. In those cases, key factors resulted in promoting innovations and sustaining pedagogical innovations in academic practice over the years.

Results show how "sustainability actions" were applied in projects, namely: in applying for project funding (before the funding period); in the project's implementation (during the funding period); and after the project's implementation (after the funding period).

Although we cannot generalise these outcomes to PPHEIs, due to the qualitative and non-representative sample of funded research in the analysed period (2004 to 2013), the study allowed us to propose some recommendations for the future at the political (recommendations aimed at politicians and sponsors), institutional (aimed at leaders) and individual levels (aimed at teachers); these recommendations impact how the sustainability of research is considered, throughout their proposals and development processes.

The political agendas should not only afford the development of pedagogical innovations in HE, but also guarantee that the project design proposal is integrated from the beginning with "sustainability actions" to disseminate and/or exploit research results, even without funding. A research funding agenda should open calls that explicitly define the goals to be achieved—produce innovations in HE (such as those of CGF) and require plans for sustainability (such as those from the European Commission). However, the funding period of the research needs to be adequate for such an endeavour. Short periods (e.g., one year) do not seem to enable conceiving, implementing, and assessing innovation in HE.

Decisions at the institutional level should be aligned with national and international political and research agendas, particularly those that are specifically related to the importance of promoting pedagogical innovations. For that, it is important that host institutions (e.g., universities) continue to provide "internal support", not necessarily just funds, to teams that have already produced knowledge from previous funded research-based education projects. Different institutional initiatives could be implemented, such as: making diverse institutional resources (e.g., the creation of "staff development offices") available to these research teams; giving internal extra funding for the development of follow-up research-based education projects; and valuing the involvement of research-based education projects in teachers' professional evaluation process (e.g., investing in "staff development of teachers").

Finally, the projects' internal dynamics are closely related to the coordinator's leadership, who may apply different "sustainability actions" during the development process of the project, namely: in the proposal project, where sustainability should be incorporated (e.g., by empowering the participants teachers, by planning effective dissemination of the project results); and during the project implementation, project coordinators should create an environment that ensures the participation of teachers in critically reflecting on their practices, prepare the conditions for the continuation of innovation after the funding period (e.g., negotiate such conditions at an institutional level), including "program champions" (e.g., directors of courses where the innovation will be implemented) and members with different profiles (e.g., educational researchers and teachers in line with the specialties of the courses where the innovation will occur). Nevertheless, projects should be developed in such a way that results/outcomes (e.g., pedagogical innovations) can be tailored to the needs of others, transferred to new areas, sustained after the funding period has ended and/or be used to influence future policy and practice.

The "sustainable actions" identified in our study can provide guidelines for the conditions necessary for the sustainability of research-based education projects, and expand their scope to other HE scenarios (course units, courses and/or institutions). In particular, this study reveals that different "key actors" are involved in the creation and sustainability of pedagogical innovations.

Although results cannot be generalised statistically—because generalisations are made for theory and not for a population [67]—the identified "sustainability actions" could be performed, as [24,25,28,31,44] highlight, by key actors of three higher education spheres—macro (policy and decision-makers in research-sponsoring agencies), meso (institutional leaders), and micro (teamwork dynamics). Each "key-actor" was linked to

sustainability factors that have influenced (or not influenced) the sustainability of the research, namely:

- In the national research agencies (e.g., funding lines focused on pedagogical innovations in higher education);
- The host institutions (e.g., institutional leaders who have responsibilities in curricular development);
- The projects (e.g., researchers and teachers who, in the light of the policy guidelines, aim to introduce pedagogical innovations in their academic practices).

At an upper level of sustainability of research, upscaling it is necessary, meaning, for example, dialogue with diverse "key actors" to reinforce definitions from political and research agendas, so that research can be developed in a way so as to promote pedagogical innovation developed through research-based education projects in HE, and requiring its sustainability.

Ultimately, what is important is the presence of leadership competencies of "key actors" in each higher education sphere, because each "key-actor" could play a different, but relevant, role in influencing, promoting, and sustaining pedagogical innovations through educational research.

Author Contributions: Conceptualisation, C.G. and N.C.; methodology, C.G. and N.C.; validation, C.G. and N.C.; formal analysis, C.G.; investigation, C.G.; resources, C.G.; data curation, N.C.; writing—original draft preparation, C.G. and N.C.; writing—review and editing, C.G. and N.C.; visualisation, C.G. and N.C.; supervision, N.C.; project administration, C.G.; funding acquisition, C.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work is financially supported by National Funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., under the project UIDB/00194/2020 (CIDTFF) and the work of the first author is funded by national funds (OE), through FCT – Fundação para a Ciência e a Tecnologia, I.P., in the scope of the framework contract foreseen in the numbers 4, 5 and 6 of the article 23, of the Decree-Law 57/2016, of August 29, changed by Law 57/2017, of July 19.

Institutional Review Board Statement: Ethical review and approval were waived for this study, due to being a study involving a small number of healthy adults, participating under informed consent, and with no sensitive data collection.

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

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy issues.

Acknowledgments: We acknowledge all the institutions and colleagues who, very kindly, prepared and sent us the cases of good practice. The authors thank the participating teachers for accepting taking part of this study. The authors also thank the reviewers for their insightful comments.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Major, J.; Tait-McCutcheon, S.L.; Averill, R.; Gilbert, A.; Knewstubb, B.; Mortlock, A.; Jones, L. Pedagogical Innovation in Higher Education: Defining What We Mean. *Int. J. Innov. Teach. Learn. High. Educ.* **2020**, *1*, 1–18. [CrossRef]
2. Hero, L.-M.; Lindfors, E.; Taatila, V. Individual Innovation Competence: A Systematic Review and Future Research Agenda. *Int. J. High. Educ.* **2017**, *6*, 103. [CrossRef]
3. Konst, T.; Kairisto-Mertanen, L. Developing innovation pedagogy approach. *Horizon* **2020**, *28*, 45–54. [CrossRef]
4. Kim, J.; Maloney, E.J. *Learning Innovation and the Future of Higher Education*; JHU Press: Baltimore, MD, USA, 2020.
5. Tight, M. Student retention and engagement in higher education. *J. Furth. High. Educ.* **2019**, *44*, 689–704. [CrossRef]
6. Biggs, J. *Teaching for Quality Learning at University*; Open University Press: Philadelphia, PA, USA, 1999.
7. Dijkstra, I.S.; Pols, J.; Remmelts, P.; Rietzschel, E.F.; Cohen-Schotanus, J.; Brand, P.L. How educational innovations and attention to competencies in postgraduate medical education relate to preparedness for practice: The key role of the learning environment. *Perspect. Med. Educ.* **2015**, *4*, 300–307. [CrossRef] [PubMed]

