iscte

**INSTITUTO
UNIVERSITÁRIO
DE LISBOA**

Evaluation patterns during the phase of "emergency remote teaching"

Carlota Patrícia Aveiro Rodrigues

Master Degree in Computer Science and Business Management

Supervisor:
PhD Joana Martinho de Almeida Costa Pardal, Invited Assistant Professor,
ISCTE-IUL

Co-Supervisor:
PhD Sérgio Miguel Carneiro Moro, Associate Professor with Habilitation,
ISCTE-IUL

October, 2021

Department of Information Science and Technology

Evaluation patterns during the phase of "emergency remote teaching"

Carlota Patrícia Aveiro Rodrigues

Master Degree in Computer Science and Business Management

Supervisor:

PhD Joana Martinho de Almeida Costa Pardal, Invited Assistant Professor,
ISCTE-IUL

Co-Supervisor:

PhD Sérgio Miguel Carneiro Moro, Associate Professor with Habilitation,
ISCTE-IUL

October, 2021

Thank you notes

I must give a special thanks to my advisor teacher Joana, who oriented me and was very understanding during the whole execution and writing process of the dissertation. Likewise, want to thank my co-advisor, professor Sergio. Thankful for my mother and father who have always supported me, for my family and friends who have always been by my side, a special thanks to Elen and Pedro that were the greatest colleagues and friends I could ever have during the execution of this master, and, finally, for all my colleagues and professors that somehow contributed for my professional and personal growth. Thank you all. Glad to be able to conclude another personal goal.

Resumo

O COVID-19 trouxe, certamente, muito mais aspetos negativos que positivos para a educação, uma vez que, lacunas que antes eram inexistentes foram descobertas e impactaram o ensino. A verdade é que, a pandemia acelerou um processo entre tecnologia e a educação que já estava em curso (Rosa & Zaboroski, 2020). Esta é, também, uma oportunidade para prevenir este tipo de situação e evoluir de acordo com o que foi aprendido no processo. Por muito negativamente que o COVID-19 tenha afetado a educação em termos de ensino, também criou uma oportunidade para a educação evoluir e aprimorar o uso de tecnologia e outros instrumentos digitais nas aulas. É em ocasiões como estas que nascem estudos e investigações como esta dissertação.

O uso de ferramentas digitais para recolher e validar dados, assim como, criar padrões entre os dados recolhidos é essencial para que estudos como este ocorram conforme o esperado. Supondo, a título de exemplo, que, enquanto à distância, um aluno tende a ser menos concentrado e tenta ludibriar o professor durante as tarefas de avaliação, esta situação cria uma maior dificuldade no trabalho do professor, pois os alunos justificam o desinteresse ou a falta de presença nas aulas com falta de internet ou outras desculpas difíceis de provar. Isso também pode indicar que, uma vez que ocorreu a passagem das aulas presenciais para a educação à distância, o instrumento de avaliação utilizado não se adequava ao propósito de avaliação. Estas são opiniões que podem e devem ser correlacionadas com outras para que exista a possibilidade de extrair padrões de conhecimento interessantes e conclusões sobre como evoluir a educação em termos de qualidade. Daí a importância desta dissertação, ou seja, analisar essas variáveis que só surgiram durante a pandemia do COVID19 e que, com certeza, vieram para ficar.

Palavras-chave: Ensino Remoto de Emergência; COVID19; Avaliação; Avaliação Online

Abstract

COVID-19 certainly brought many more negative aspects than positives ones to education. Gaps that did not exist before were now discovered and impact education. The truth is that the pandemic accelerated a process between technology and education that was already underway (Rosa & Zaboroski, 2020). This is an opportunity to prevent this sort of situation and evolve according to what was learned in the process. As negatively as COVID-19 affected education in terms of teaching, it also created an opportunity for education to evolve and enhance the use of technology and other digital instruments in class. It is in opportunities such as this that studies and investigations like the one on this dissertation are born.