8. Getz, M.; Siegfried, J.J.; Anderson, K.H. Adoption of innovations in higher education. *Q. Rev. Econ. Financ.* **1997**, *37*, 605–631. [CrossRef]
9. Hoidn, S.; Kärkkäinen, K. *Promoting Skills for Innovation in Higher Education: A Literature Review on the Effectiveness of Problem-based Learning and of Teaching Behaviours*; OECD Education Working Papers, No. 100; OECD Publishing: Paris, France, 2014. [CrossRef]
10. Nevgi, A.; Löfström, E. The development of academics' teacher identity: Enhancing reflection and task perception through a university teacher development programme. *Stud. Educ. Eval.* **2015**, *46*, 53–60. [CrossRef]
11. Tabata, L.N.; Johnsrud, L.K. The impact of faculty attitudes toward technology, distance education, and innovation. *Res. High. Educ.* **2008**, *49*, 625–646. [CrossRef]
12. Zeegers, P. A revision of the biggs' study process questionnaire (R-SPQ). *High. Educ. Res. Dev.* **2002**, *21*, 73–92. [CrossRef]
13. Herodotou, C.; Sharples, M.; Gaved, M.; Kukulska-Hulme, A.; Rienties, B.; Scanlon, E.; Whitelock, D. Innovative pedagogies of the future: An evidence-based selection. *Front. Educ.* **2019**, *4*, 13. [CrossRef]
14. Mendonça, M.; Popov, O.; Frånberg, G.-M.; Cossa, E.F.R. Introducing a student-centred learning approach in current curriculum reform in mozambican higher education. *Educ. Inq.* **2012**, *3*, 37–48. [CrossRef]
15. Erol, M.; Özcan, A.; Luft, J.A. Six reasons to teach undergraduate courses in SCALE-UP classrooms: suggestions for higher education authorities and instructors. *Educ. Sci. Egit. Ve Bilim* **2016**, *41*. [CrossRef]
16. Liu, Q.; Geertshuis, S.; Grainger, R. Understanding academics' adoption of learning technologies: A systematic review. *Comput. Educ.* **2020**, *151*, 103857.
17. Savery, J.R. Overview of problem-based learning: Definitions and distinctions. *Interdiscip. J. Probl. Learn.* **2006**, *1*, 3. [CrossRef]
18. Pedrosa-de-Jesus, H.; Guerra, C.; Feehily, R.; Williams, H.; Kyomuhendo, F.C.; Seeam, A.K.; Congo-Poottaren, N.; Beebeejaun-Roojee, S.; Betchoo, N.K.; Lun, K.C.K.T.; et al. Teachers' written formative feedback on students' critical thinking: A case study. *MIE J. Educ.* **2018**, *9*, 3–22.
19. Pedrosa-De-Jesus, H.; Guerra, C.; Watts, M. University teachers' self-reflection on their academic growth. *Prof. Dev. Educ.* **2016**, *43*, 454–473. [CrossRef]
20. Serdyukov, P. Innovation in education: What works, what doesn't, and what to do about it? *J. Res. Innov. Teach. Learn.* **2017**, *10*, 4–33. [CrossRef]
21. Cleaver, E.; Maxine, L.; McLinden, M. *Teaching and Learning in Higher Education: Disciplinary Approaches to Educational Enquiry*, 2nd ed.; Sage Publishing: London, UK, 2018.
22. Schratz, M. Researching while teaching: An action research approach in higher education. *Stud. High. Educ.* **1992**, *17*, 81–95. [CrossRef]
23. Tong, V.C.H.; Standen, A.; Sotiriou, M. *Shaping Higher Education with Students: Ways to Connect Research and Teaching*; UCL Press: London, UK, 2018.
24. Costa, N.; Guerra, C. Concluding Remarks: Contribution of the Book for Future Directions to Advance the Sustainability of Pedagogical Research and Practice. In *Sustainable Pedagogical Research in Higher Education*; Routledge: Oxford, UK, 2020; pp. 113–118.
25. Guerra, C.; Franco, A. *Seabra, M. Sustainable Pedagogical Research in Higher Education: The Political, Institutional and Financial Challenges*; Seabra, M., Ed.; Routledge: Oxford, UK, 2020.
26. Bloch, C.; Sørensen, M.P.; Graversen, E.K.; Schneider, J.W.; Schmidt, E.K.; Aagaard, K.; Mejlgaard, N. Developing a methodology to assess the impact of research grant funding: A mixed methods approach. *Eval. Program Plan.* **2014**, *43*, 105–117. [CrossRef] [PubMed]
27. Watt, B. Funded Higher Educational Research: Sustainable and Influential. In *Sustainable Pedagogical Research in Higher Education*; Routledge: Oxford, UK, 2020; pp. 20–28.
28. Guerra, C. Educational Research Sustainability in Higher Education: Reflections on the Concept, Factors, and Actions for its Enhancement. In *Sustainable Pedagogical Research in Higher Education*; Routledge: Oxford, UK, 2020; pp. 7–19.
29. Oztaysi, B.; Onar, S.C.; Goztepe, K.; Kahraman, C. Evaluation of research proposals for grant funding using interval-valued intuitionistic fuzzy sets. *Soft Comput.* **2015**, *21*, 1203–1218. [CrossRef]
30. David, F.; Abreu, R.; Segura, L.; Formigoni, H.; Mantovani, F. Impact of the Economic Crisis on the Higher Education: The Case of Portugal. In Proceedings of the INTED 2015 Conference, Madrid, Spain, 2–4 March 2015; pp. 6053–6063, ISBN 978-84-606-5763-7.
31. Guerra, C.; Costa, N. Educational innovations in Engineering education: sustainability of funded projects developed in Portuguese higher education institutions. In Proceedings of the 2018 3rd International Conference of the Portuguese Society for Engineering Education (CISPEE), Aveiro, Portugal, 27–29 June 2018; pp. 1–6.
32. Huet, I.; Costa, N. Challenges for promoting teacher development in higher education: The Portuguese context. Learning to Teach in Higher Education. In *Learning to Teach in Higher Education. Approaches and Case Studies in Europe*; University of Aveiro: Aveiro, Portugal, 2010; pp. 83–94.
33. Oliveira, C.G.; Oliveira, P.C.; Costa, N. Students' and teachers' perspectives about quality of engineering education in Portugal. *Eur. J. Eng. Educ.* **2012**, *37*, 49–57. [CrossRef]
34. Steckler, A.; Goodman, R.M. How to Institutionalize Health Promotion Programs. *Am. J. Heal. Promot.* **1989**, *3*, 34–43. [CrossRef]
35. Pluye, P.; Potvin, L.; Denis, J.-L.; Pelletier, J.; Mannoni, C. Program sustainability begins with the first events. *Eval. Program Plan.* **2005**, *28*, 123–137. [CrossRef]

36. Scheirer, A. Is sustainability possible? A review and commentary on empirical studies of program sustainability. *Am. J. Eval.* **2005**, *26*, 320–347. [CrossRef]
37. Luukkonen, T. The European research council and the European research funding landscape. *Sci. Public Policy* **2013**, *41*, 29–43. [CrossRef]
38. Da Costa, A.F.; Pegado, E.; Ávila, P.; Coelho, A.R. Mixed-methods evaluation in complex programmes: The national reading plan in Portugal. *Eval. Progr. Plan.* **2013**, *39*, 1–9. [CrossRef]
39. Fixsen, D.L.; Naoom, S.F.; Blase, K.A.; Friedman, R.M.; Wallace, F.; Burns, B.; Carter, W.; Paulson, R.; Shern, D.; Dean, L.; et al. *Implementation Research: A Synthesis of the Literature (FMHI Publication No. 231)*; University of South Florida, Louis de la Parte Florida Mental Health Institute, National Implementation Research Network: Tampa, FL, USA, 2005.
40. Marques, M. Does 'programmification' foresee the sustainability of research results? Discussing the effects of thematic national programmes in pedagogical innovation in higher education. In *Sustainable Pedagogical Research in Higher Education*; Routledge: Oxford, UK, 2020; pp. 29–40.
41. Penno, L.N.; Davies, B.; Graham, I.D.; Backman, C.; Macdonald, I.; Bain, J.; Johnson, A.M.; Moore, J.; Squires, J. Identifying relevant concepts and factors for the sustainability of evidence-based practices within acute care contexts: a systematic review and theory analysis of selected sustainability frameworks. *Implement. Sci.* **2019**, *14*, 1–16. [CrossRef]
42. Scheirer, M.A. Linking sustainability research to intervention types. *Am. J. Public Heal.* **2013**, *103*, e73–e80. [CrossRef] [PubMed]
43. Scheirer, M.A.; Dearing, J.W. An agenda for research on the sustainability of public health programs. *Am. J. Public Heal.* **2011**, *101*, 2059–2067. [CrossRef] [PubMed]
44. Guerra, C.; Costa, N. Sustentabilidade da investigação educacional: Contributos da literatura sobre o conceito, fatores e ações. *Rev. Lusófona Educ.* **2016**, *34*, 34.
45. Fishman, B.J.; Krajcik, J. What does it mean to create sustainable science curriculum innovations? A commentary. *Sci. Educ.* **2003**, *87*, 564–573. [CrossRef]
46. Fishman, B.; Penuel, W.R.; Hegedus, S.; Roschelle, J. What happens when the research ends? Factors related to the sustainability of a technology-infused mathematics curriculum. *J. Comput. Math. Sci. Teach.* **2011**, *30*, 329–353.
47. Trueman, J.H.; Evans, R.I.; Leppmann, P.K.; Trenaman, J.M. Resistance to innovation in higher education: A social psychological exploration focused on television and the establishment. *J. High. Educ.* **1968**, *39*, 414. [CrossRef]
48. Rogers, E.M.; Singhal, A.; Quinlan, M.M. Diffusion of Innovations 1. In *An Integrated Approach to Communication Theory and Research*; Routledge: Oxford, UK, 2019; pp. 415–434.
49. Harrison, R.; Hutt, I.; Thomas-Varcoe, C.; Motteram, G.; Else, K.; Rawlings, B.; Gemmell, I. A cross-sectional study to describe academics' confidence, attitudes, and experience of online distance learning in higher education. *J. Educ. Online* **2017**, *14*, n2. [CrossRef]
50. Teles, R.; Valle, A.; Rodríguez, S.; Piñeiro, I.; Regueiro, B. Perceived stress and indicators of burnout in teachers at Portuguese higher education institutions (HEI). *Int. J. Environ. Res. Public Health* **2020**, *17*, 3248. [CrossRef] [PubMed]
51. Bonwell, C.C.; Eison, J.A. *Active Learning: Creating Excitement in the Classroom. 1991 ASHE-ERIC Higher Education Reports*; ERIC: Washington, DC, USA, 1991.
52. Sagardia, A.E.; Urdin, J.A.; Fernández, I.F. Active-collaborative learning as best practices in the development of cross-curricular competencies in Basque Country vocational training. *Educar* **2018**, *54*, 331–349.
53. Gholami, K.; Sayadi, Y. The faculty's perception of web-based instruction application in Iran's higher education. *Int. Educ. Stud.* **2012**, *5*, 204–211. [CrossRef]
54. Antonova, N.; Merenkov, A.; Yeltsin Ural Federal University. Flipped learning in higher education: Problems and contradictions. *Integr. Educ.* **2018**, *22*. [CrossRef]
55. Johnson, K.; Hays, C.; Center, H.; Daley, C. Building capacity and sustainable prevention innovations: A sustainability planning model. *Eval. Program Plan.* **2004**, *27*, 135–149. [CrossRef]
56. Moore, J.E.; Mascarenhas, A.; Bain, J.; Straus, S.E. Developing a comprehensive definition of sustainability. *Implement. Sci.* **2017**, *12*, 1–8. [CrossRef]
57. Brundtland, G.H. World Commission on environment and development. *Environ. Policy Law* **1985**, *14*, 26–30. [CrossRef]
58. De Castell, S.; Egan, K.; Beck, K.; Ilieva, R.; Waterstone, B.; Nilson, M.; Paterson, D.; O'Niell, K.; Smith, S.; Blenkinsop, S.S.; et al. *Sustainable Educational Ecologies: Final Report*; Simon Fraser University: Burnaby, BC, Canada, 2010.
59. Shediac-Rizkallah, M.C.; Bone, L.R. Planning far the sustainability of community-based health programs: Conceptual frameworks and future directions for research, practice and policy. *Health Educ. Res.* **1998**, *13*, 87–108. [CrossRef]
60. Southwell, D.; Gannaway, D.; Orrell, J.; Chalmers, D.; Abraham, C. Strategies for effective dissemination of the outcomes of teaching and learning projects. *J. Higher Educ. Policy Manag.* **2010**, *32*, 55–67. [CrossRef]
61. Savaya, R.; Elsworth, G.; Rogers, P. Projected sustainability of innovative social programs. *Eval. Rev.* **2008**, *33*, 189–205. [CrossRef] [PubMed]
62. Stirman, S.W.; Kimberly, J.; Cook, N.; Calloway, A.; Castro, F.; Charns, M. The sustainability of new programs and innovations: A review of the empirical literature and recommendations for future research. *Implement. Sci.* **2012**, *7*, 17. [CrossRef] [PubMed]
63. Guerra, C. Sustentabilidade da investigação em educação: Da concepção à implementação de um referencial. *Práxis Educ.* **2021**, *17*, 48. [CrossRef]

64. Creswell, J.W.; Creswell, J.D. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*; Sage Publications: Thousand Oaks, CA, USA, 2017.
65. Pinto-Llorente, A.M.; Sánchez-Gómez, M.C.; Pedro Costa, A. Qualitative and Mixed Methods Research in Social Sciences. In Proceedings of the Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality, Salamanca, Spain, 21–23 October 2020; pp. 193–196.
66. Cohen, L.; Manion, L.; Morrison, K. *Research Methods in Education*; Routledge: Oxford, UK, 2013.
67. Yin, R.K. Validity and generalization in future case study evaluations. *Evaluation* **2013**, *19*, 321–332. [CrossRef]
68. Corbin, J.M. Grounded theory methodology. *Grounded theory methodology. Handbook of Qualitattive Research.*. **2011**, 273–285.
69. Mayring, P. Qualitative content analysis. In *A Companion to Qualitative Research*; Sage Publications: London, UK, 2004; pp. 159–176.
70. Neuendorf, K.A. Defining Content Analysis. In *The Content Analysis Guidebook*; Sage Publications: Thousand Oaks, CA, USA, 2016; pp. 1–31.
71. Bardin, L. Content Analysis. *Psychol. Bull.* **1977**, *84*, 289–297.
72. Tavares, O.; Sin, C.; Videira, P.; Amaral, A. Academics' perceptions of the impact of internal quality assurance on teaching and learning. *Assess. Eval. High. Educ.* **2017**, *42*, 1293–1305. [CrossRef]
73. Fonseca, M.; Encarnação, S.; Justino, E. Shrinking Higher Education Systems: Portugal, Figures, and Policies. In *Higher Education in Societies*; Brill Sense: Leiden, Holland, 2014; pp. 127–147.
74. Santiago, R.A.; Carvalho, T. Effects of managerialism on the perceptions of higher education in Portugal. *High. Educ. Policy* **2004**, *17*, 427–444. [CrossRef]
75. Bennet, A.; Bennet, D.; Fafard, K.; Fonda, M.; Lomond, T.; Messier, L.; Vaugeois, N. *Knowledge Mobilization in the Social Sciences and Humanities*; Mqi Press: Frost, WV, USA, 2007.

education sciences

MDPI

Article

Assessment of the Application of Content and Language Integrated Learning in a Multilingual Classroom

Tatiana Baranova, Dmitriy Mokhorov, Aleksandra Kobicheva * and Elena Tokareva

Institute of Humanities, Peter the Great St. Petersburg Polytechnic University, 195251 Saint Petersburg, Russia; baranova_ta@spbstu.ru (T.B.); mokhorov_da@spbstu.ru (D.M.); tokareva.euy@gmail.com (E.T.)
* Correspondence: kobicheva92@gmail.com

Abstract: (1) Background: based on the constantly increasing requirements for modern university graduates, we have developed an educational model that allows us to introduce content and language integrated learning into classes with a multilingual approach, which will allow students to use several foreign languages in the process of professional communication. The purpose of the paper is to evaluate the efficiency of a newly introduced integrated learning model from the perspective of students, to identify the impact of such a model on students' professional discipline learning outcomes and to determine if the learning model contributes to an improvement in foreign language proficiency. (2) Methods: for our research we used qualitative and quantitative data from students' records of professional discipline and Spanish testing, as well as surveys and interviews on proposed learning model efficiency. Two groups of students took part in the experiment (N = 23 and N = 24). (3) Results: results on students' Spanish proficiency showed that the proposed learning model had a positive influence. Students from the experimental group got higher results on Listening, Reading and Speaking. According to results on professional discipline, both groups achieved approximately equal scores. Moreover, students described such a proposed learning model as efficient and progressive, giving a lot of advantages. (4) Conclusions: the experiment conducted confirmed the efficiency of the proposed learning model. In conclusion, it can be recommended for the realization of a multilingual approach, as well as the learning of a professional discipline.

Keywords: assessment of student learning; multilingualism; language teaching; second language acquisition; CLIL

Citation: Baranova, T.; Mokhorov, D.; Kobicheva, A.; Tokareva, E. Assessment of the Application of Content and Language Integrated Learning in a Multilingual Classroom. *Educ. Sci.* **2021**, *11*, 808. https://doi.org/10.3390/educsci11120808

Academic Editors: Orhan Agirdag, Sandra Raquel Gonçalves Fernandes, Marta Abelha and Ana Teresa Ferreira-Oliveira

Received: 15 September 2021
Accepted: 10 December 2021
Published: 14 December 2021

1. Introduction

1.1. Background

Despite significant investments in foreign language teaching, student competencies continue to fall short of expected levels (students' foreign language proficiency in Russia is worse than in Europe) [1]. With an obvious decrease in the problems associated with the availability of learning a foreign language within educational institutions, the effectiveness of language education is still unsatisfactory. Despite the fact that there are a lot of learning hours dedicated to learning a foreign language at university, the level of proficiency is still not satisfactory. Currently, it is necessary to take a fresh look at the quality of teaching foreign languages. The implementation of changes in modern higher education in the context of a modernizing society presupposes the need to implement not only a competence-based approach, but also the introduction of interdisciplinarity and meta-discipline in the learning process at a higher educational institution [1]. To achieve interdisciplinarity in education, a contemporary approach is used: content and language integrated learning. This approach allows the study of two subjects within the same discipline, namely, professional and linguistic, which forms both linguistic competencies and professional competencies [2,3]. Content and language integrated learning is already used in a number of educational institutions in Europe [4], but most teachers and methodologists in Russia still do not fully

understand the mechanism of this approach and its implementation in the curricula of schools and non-linguistic universities. This is due to the low level of foreign language competence among engineering professors in Russia. In the Russian education system, the method of integrating foreign languages into the general outline of the educational process is practically not used. Undoubtedly, the development of language competence in the universities of the Russian Federation could move to a qualitatively new level if this approach was thoroughly studied and the mechanisms of its implementation in the learning process were mastered.