The use of digital tools to collect and validate data, as well as create patterns between the data collected is essential for studies such as this to go as expected. Assuming, as an example, that, while at distance, a student tends to be less focused and tries to cheat during an evaluation task, this situation will create a greater difficulty in the teacher's job, because, based on teachers' opinions, students justify their lack of interest or even presence in classes with lack of internet or other excuses that are hard to prove wrong, such as cheating. This may also indicate that, once there was a passage from face-to-face classrooms to distant learning, that the used evaluation tool was not suited for the purpose of evaluation. These are opinions that can and must be correlated with others to extract patterns of knowledge from the data collected and to be able to obtain conclusions about how to evolve education. Hence the importance of this dissertation, that is, to analyze these variables that only appeared during the COVID19 pandemic and have, for sure, come to stay.

Key-words: Emergency Remote Teaching; COVID19; Evaluation; Online evaluation Methods

Index

Figure Index

Table Index

Chapter 1 – Introduction

1.1.Context

The appearance of the COVID19 pandemic in the middle of a school term forced social isolation and caused a radical transformation in education and teachers' evaluation methods. The enhanced changes of the teaching practice, due to the pandemic in 2020, have considerably changed the way teachers have recognized themselves in virtual environments, especially in terms of technological instruments and tools (G. S. Vieira, 2020).

 "Students also appear to be unable to concentrate well and are constantly distracted – both during online lessons and afterward – by all their social media" (Kirschner & Mirjam, 2020, p. 9). This generates a need for teachers to manage and address what used to be an internal issue in a classroom, externally, as well as use new, dynamic and effective assessment tools. According to (Moraes et al., 2020), an online lecture lasts an hour or more, it tires students and tends to not to be productive. So, if teachers are not careful in choosing the correct techniques and adopting an active teaching strategy, it can be even less productive. This mandates changes in the traditional teaching and evaluation modes. Despite the pandemic affecting students because they must adapt to a new form of studying, this change impacts teachers the most, who now must change the way they teach, evaluate, as well as coordinate students' integration and motivation to this new type of learning. The decision to change to the Emergency Remote Education (ERE) had "impacted the planning and execution of activities developed by teachers, making it essential to rethink teaching practices to adapt them to the ERE with the use of Digital Information and Communication Technologies (TDIC)" (Moraes et al., 2020, p. 182).

There are motivating studies in the scope of online evaluation techniques. Finding pedagogical patterns in online education instruments is a study presented by Seoane-Pardo and García-Peñalvo (2014) that aims to study pedagogical patterns based on data collected from years of experience on eLearning education focusing on the several aspects of education as teaching, it concludes interesting facts about the use of online education

instruments such as quizzes (Seoane-Pardo & García-Peñalvo, 2014). There is also an interesting investigation made on a Middle East College school term about student and teacher interaction and the analyses of the education processes such as evaluation. In this case, the author even concludes, after analyzing teachers' answers to questionnaires and interviews that, on average, teachers find that evaluation should be majorly based on asynchronous activities, in other words, with no direct interaction between teachers and students, and that synchronous activities should only be online classes or when in need to give an opportunity to a student who failed (Mohmmed, Khidhir, Nazeer, & Vijayan, 2020). Even though both studies are interesting. The truth is, none of them answer the question: "Are there any online evaluation patterns during the Emergency Remote Teaching (ERT) state on Portuguese schools?" The reason for this is simply because, this has not been tested. Even though this data might have been collected in the past for other investigations, it was not exploited in this way, therefore, no study was found regarding this specific matter.

Only when the COVID19 pandemic was declared a calamity case in Portugal, distant teaching was implemented for the first time. Never before had existed a time where remote teaching was mandatory or even possible. ERT forced the use of new technologies and the learning of new digital skills and opened a whole new list of possibilities for studies about this new matter that was distant teaching. Hence the purpose of this dissertation and the importance of analyzing teachers' opinions about online and distant evaluation modes and finding patterns that define the evaluation on several levels during the ERT. This issue is massively important because it can help to ease the adaptation and diminish the impact of this type of evaluation when it happens again. Nowadays, there are no studies concerning teachers' opinions about modes of evaluation during ERT in Portugal. So, not much has been done regarding this subject. Therefore, the main goal of this dissertation is to find the best evaluation patterns, in terms of modes of evaluation, used by teachers during the several Remote Emergency Teaching states of mandatory schools throughout the COVID19 pandemic.

1.2. Motivation and theme relevance

This proposal will address the needs, not yet fully stated on the problem, of trying to discover which of the evaluation patterns was better in evaluating students by teachers during the Remote Emergency Education.

The objective is to analyze teachers' answers in a survey and their opinion on social media about evaluation patterns during this pandemic.

Remote Emergency Education is a temporary situation. Although, this does not mean that it will not happen again. On the contrary, it is expected that, according to the growth and evolution of the humankind, it will happen more often, hence the relevance of this study. It is important to have defined evaluation methodologies that have been investigated and proven to be successful. This study will help teachers in the future to choose the most adequate evaluation techniques and evaluate their students more successfully, using tools that were proven to be effective, in the case of having to change from a face-to-face education to a Remote Emergency Education again.

1.3. Investigation Goals and methodologic approaches

This dissertation's main goal is to analyze how the evaluation in mandatory education was implemented during the emergency remote education state. As specific goals, this dissertation will investigate, as an example, if:

- students of primary schools kept their grades during the remote emergency education and if there was an evaluation pattern.

- students of secondary schools kept their grades during the Remote Emergency Teaching and if there was an evaluation pattern.

- the difficulty degree in implementing the evaluation at different levels of education was similar among teachers.

- the remote emergency education caused changes in grades.

- there was an evolution in learning during the remote emergency education.

1.4. Hypothesis of investigation

To investigate deeper, correlations, and associations between the collected data must be done. The best way to do so is to start by defining some of the possible outcomes that could be accepted and considered successful because of this investigation, in other words, the research hypothesis. For this investigation, the following hypothesis have been defined:

- This evaluation pattern resulted in good grades and was easy to implement.

- This evaluation pattern resulted in average grades and was medium-hard to implement.

- This evaluation pattern resulted in bad grades and was hard to implement.

- The evaluation pattern was adequate.

- No specific evaluation pattern was used.

1.5. Structure and Organization of the dissertation

This dissertation has five chapters that will validate the different phases of this research.

Chapter 1 holds the Introduction, explaining the context of the theme and the expected results. Chapter 2 is the narrative review of literature on the theme and goes through every dissertation topic. Chapter 3 explores the research sample and explains the Methodologies used to exploit the data from the questionnaire. Chapter 4 describes the deep analysis of data extracted based on several applied quantitative techniques on SPSS. Chapter 5 sums all the information gathered in Chapters 3 and 4, reflects on the conclusions of the described research, presents the limitations and paths to future work.

Chapter 2 – Narrative Literature Review

2.1. Emergency Remote Teaching

Emergency Remote Teaching (ERT) is a temporary state declared by the government. It happens when face-to-face education must change to remote due to situations of crisis (Ferri, Grifoni, & Guzzo, 2020). It is an unexpected situation caused by several factors like natural disasters, wars, diseases and other calamities that may occur and put our nation's health or safety at risk.

The COVID-19 pandemic impacted the world and consequently the educational field. This mandates a change in the traditional teaching models we know because teaching must be at distance during this state (Castaman & Szatkoski, 2020; Gonzalez et al., 2020a).

2.1.1. The COVID-19 pandemic and implementation of ERT

In March 2020, COVID-19, also known as Sars-Cov-2, was classified as a global pandemic becoming the first crisis in the digital era that we assist to (Bozkurt & Sharma, 2020).

Meanwhile, the world finds itself dealing with the COVID-19 pandemic and its socio-cultural, economic and political impacts, and schools were forced to implement this challenging new type of teaching: the Emergency Remote Teaching (ERT) (Bozkurt & Sharma, 2020). ERT is, to put it more simply, the passage from face-to-face to remote online learning when this state is activated by the government. As soon as this state is declared, teachers must still lecture their classes by adapting and transforming their presential planned course to a remote environment with no preparation or training to do so. **"**teachers had to move from 100% face to face to 100% online learning, sometimes within 24 hours. They had to do so without any training in distance teaching, with nothing in their prior education on it, and no experience in providing for distance learning. And almost every single one of them worked themselves to the bone." (A. Kirschner & Mirjam., 2020, para. 4). In Portugal, the presential part was given, online by teachers, within 2 hours, daily, and the rest was exposed by teacher around the country in national TV , which resulted very well for students.