However, the development of the educational system focuses not only on interdisciplinarity. One of the modern requirements for future specialists is a knowledge of several languages [5,6]. Thus, future university graduates will have to be multilingual. Such a requirement imposes a great burden on students, and also requires from the educational system innovative educational models that allow students to learn several foreign languages, as well as to be able to use these languages in the field of professional communication.

Thus, we have developed an educational model that allows us to introduce content and language integrated learning into classes with a multilingual approach, which will allow students to use several foreign languages in the process of professional communication.

To test our proposed educational model, we conducted an experiment among students of Peter the Great St. Petersburg Polytechnic University. The experiment was attended by 3rd year students of the international educational program "International Business". The choice is due to the following facts:

- the groups consist of students of different nationalities,
- teaching takes place in English (not native for students and the teacher),
- the curriculum includes the discipline "Spanish", which is studied in English, and is suitable for the use of content and language integrated learning and a multilingual approach. The point was to learn Spanish. English was the teaching language, as the students came from different linguistic backgrounds – their L1s were different, teaching was on their L2s, and they were learning an L3.

The objective of the research is to assess the application of the developed educational model, which includes content and language integrated learning and a multilingual approach.

1.2. Literature Review

1.2.1. Content and Language Integrated Learning and Its Forms

In many countries (including Russia), universities are adapting their educational programs in accordance with the growing demand for specialists with high knowledge of the English language by developing and applying in practice various bilingual educational programs. One of the most visible and most common approaches to these programs is Content and Language Integrated Learning (CLIL) [7]. There has been a large body of research on the efficacy of CLIL classes in terms of their impact on learners' language and subject knowledge [8].

CLIL is an approach that involves studying the content of a non-linguistic subject (e.g., history or geography) taught in a foreign language, and therefore learning that foreign language by studying the content of that subject [7,9]. According to the European Commission, CLIL "seeks to develop knowledge in both a non-linguistic subject and the language in which it is taught, giving the same meaning to each of them" [10] (p. 7). This means that a distinctive feature of this approach is the setting of a double goal, namely, improving the level of competency in a foreign language as well as knowledge in a non-linguistic discipline.

Many CLIL studies have found beneficial effects on English proficiency [4,5,8,11–17]. However, they should be interpreted with caution, as most CLIL studies do not sufficiently control selection effects and pre-existing differences between CLIL students and other students [18].

There is a rather contradictory picture of the efficacy of bilingual education in the development of subject knowledge. While some studies found no difference in knowledge of content and subject matter [19,20], others found benefits for bilingual students [21–23] or for monolingual students [24–27].

Piesche et al. [28] studied the influence of a bilingual and monolingual approach on the assimilation of the content of a professional discipline. The results of this study showed that monolingual students were better than bilingual ones by about a fifth of the standard deviation. This was the first time that bilingual learners had taken part in bilingual education, and Piesche et al. [28] suggested that this was the main reason for their result. In turn, studies by Dallinger et al. [14,29] found no difference between CLIL and non-CLIL students, using a history course as an example. It is worth noting, however, that the bilingual course taught history for 3 h, not 2 h (monolingual course) per week. Dallinger et al. [30] also found positive effects from a more frequent use of English, as well as a more frequent use of a second foreign language to introduce new terms (to improve knowledge of the professional discipline), which supported the idea of the deliberate use of multiple languages.

1.2.2. Multilingualism

In this study, we define multilingualism at the individual level as a person's ability to use two or more languages and "easily switch from one language to another" [31] (p. 158). From the point of view of multilingualism, languages are considered as separate language systems, and not dialects, styles or registers within one language system [32]. Teachers may not always be multilingual, as some native speaker teachers speak only one language, especially those who teach English [33]. In contrast, non-native speakers are always multilingual because they teach a language that is not their first language. Thus, they are equally proficient in several languages in the sense that they have advanced abilities in at least two languages (i.e., in their first language and the language they teach), which is not always the case with native teachers [33].

In teaching foreign languages, it has often been assumed that one teacher teaches only one foreign language, which was the norm in schools in most countries. It was relatively rare to find people teaching two or more foreign languages in Russia [34]. Research on foreign language teachers has also focused on the identity, beliefs and practices of those who teach primarily one foreign language, usually English [33,34]. Recently, researchers have begun to pay increasing attention to the benefits of introducing teaching methods in language classrooms that use multilingualism as a resource, that is, multilingual teaching methods (MTP) [35–38].

The growing level of superdiversity [39] has prompted some countries to change their language teaching programs in schools and universities to promote multilingualism among the younger generation and prepare them to succeed in a globalized world in which multilingualism is considered an asset [30,40]. As part of these changes, there have also been attempts at the political level [41] to encourage teachers to implement MTP.

Multilingual teaching is an educational approach that teachers can use to raise their students' awareness and understanding of linguistic diversity and encourage them to use their knowledge of other languages and language experience when learning a new language [33,35,42]. Through this type of learning, teachers increase the motivation of students by helping them realize that they are not entirely novices, and that they already have a set of tools that they can use to learn new languages more effectively. MTPs have also been shown to improve student literacy and pragmatic knowledge, as well as overall language performance [43–45]. Examples of MTPs include translation, awakening to language action, cross-language comparisons, multilingual storytellings and language diaries [23,46,47].

This study defines multilingualism in multilingual classrooms as the process in which multilingual teachers and students engage in complex, multiple discursive practices, including translation, to communicate in and navigate multilingual classrooms [35]. Multilingual

practices can be used in a targeted and systematic way to overcome language boundaries and to improve and maintain the language skills and multilingual competence of learners [48]. In addition, this deliberate use of multilingualism enhances students' ability to analyze and compare different language systems, which contributes to their learning [49]. According to García and Silvan [35] (p. 389), multilingualism is also "part of the discursive regimes that students must perform in the 21st century," as it not only reflects the interactions of multilingual people in their daily lives, but also how an opportunity to develop their multilingual skills through translation allows students to improve their knowledge of each of their languages.

The available research on multilingualism offers educational institutions seeking to promote multilingual practices among students a limited opportunity, since the types of multilingualism studied cannot be offered to all students. Adopting the concept of multilingualism as a pedagogical resource enabling everyone to achieve multilingualism will more accurately reflect the profiles of all teachers and students in a language class. The number of studies that have used multilingualism as an educational approach is currently very limited [50]. It is also worth noting that many studies have used a small number of participants, which affects the generalizability of the results when the goal is to understand broader trends in teacher acceptance of the MTP language in a given context [34].

At the same time, research on multilingualism as a pedagogical resource is practically absent in some countries. One such country is Russia, where the government is just beginning to introduce multilingual educational practices through initiatives to learn several foreign languages in schools and universities [51,52]. In Europe, some countries also receive much less attention than others when it comes to research on pedagogical multilingualism. For example, research in Norway has focused on teachers' beliefs and practices regarding multilingualism in schools [53], which provides only a partial indication of their use of MTP. What is interesting for our study is that in Norway, a new national curriculum for English entered into force in 2019, which promotes multilingualism in language teaching and learning [54]. The updated curriculum now emphasizes the relevance and value of multilingualism [54]. The learning outcomes of the new curriculum likewise highlight the development of an awareness of different languages and the use of the languages spoken by students to find similarities at different linguistic levels, from vocabulary and expressions to more complex language similarities and differences [54].

In addition, since multilingual users are expected to use different languages in different situations for different purposes, they may need to use all components of the communication competence; however, there is often an asymmetric development of these components, that is, they do not necessarily develop all competencies in each of these languages at the same level [55]. Consequently, the successful study of a foreign language presupposes the ability to correctly choose and use communication strategies from a linguistic repertoire [34]. Appropriate language learning strategies are tools that are believed to encourage learners to take responsibility for their own learning and lead to increased language proficiency and greater self-confidence [34].

Analysis of the literature showed that the use of subject language integrated learning for a multilingual group of students has not been studied enough. Several studies [56,57] are examining the use of CLIL in multilingual groups for the study of professional disciplines (e.g., business English) and English. However, no studies were found examining multilingual courses combining Spanish and English. Thus, the purpose of this study is to assess the efficiency of the content language integrated learning in a multilingual class.

2. Materials and Methods

Our research involved 3rd year undergraduate students ($N = 47$) studying the program "International business (international educational program)" at the Peter the Great St. Petersburg Polytechnic University. To implement the experiment, the discipline "Spanish language", which is taught in English, was used. Within the framework of this research, a multilingual approach to the study of the Spanish language and the basics of international

business in Spanish was implemented in the context of content and language integrated learning, as content and language integrated learning involves dual goal-setting, namely the study of two disciplines within one subject.

The multilingual approach is supposed to use English and Spanish in teaching materials, given that these languages are not the native languages of the students. Thus, students used only languages foreign to them.

Figure 1 shows a model of teaching undergraduate students in the department of "International Business" in the discipline "Spanish language", taught in English. The work on the presented model is suitable for students with a level of Spanish knowledge A2-B1 and a minimum level of English knowledge B2. The experiment lasted 1 semester (February 2021–June 2021). During the semester, students studied 5 topics. For each topic, work was carried out in a specific teaching method.

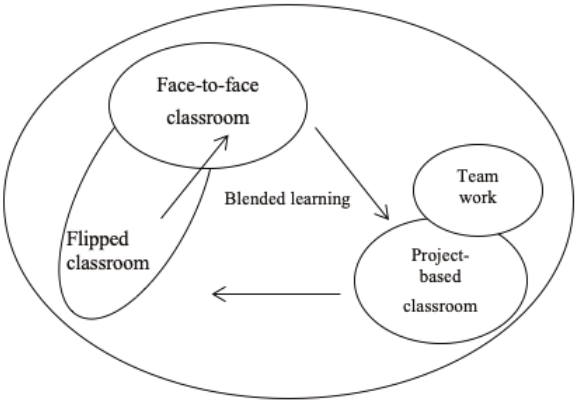

Figure 1. Learning model for the discipline "Spanish language".

The first stage of the work consists in the independent work of students with new vocabulary on a certain topic. The new vocabulary is presented in the form of a dictionary with an English-Spanish translation. Students need to familiarize themselves with the vocabulary (e.g., vocabulary for such topics as places of employment, professions, office equipment, H.R., international trade, negotiations) and, if necessary, translate it into their native language before the classroom session ("flipped classroom"). The electronic educational platform Moodle is also used, where exercises for learning vocabulary are presented (for example, exercises based on matching a term in Spanish with a definition in English), as well as the opportunity to hear the pronunciation of words in both languages.

In a classroom lesson (2nd stage), students are offered various tasks (e.g., "Read the texts and complete the gaps with the words from the list" or "Choose the verb to complete the phrases") to practice the previously studied vocabulary. During the lesson, students discuss all vocabulary, including professional terminology, in English (possibly using Spanish) that is incomprehensible to them during independent study (face-to-face classroom), making up definitions for basic economic terms on the topic. Also, during the class, the grammatical material proposed for study within the framework of the Spanish language discipline is discussed.

The basis of the 3rd stage is teamwork. Students are encouraged to independently study the theoretical material of professional content in Spanish. It is necessary to divide the group into 3–4 small subgroups of students, each of which studies the proposed section using theoretical material. Then, during the classroom session, students talk to other groups about the material they have learned in English, providing key phrases, concepts, and terms in Spanish. At the end of each block of theoretical material, questions for thought and analysis are presented (e.g., "Twenty years from now we will have seen a

huge global market emerge for standardized consumer products. Do you agree with this statement? Justify your answer."). The whole group is invited to collectively answer the questions posed (teamwork). The Moodle contains video and audio materials in Spanish with English subtitles on the topics studied, which make it possible to better understand the theoretical material.

The 4th and final stage consists of project work. Each topic of the manual ends with a final task (case study). Students are offered a task in the format of a case with questions (e.g., a case about the hiring policy applied by Lenovo). The assignment is presented as a text in Spanish describing a problematic situation in business. Questions are attached to the text to provide a solution to the problem. After studying the case, students prepare a presentation in Spanish, either in their teams or individually, about their proposed solution to the problem, answering questions in Spanish (project-based classroom).

According to this model, students studied all 5 topics during the semester.

In order to obtain a detailed analysis of the results of the experiment, preliminary testing of all 3rd year students (5 groups of 23–25 people) for knowledge of English, Spanish and the theoretical foundations of the discipline of International Business was carried out. The tests of Spanish and English proficiency consisted of 4 parts: Listening, Reading, Writing and Speaking. The tests were conducted partly through the online platform Moodle—being developed for the St. Petersburg Polytechnic University (Listening, Reading, Writing)—and partly via seminars (Speaking). The theoretical test was conducted through Moodle and consisted of 20 questions, presented in English.

Based on the most similar results in preliminary tests, we chose a section of the participants and created 2 groups. In one group, Spanish classes were held in a traditional form in English, without content and language integrated learning and a multilingual approach (N = 23). In the second group, classes were held according to our proposed educational model (N = 24). The native languages of the students of the experimental group were Chinese (7 students), Arabic (8 students), Urdu (2 students), French (2 students) and Kazakh (5 students). The native languages of the control group students were Chinese (5 students), Arabic (7 students), Urdu (4 students), French (1 student), Kazakh (5 students) and Azerbaijani (1 student).

At the end of the course, students from the two groups were re-tested in Spanish to determine the impact of the proposed educational model on the learning of Spanish.

The students in the experimental group also underwent testing on the theoretical foundations of International Business to determine the effectiveness of the use of the CLIL method in order to study a professional discipline along with the study of a foreign language. Testing in a professional discipline was carried out in English.

In addition, students from the experimental group completed a survey and a short interview, where they expressed their opinion about the course studied and the educational model applied.

This paper is based on the following research questions:

1. Is there a significant difference in the level of Spanish proficiency before and after the course?
2. Does the proposed educational model help to improve knowledge in the discipline of "International Business"?
3. Is the proposed learning model effective from the students' perspective?

The hypothesis of the study is that the proposed educational model positively influences students' Spanish proficiency and professional discipline knowledge.

To obtain the results we used both quantitative and qualitative data (Table 1).

Table 1. Data collection.

Results	Sort of Data Collection	Type of Data
Spanish proficiency	Scores on testing (N = 47)	quantitative
Professional discipline knowledge	Scores on testing (N = 47)	quantitative
Efficiency of the learning model from students' perspective	Interview (N = 24)	qualitative
	Survey (N = 8)	quantitative

For the analysis, descriptive statistics and pair-samples of students' *t*-tests were conducted.

3. Results

3.1. Learning Results

3.1.1. Spanish Testing

Testing on Spanish proficiency was conducted twice: once before the course and once after it, for two groups (experimental and control). Before the experiment, we asked the groups of students to identify the level of their Spanish proficiency. The test included the assessment of 4 categories: listening, reading, writing and speaking. When the course finished, students were tested again.

In general, the overall quality of students' Spanish knowledge in four categories improved (Figures 2 and 3).

Figure 2. Descriptive results of the pre-test and the post-test on Spanish proficiency (Experimental group).

The *t*-value test allowed us to discover whether the difference between the pre-test and post-test was significant in both groups, and thus whether it was possible to make a conclusion on the positive, neutral or negative effect of the proposed educational model.

A comparison of the results of the two tests (before and after the course) taken by the participants (experimental group) in the experiment indicates that the improvements in listening, reading and speaking were significant at the $p < 0.001$ level. In the writing category, students in the experimental group showed fewer progressive achievements, but due to Student's *t*-test they were also significant at the $p < 0.05$ level (Table 2). This difference in the development of writing skills is explained by the fact that in the CLIL group, more attention was paid to the development of communicative oral skills and group interaction. In the traditional teaching model, more emphasis was placed on writing. Taking into account the results of control group, here students showed fewer progressive achievements

in reading. Generally, however, both groups indicated a high level of improvement. Hence, we can firstly confirm the efficiency of such an integrated learning model for the purposes of learning Spanish.