This situation represents a major change in teachers' and students' lives. Teachers found themselves in an unpredictable, mandatory change on teaching and must adapt every single

element of their teaching, including assessment, as well as still coordinate their other daily tasks. "Without time to prepare, they suddenly had to teach in ways they had never taught before, with no experience, with minimal equipment, little to no support, and so on" (Kirschner & Mirjam, 2020, p. 7). Students also must adapt to this new form of learning and studying at home, which cannot be the most suitable environment for learning. Nevertheless, it must be kept in mind that learning is an interaction, it's not only up the instructor to determine the learning quality experience during the Emergency Remote Teaching. It depends on the student and his or her effort just as much   (Kirschner & Mirjam, 2020).

According to Kirschner and Mirjam, (2020), students are not yet capable of managing their time remotely as supposed and this is due to not having the correct and necessary tools and knowledge to do so. Therefore, their learning is being affected. Students complain about lack of teacher guidance and they claim that it takes away their focus on studying (Kirschner & Mirjam, 2020). Unfortunately, not all students have the capability or possibility to stay focused at home, whether it is because of not having the necessary instruments to study such as computer, internet, space, silence or even because they get distracted elsewhere than school. The truth is, most of the investigations on this subject show that in average most of the students increased or sustained their grades [as showed on the examples of Table1]. Sadly, in Portugal there is no such data about this subject.

Study	Source
A study made on the University of Valladolid in Spain found out that students had changed their learning habits with the appearance of COVID-19. They became more efficient, and their results improved. When asked about their personal life, students said that being remote impacted them personally as well as their social relationships and because of that they felt more stressed, anxious and depressed during this time.	(Odriozola-González, Planchuelo-Gómez, Irurtia, & de Luis-García, 2020)
An investigation made on the Netherlands, Vrije Universiteit Amsterdam, during the COVID-19 semester (February-June 2020) discovered that, after asking a group of students to rate their motivation, effort and academic results during the ERT state and compare it to before this period, this students motivation had decreased, they had attended less meetings, were not satisfied with classes, were not very active in meetings and spent fewer hours studying. However, somehow, their grades had gotten slightly better on average.	(Meeter et al., 2020)
A study made at the Universidad Autónoma de Madrid (UAM-Spain) found out that students from three different courses had increased their grades while on ERT state [as illustrated on Figure1] after increasing autonomous tasks (assignments, online researches, etc) then they did before.	(Gonzalez et al., 2020b, Chapter 3.2.1)

Table 1 - Studies about students' success during ERT
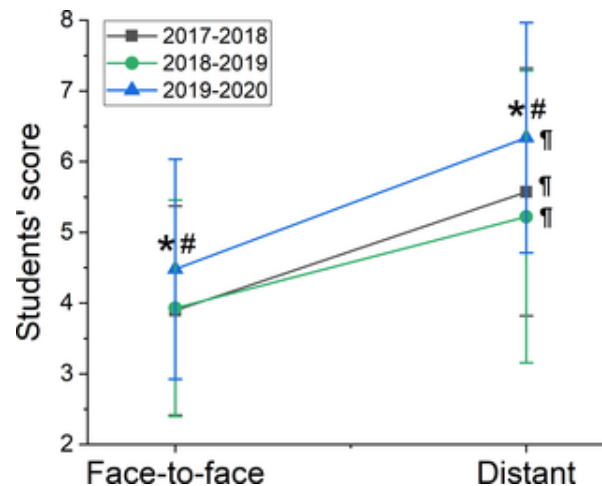
Figure 1– Students success rate was higher during ERT (UAM study -Spain)

These studies [shown in Table1 and illustrated in Figure1] demonstrate that there is an opportunity when it comes to distance learning to learn as much as in face-to-face classrooms. Though students seem to be personally unsatisfied with this situation, the fact is that student retention is lower when compared to presential learning. However, this is concerning because it implies that the dropout rate will be more prominent in remote education (Kirschner & Mirjam, 2020).

As negatively as COVID-19 affected education in terms of teaching, specially, it also created an opportunity for education to evolve and enhance the use of technology and other digital instruments in classes.

2.1.2. From Presential Education to Distant Education

Education can be divided into three groups, Online, Distant or Face-to-Face. It can be classified through its synchronicity, the duration of the simultaneous interaction between a student and a teacher, and its modality, the location of the interaction of student and teacher - online or face-to-face (Ebner & Gegenfurtner, 2019). Presential Education is a type of education with a pre-defined schedule of face-to-face interaction between teachers and students (Moraes et al., 2020). Online Education is a sort of education with online autonomous tasks and no pre-defined calendar (Ferri et al., 2020). Distant education is the most complex to define because of being an alternative type of teaching, but most of the

cases have a pre-defined calendar of virtual classes and assessments. However, there can be the case of a student or even a teacher who is unable to operate in a virtual environment due to lack of training or devices. To do so and, in that case, the alternative would have to be an offline and asynchronous education in terms of synchronicity and modality. In this case, communication would have to be made offline (Ebner & Gegenfurtner, 2019).

The initial plan designed by teachers at the beginning of any school term is focused on face-to-face learning activities in the classroom specially on primary and secondary education, according to Moraes et al., (2020). The fact is that teachers usually plan courses at the beginning of every school term and based on their previous experiences (Moraes et al., 2020). COVID-19 was the first pandemic of the digital era. Portuguese schools had never been exposed to a state of emergency in a context of 'public health matter' that demanded absence from schools, at least, during this era, as said before. However, that moment had arrived and schools were now experiencing ERT. This brought an environment of insecurity for schools that had to adapt in no time to a remote education without knowing how long this situation would affect schools (Rasmitadila et al., 2020).

In addition, schools, as well as instructors, not only have to change to a remote environment and adapt to this new digital teaching temporary state, as well as, accompany each student and making sure they have the necessary equipment and the capability to install the technologies required (Bozkurt & Sharma, 2020).

According to INE (2020), there is still a big percentage of the population, more precisely 16,5% (which is approximately 2 million Portuguese), who do not have access to the internet from their home. This implies that there are students along the country who are unable to have access to the World Wide Web having troubles connecting to teachers and colleagues (Pereira, Narduchi, & Miranda, 2020). This also affects teachers that must discover how to communicate with students who are in this situation.

It is clear that, in order for distance education to be successful in the future, schools must have a backup educational plan that should be organized in such a way that, in case of necessity to move to a remote environment, the communication path between teacher and students is already digitally established and the means and instruments made available to all (Castaman & Szatkoski, 2020). There is a need for analyzing evaluation of what has already been made to improve future ERT plans.

The transition from presential to online can be hard to implement but can surely have many benefits if done correctly. Distance education compels the learning of new skills, knowing new digital and being in touch with new technological tools and instruments to be able to work, study, learning or teaching digitally. This also opens the doors to establishing trust in remote work so that in the future it can be used in situations of need (Bernard et al., 2004).

2.1.2.1. Distant Teaching (ERT) vs Online Teaching: Major differences

The idea that distant education and online teaching are the same thing is often assumed by most of the population (Khvilon & Patru, 2002). The fact is distance education is not the same as online teaching. Distant education is a forced temporary state and its only intention is to substitute presential education. Meanwhile online education is a form of education (Fojtik, 2015).

"The term emergency remote teaching (ERT) was proposed to distinguish it from established, well-designed and long term online teaching" (Abdul Rahim, 2020, p. 59)

According to Bozkurt & Sharma, (2020) the adoption of a remote emergency education is a temporary issue and reflects the urgent passage of a system that was face-to-face to a remote system due to unexpected situations that may cause absence from school. This case was what happened during the COVID-19 pandemic. The disease demanded social isolation and schools were forced to close and change from a face-to-face education to a distant education.