Figure 3. Descriptive results of the pre-test and the post-test on Spanish proficiency (Control group).

Table 2. Descriptive results of pre-test and the post-test on Spanish proficiency.

Group	Category	Test	Mean (SD)	*t*-Value
Experimental	Listening	Pre-test	13.5 (1.87)	5.3 ***
		Post-test	16.71 (1.94)	
	Reading	Pre-test	15.32 (2.1)	4.1 ***
		Post-test	17.44 (1.97)	
	Writing	Pre-test	15.1 (1.79)	2.2 *
		Post-test	16.9 (1.88)	
	Speaking	Pre-test	14.54 (1.74)	5.2 ***
		Post-test	17.88 (1.78)	
Control	Listening	Pre-test	13.7 (1.83)	4.3 ***
		Post-test	15.98 (2.02)	
	Reading	Pre-test	15.1 (1.99)	2.1 *
		Post-test	16.88 (1.85)	
	Writing	Pre-test	15.05 (1.89)	4.2 ***
		Post-test	17.2 (1.94)	
	Speaking	Pre-test	14.37 (1.69)	4.5 ***
		Post-test	16.91 (1.78)	

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

3.1.2. Professional Discipline Testing

Assessment in the professional discipline (the International Business course) took the form of final testing that consisted of 25 closed questions in English (e.g., "What is the absolute advantage theory?" or "According to the theory of comparative advantage, how does opening a country to free trade affect its economic growth?"). This test was only done at the end of the course for both groups. The test was performed through Moodle with learners in both groups (experimental and control groups). The average test results are presented below (Table 3).

Table 3. Descriptive results of testing for professional discipline proficiency.

Group	Testing Results	Mean	SD	*t*-Value
Experimental	Professional discipline	71.87	5.72	1.7
Control	Professional discipline	73.44	6.13	

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

According to the results of the professional discipline testing, both groups achieved approximately equal scores. Control group students passed the test better, but the difference in the results is not significant, so we can confirm a neutral effect on professional discipline outcomes (the proposed educational model has the same efficiency as a traditional one). Thus, it can be concluded that the proposed learning model can be used for professional discipline learning purposes as well.

3.1.3. Efficiency of the Learning Model from Students' Perspective

The survey on the efficiency of the learning model consisted of 5 statements and was given to the experimental group of students. In this part, the students were asked to rate statements according to the following scale: 1-strongly disagree, 2-disagree, 3-unsure, 4-agree and 5-strongly agree. The results are shown in Table 4.

Table 4. Results on the efficiency of the learning model from the students' perspective.

Statement	Mean	SD
The current learning model can give more advantages rather than disadvantages to my academic achievement	4.41	0.32
The current learning model can enhance my multilingual competency	3.87	0.39
The current learning model provides complete content in my learning with good exercise	3.95	0.28
The current learning model provides me with different learning styles and can make my learning more fun	4.11	0.42
The current learning model helps to make my lesson more effective compare to traditional learning model	4.34	0.33
Average efficiency level	4.14	

Results of the survey showed that students perceived the proposed learning model as effective and more productive in comparison to traditional learning models. What is more, from their perspective it helps to develop multilingual competency and give more advantages rather than disadvantages to their academic achievements.

In addition, students from the experimental group were asked to participate in interviews. 11 people agreed to the interview and the subsequent processing of this data. Among these were students with different academic performance. For a reasonable assessment, 8 students were selected from the volunteers:

- Two students with high scores on the results of two tests (professional discipline and Spanish),
- Two students with a high score in Spanish and a low score in professional discipline,
- Two students with high scores in professional discipline and a low score in Spanish,
- Two students with low scores on two tests (professional discipline and Spanish).

This choice of students is justified by the need to get feedback from an equal number of students with each level of academic performance. Each student was asked to answer five questions related to their assessment of learning according to the proposed model:

- Give a brief assessment of the course passed;
- Highlight the advantages of learning according to the proposed model;
- Highlight the shortcomings of training according to the proposed model;
- Do you consider it expedient to introduce the discipline "international business" into Spanish classes?
- Do you find the use of several languages in the learning process useful?

It should be noted that students with poor academic performance gave more negative feedback than students who excelled. Since students with poor results are used to looking for a cause among external factors, they pointed out the shortcomings of the educational model. However, special attention should be paid to the reasons for this attitude among students. All students noted that it became much more difficult to study according to the new model. This is due to several factors:

1. An unusual shape: over the past 2 years of studying the discipline "Spanish", students have become accustomed to the traditional form.
2. For the successful completion of the discipline, professional knowledge is required, which not all students have in the same amount.
3. Working in a new format requires a good knowledge of the Spanish language. Students experienced difficulty using the Spanish language.

As a result, five out of eight students expressed satisfaction with the passed course. Two students remained indifferent. One student spoke negatively: "The idea of studying international business in Spanish without knowing Spanish is a bad idea."

Among the advantages of the educational model, students noted:

- Interesting form of presentation of material, including various forms of activity and work ("working with case studies was the most interesting part, classes became much more interesting");
- Open recognition of the importance of using the mother tongue on an equal footing with English for learning a second foreign language ("reliance on the native language greatly simplified the study L3");
- Study of professional vocabulary in Spanish ("we were able to learn really useful Spanish");
- Obtaining skills in the use of Spanish in business communication ("it was useful to learn how to apply Spanish in future professional activities and negotiations");
- Significant progress in the development of speaking skills ("this semester a lot of attention was paid to the spoken aspect, which is usually very lacking");
- Increasing motivation to learn due to the lack of monotony in the educational process ("tasks were interesting, which motivated to complete them all").

The disadvantages were also noted:

- Unnecessary complexity ("too difficult, no one knows why");
- Insufficient knowledge of Spanish to implement such training ("we do not have such a level of Spanish to master professional content");
- Insufficient attention is paid to developing writing skills and explaining grammar ("it was better to do more grammar, not business");
- Overloaded educational process, work on cases in a group takes more time ("such a program requires a lot of time, which we do not have, since we study other disciplines").

The use of multilingual education was very popular among interviewed students. The respondents noted that at the beginning of the course they could not get used to the constant switching between languages, since in the last two years Spanish was the most common language used in the classroom. By the end of the course, however, the students appreciated the idea of using multiple languages and noted that it helped them to understand the rules of the Spanish language, as there was a comparison with their native language and with the English language.

An important aspect that the respondents drew attention to was the advantage of the multilingual approach in the study of professional discipline. Since students study according to an international program, all theoretical knowledge is accordingly exchanged exclusively in English. The proposed learning model allowed them to look at professional content differently, from the perspective of a different language, which expanded their knowledge and immersion.

4. Discussion

After analyzing the results of the experiment, we can draw conclusions about the positive experience of implementing the proposed educational model, thus confirming our hypothesis.

As a result of training according to this model, the Spanish language outcomes in the experimental group were better than in the control group. Consequently, the application of this model contributes to the study of the Spanish language. It is worth noting that the results of the experimental group showed that writing skills developed to a lesser extent than others. This difference in the development of writing skills is explained by the fact that in the CLIL group more attention was paid to the development of communicative oral skills and group interaction. In the traditional teaching model, more emphasis was placed on writing. In further work on the educational model, it is necessary to take this into account and think over additional activities for better progress in writing skills.

The study of professional discipline in the Spanish language class has become a controversial point in discussion with students. Some students responded negatively. However, according to the results of testing students in the course of international business (professional discipline), the control group coped better than the experimental group. It is important to emphasize that the difference in the results of the control and experimental groups turned out to be insignificant. Thus, the application of the educational model under consideration contributes to the study of professional content, albeit insignificantly. With the further development of the design of the model under consideration, it is worth taking this into account and finalizing the block of professional discipline for its greater efficiency.

As a result of the survey (N = 24), as well as interviews (N = 8) with students, it was revealed that students evaluate this model as effective. The development of multilingualism and switching between languages was especially noted. The students also noted the increased motivation to study and the emerging opportunity to use the Spanish language in business communication. However, some students also noted the following negative aspects: excessive complexity, unfamiliarity and insufficient knowledge of the Spanish language for effective study.

It is important to note that this study has limitations. The study involved only one experimental group of students. This is due to the fact that the proposed educational model is very new; we have not found research that introduces the approved model into the educational process. Thus, the result of the experiment was impossible to predict, which carried risks for the effectiveness of the educational process.

An important difference between our study and other works listed in the literature review is the simultaneous introduction of CLIL and the multilingual approach into one educational model. Thus, the students studied the Spanish language and the professional discipline (International Business), while using multilingualism and their mother tongue (i.e., English and Spanish).

5. Conclusions

In connection with the tightening requirements for university graduates from the business side, there are changes in programs and teaching methods in higher educational institutions. The graduate must have a whole pool of professional competencies, knowledge and skills. In addition, the modernized education of the 21st century is based on humanism and openness towards the outside world. Multilingual education focused on the harmonization of different cultural spaces is a necessary part of a modern vision of the world. There is a gradual introduction of the student to a fully-fledged existence in the conditions of a modern multilingual and heterogeneous society (including the world of modern communications, social networks, e-mail, etc.)

Earlier studies [8,11,13,16,28] presented in the literature review indicate the feasibility of using both subject-language integrated learning and multilingual learning. A distinctive feature of our study was an attempt to combine these two increasingly popular types

of education. It is also worth noting that the study was conducted in Russia, where multilingual practices are not common.

Despite the growing popularity all over the world of CLIL, and its assignment by the European Union, the status of one of the leading effective means of implementing the international language policy of multilingualism, namely, the mechanisms and forms of organizing such training, have not yet found proper understanding among some representatives of teaching staff in Russian higher education, which is largely associated with national characteristics and traditions. Nevertheless, in the globalizing world space, CLIL didactics is becoming a promising area of higher education and an effective way of acquiring and improving linguistic and communicative competencies, ensuring the further successful professionalization of students and facilitating their career advancement in their chosen field of activity.

In the current study, we assessed the efficiency of the learning model based on CLIL methodology from the students' perspective, as well as analyzing the impact of such a model on students' Spanish language improvement and professional discipline knowledge.

According to the results, students from the experimental group succeeded in Spanish language learning and scored highly in final testing on professional discipline (the difference between the results of both groups—experimental and control—was not significant). The results confirmed the efficiency of the proposed learning model for both foreign language and professional discipline learning. The survey on the efficiency of the learning model from students' perspective also showed positive results. Students assessed the learning model as more beneficial and prosperous as it has many advantages, develops multilingual competence and provides them with different learning styles.

Author Contributions: Conceptualization, T.B. and A.K.; data curation, E.T.; formal analysis, A.K. and E.T.; investigation, A.K. and E.T.; methodology, E.T.; project administration, T.B. and D.M.; resources, E.T.; supervision, T.B. and D.M.; validation, A.K. and E.T.; writing—original draft, A.K. and E.T.; writing—review and editing, T.B. and A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Baranova, T.; Kobicheva, A.; Tokareva, E. Web-based environment in the integrated learning model for CLIL-learners: Examination of students' and teacher's satisfaction. In *Advances in Intelligent Systems and Computing*; Antipova, T., Rocha, A., Eds.; Springer: Cham, Switzerland, 2020; Volume 1114, pp. 263–274.
2. Baranova, T.A.; Tokareva, E.Y.; Kobicheva, A.M.; Olkhovik, N.G. Effects of an integrated learning approach on students' outcomes in St. Petersburg Polytechnic University. In Proceedings of the 2019 the 3rd International Conference on Digital Technology in Education, Yamanashi, Japan, 25–27 October 2019; pp. 77–81. [CrossRef]
3. Volodarskaya, E.B.; Grishina, A.S.; Pechinskaya, L.I. Virtual learning environment in lexical skills development for active vocabulary expansion in non-language students who learn English. In Proceedings of the 2019 12th International Conference on Developments in eSystems Engineering (DeSE), Kazan, Russia, 7–10 October 2019; pp. 388–392.
4. Admiraal, W.; Westhoff, G.; De Bot, K. Evaluation of bilingual secondary education in the Netherlands: Students' language proficiency in English. *Educ. Res. Eval.* **2006**, *12*, 75–93. [CrossRef]
5. Baranova, T.; Khalyapina, L.; Kobicheva, A.; Tokareva, E. Evaluation of students' engagement in integrated learning model in a blended environment. *Educ. Sci.* **2019**, *9*, 138. [CrossRef]
6. Baranova, T.; Kobicheva, A.; Tokareva, E.; Vorontsova, E. Application of translinguism in teaching foreign languages to students (specialty "Ecology"). *E3S Web Conf.* **2021**, *244*, 11034. [CrossRef]
7. Coyle, D.; Hood, P.; Marsh, D. *CLIL. Content and Language Integrated Learning*; Cambridge University Press: Cambridge, UK, 2010.

8. Ohlberger, S.; Wegner, C. Bilingualer Sachfachunterricht in Deutschland und Europa-Darstellung des Forschungsstands (Bilingual content learning in Germany and Europe–outline of the state of research). *Herausford. Lehr.* Innenbildung-Z. Zur Konzept. Gestalt. Und Diskuss.* **2018**, *1*, 45–89. [CrossRef]
9. Coyle, D. Content and language integrated learning: Towards a connected research agenda for CLIL pedagogies. *Int. J. Biling. Educ. Biling.* **2007**, *10*, 543–562. [CrossRef]
10. European Commission. *Content and Language Integrated Learning (CLIL) at School in Europe*; Eurydice: Brüssel, Belgium, 2006.
11. Arribas, M. Analysing a whole CLIL school: Students' attitudes, motivation, and receptive vocabulary outcomes. *Lat. Am. J. Content Lang. Integr. Learn.* **2016**, *9*, 267–292. [CrossRef]
12. Belyaeva, I.G.; Samorodova, E.A.; Voron, O.V.; Zakirova, E.S. Analysis of innovative methods' effectiveness in teaching foreign languages for special purposes used for the formation of future specialists' professional competencies. *Educ. Sci.* **2019**, *9*, 171. [CrossRef]
13. Canga Alonso, A.; Arribas Garcia, M. The benefits of CLIL instruction in Spanish students' productive vocabulary knowledge. In *Encuentro: Revista de Investigación e Innovación en la Clase de Idiomas*; Universidad de Alcalá de Henares: Madrid, Spain, 2015; Volume 24, pp. 15–31.
14. Dallinger, S.; Jonkmann, K.; Hollm, J.; Fiege, C. The effect of content and language integrated learning on students' English and history competences–Killing two birds with one stone? *Learn. Instr.* **2016**, *41*, 23–31. [CrossRef]
15. Lancaster, N.K. Extramural exposure and language attainment: The examination of input-related variables in CLIL *Programmes. Porta Ling. Rev. Interuniv. Didáctica Leng. Extranj.* **2018**, *29*, 91–114. [CrossRef]
16. Agudo, J.D.D.M. The impact of CLIL on English language competence in a monolingual context: A longitudinal perspective. *Lang. Learn. J.* **2019**, *48*, 36–47. [CrossRef]
17. Pablo, M.N. Are CLIL students more motivated? An analysis of affective factors and their relation to language attainment. *Porta Ling. Rev. Interuniv. Didáctica Leng. Extranj.* **2018**, *29*, 71–90. [CrossRef]
18. Dallinger, S.; Jonkmann, K.; Hollm, J. Selectivity of content and language integrated learning programmes in German secondary schools. *Int. J. Biling. Educ. Biling.* **2016**, *21*, 93–104. [CrossRef]
19. Koch, A.; Bünder, W. Bilingualer Unterricht in den Naturwissenschaften (Bilingual classes in science). *Math. Und Nat. Unterr.* **2008**, *61*, 4–11.
20. Haagen-Schützenhöfer, C.; Mathelitsch, L.; Hopf, M. Fremdsprachiger Physikunterricht: Fremdsprachlicher Mehrwert auf Kosten fachlicher Leistungen? (Learning physics in a foreign language: Added value at cost of subject achievements?). *Z. Didakt. Nat.* **2011**, *17*, 223–260.
21. Lamsfuß-Schenk, S. *Fremdverstehen im Bilingualen Geschichtsunterricht: Eine Fallstudie (Foreign Understanding in Bilingual History Classes: A Case-Study)*; Peter Lang: Frankfurt, Germany, 2008.
22. Madrid, D. Monolingual and bilingual students' competence in social sciences. In *Studies in Bilingual Education*; Madrid, D., Hughes, S., Eds.; Peter Lang: Bern, Switzerland, 2011; pp. 195–222.
23. Surmont, J.; Struys, E.; Noort, M.V.D.; Van De Craen, P. The effects of CLIL on mathematical content learning: A longitudinal study. *Stud. Second Lang. Learn. Teach.* **2016**, *6*, 319–337. [CrossRef]
24. Anghel, B.; Cabrales, A.; Carro, J.M. Evaluating a bilingual education program in Spain: The impact beyond foreign language learning. *Econ. Inq.* **2016**, *54*, 1202–1223. [CrossRef]
25. Fernández-Sanjurjo, J.; Fernández-Costales, A.; Blanco, J.M.A. Analysing students' content-learning in science in CLIL vs. non-CLIL programmes: Empirical evidence from Spain. *Int. J. Biling. Educ. Biling.* **2019**, *22*, 661–674. [CrossRef]
26. Kondring, B.; Ewig, M. Aspekte der Leistungsmessung im bilingualen Biologieunterricht (Aspects of performance assessment in bilingual biology classes). *IDB-Ber. Des. Inst. Didakt. Biol.* **2005**, *14*, 49–62. [CrossRef]
27. Madrid, D.; Barrios, E. A comparison of students' educational achievement across programmes and school types with and without CLIL Provision. *Porta Ling.* **2018**, *29*, 29–50. [CrossRef]
28. Piesche, N.; Jonkmann, K.; Fiege, C.; Keßler, J.-U. CLIL for all? A randomised controlled field experiment with sixth-grade students on the effects of content and language integrated science learning. *Learn. Instr.* **2016**, *44*, 108–116. [CrossRef]
29. Canz, T.; Piesche, N.; Dallinger, S.; Jonkmann, K. Test-language effects in bilingual education: Evidence from CLIL classes in Germany. *Learn. Instr.* **2021**, *75*, 101499. [CrossRef]
30. Dallinger, S.; Jonkmann, K.; Hollm, J.; Fiege, C. Merkmale erfolgreichen bilingualen Sachfachunterrichts. Die Rolle der Sprachen im deutsch-englischen Geschichtsunterricht (Aspects of successful bilingual subject education. The role of languages in German-English history lessons). In *Die Wirksamkeit Bilingualen Sachfachunterrichts: Selektionseffekte, Leistungsentwicklung und die Rolle der Sprachen im Deutsch-Englischen Geschichtsunterricht*; Dallinger, S., Ed.; Ludwigsburg University of Education: Ludwigsburg, Germany, 2015; pp. 145–165.
31. Lüdi, G.; Py, B. To be or not to be a plurilingual speaker. *Int. J. Multiling.* **2009**, *6*, 154–167. [CrossRef]
32. Kirss, L.; Säälik, Ü.; Leijen, Ä.; Pedaste, M. School effectiveness in multilingual education: A review of success factors. *Educ. Sci.* **2021**, *11*, 193. [CrossRef]
33. Calafato, R. The non-native speaker teacher as proficient multilingual: A critical review of research from 2009–2018. *Lingua* **2019**, *227*, 102700. [CrossRef]
34. Calafato, R. Teachers' reported implementation of multilingual teaching practices in foreign language classrooms in Norway and Russia. *Teach. Teach. Educ.* **2021**, *105*, 103401. [CrossRef]

35. García, O.; Sylvan, C.E. Pedagogies and practices in multilingual classrooms: Singularities in pluralities. *Mod. Lang. J.* **2011**, *95*, 385–400. [CrossRef]
36. Hall, G.; Cook, G. Own-language use in language teaching and learning. *Lang. Teach.* **2012**, *45*, 271–308. [CrossRef]
37. Heuzeroth, J.; Budke, A. The Effects of multilinguality on the development of causal speech acts in the geography classroom. *Educ. Sci.* **2020**, *10*, 299. [CrossRef]
38. Spotti, M.; Kroon, S. Multilingual classrooms at times of superdiversity. In *Discourse and Education. Encyclopedia of Language and Education*, 3rd ed.; Wortham, S., Kim, D., May, S., Eds.; Springer: Cham, Switzerland, 2017; pp. 97–109. [CrossRef]
39. Liwiński, J. The wage premium from foreign language skills. *Empirica* **2019**, *46*, 691–711. [CrossRef]
40. Etxebarrieta, G.R.; Pérez-Izaguirre, E.; Langarika-Rocafort, A. Teaching minority languages in multiethnic and multilingual environments: Teachers' perceptions of students' attitudes toward the teaching of basque in compulsory education. *Educ. Sci.* **2020**, *10*, 25. [CrossRef]
41. Raud, N.; Orehhova, O. Training teachers for multilingual primary schools in Europe: Key components of teacher education curricula. *Int. J. Multiling.* **2020**, 1–13. [CrossRef]
42. Van Viegen, S.; Zappa-Hollman, S. Plurilingual pedagogies at the post-secondary level: Possibilities for intentional engagement with students' diverse linguistic repertoires. *Lang. Cult. Curric.* **2020**, *33*, 172–187. [CrossRef]
43. Brown, A. Monolingual versus multilingual foreign language teaching: French and Arabic at beginning levels. *Lang. Teach. Res.* **2021**. [CrossRef]
44. Krulatz, A.; Iversen, J. Building inclusive language classroom spaces through multilingual writing practices for newly-arrived students in Norway. *Scand. J. Educ. Res.* **2020**, *64*, 372–388. [CrossRef]
45. Matsumoto, Y. Teachers' identities as 'non-native' speakers: Do they matter in English as a lingua franca interactions? In *Criticality, Teacher Identity, and (In) Equity in English Language Teaching. Educational Linguistics*; Yazan, B., Rudolph, N., Eds.; Springer: Cham, Switzerland, 2018; Volume 35, pp. 57–79. [CrossRef]
46. Higgins, C.; Ponte, E. Legitimating multilingual teacher identities in the mainstream classroom. *Mod. Lang. J.* **2017**, *101*, 15–28. [CrossRef]
47. Schwartz, M.; Asli, A. Bilingual teachers' language strategies: The case of an Arabic–Hebrew kindergarten in Israel. *Teach. Teach. Educ.* **2014**, *38*, 22–32. [CrossRef]
48. Goodman, B.; Tastanbek, S. Making the shift from a codeswitching to a translanguaging lens in english language teacher education. *TESOL Q.* **2021**, *55*, 29–53. [CrossRef]
49. Carstens, A. Translanguaging as a vehicle for L2 acquisition and L1 development: Students' perceptions. *Lang. Matters* **2016**, *47*, 203–222. [CrossRef]
50. Haukås, Å. Teachers' beliefs about multilingualism and a multilingual pedagogical approach. *Int. J. Multiling.* **2015**, *13*, 1–18. [CrossRef]
51. Russia, M. *Curriculum for Grades 10 and 11 According to Federal Government Education Standards (FGOS)*; The Ministry of Education and Science of the Russian Federation: Moscow, Russia, 2018.
52. PIRAO. *Modernization Project for the Contents and Technologies of Teaching the Subject "Foreign Languages"*; Russian Academy of Education: Moscow, Russia, 2017.
53. Burner, T.; Carlsen, C. Teacher qualifications, perceptions and practices concerning multilingualism at a school for newly arrived students in Norway. *Int. J. Multiling.* **2019**, 1–15. [CrossRef]
54. Iversen, J.Y. Pre-service teachers' translanguaging during field placement in multilingual, mainstream classrooms in Norway. *Lang. Educ.* **2019**, *34*, 51–65. [CrossRef]
55. Cenoz, J.; Genesee, F. Psycholinguistic perspectives on multilingualism and multilingual education. In *Beyond Bilingualism, Multilingualism and Multilingual Education*; Cenoz, J., Genesee, F., Eds.; Multilingual Matters: Bristol, UK, 1998; pp. 16–34.
56. Dana Di, P. CLIL in the Business English Classroom: From Language Learning to the Development of Professional Communication and Metacognitive Skills. 2015. Available online: https://cpb-us-w2.wpmucdn.com/blog.nus.edu.sg/dist/7/112/files/2015/04/CLIL-in-the-Business-English-Classroom_editforpdf-2da6nlw.pdf (accessed on 20 July 2021).
57. Nikula, T.; Dafouz, E.; Moore, P.; Smit, U. *Conceptualising Integration in CLIL and Multilingual Education*; Multilingual Matters: Bristol, UK, 2016.

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
Tel. +41 61 683 77 34
Fax +41 61 302 89 18
www.mdpi.com

Education Sciences Editorial Office
E-mail: education@mdpi.com
www.mdpi.com/journal/education

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland

Tel: +41 61 683 77 34

www.mdpi.com

ISBN 978-3-0365-6751-8