Emergency Remote Education is a temporary state and it "should be considered a temporary solution to an immediate problem" (Aras Bozkurt & Ramesh Sharma, 2020, p. 2). It is not planned, and its only goal is to replace face-to-face classrooms while teachers and students are forced to be apart from school while Online Education will always be apart from school. These two types of teaching require a totally different approach [as illustrated on Figure 3]. Online Education is usually personalized individually for every student and there are no physical classes or calendar throughout the course. Meanwhile

in Emergency Remote Education there was a physical class that is now a remote class during this phase. The duration is not predictable, it can last for days, weeks, months or even longer. In Portugal's specific case, ERT, is more than just the passage from presential to online. There are classes on TV to assist, there is work that has to be made and accompanied by parents, as well as, there were online classes in national TV for both basic and secondary schools.
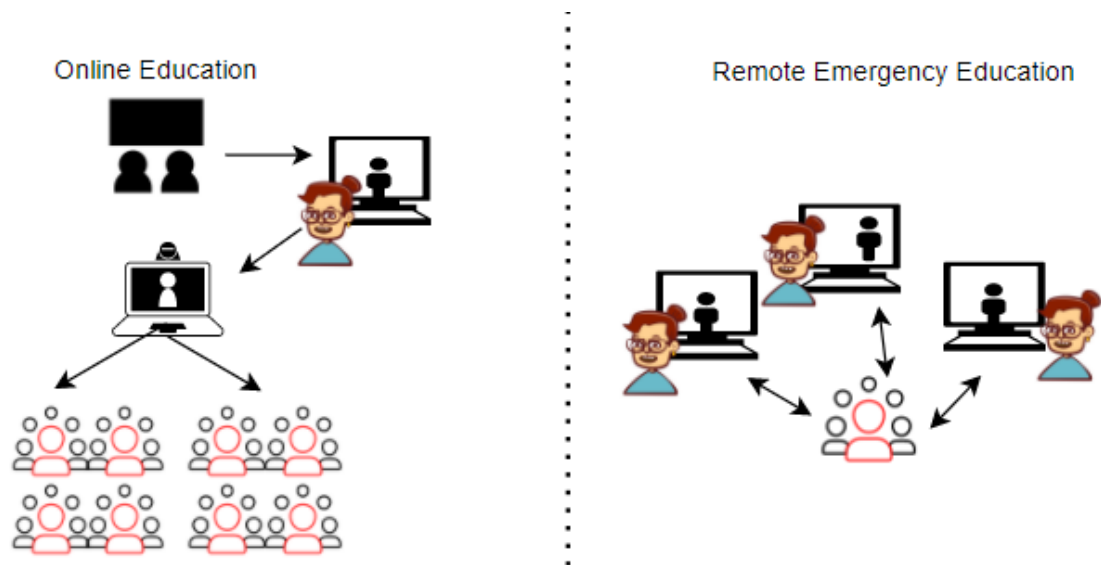
Figure 2 - Difference between Distant Education and Online Education

Despite of the scenario of uncertainty for teachers, parents and students, who are now forced to change their habits and teaching methodologies, this brings along an opportunity to improve teaching processes in order to avoid gaps in education in similar situations and enhance this type of education (Oliveira, 2020).

2.1.3. Challenges and Impacts of ERT on Schools

In order to restructure the educational system during the ERT, when this first occurred, a set of important governmental recommendations were transmitted to Portuguese schools. The most important were: Redefinition of curricular goals, Elucidate the role of the teacher in effectively supporting student learning; Guarantee support for the most vulnerable students and families and the implementation of a communication system, adapted to each student, in order to closely monitor their learning (M. de Vieira & Silva, 2020).

The interruption of face-to-face teaching imposed the challenge of this new digital era of teaching in a new educational model based on online education methodologies that make use of digital technologies (M. de Vieira & Silva, 2020).

COVID-19 certainly brought more negative than positives aspects to education. Gaps that did not exist before were now discovered and were impacting education

Impact/Challenges of COVID in Education	Source
• Schools closure and interruption of presential classes • Restructuring agenda due to distant education	(Pereira et al., 2020) (Castaman & Szatkoski, 2020)
• Uncertainty and structural limitations that challenged the implementation of effective measures. • Teachers limitation to the use of technologies	(e Oliveira, Gomes, & Barcellos, 2020) (M. de Vieira & Silva, 2020)
• No conditions to acquire a computer or internet • Growth of social disparities. • The need to reinvent teaching. • Increased teaching hours, accumulation of tasks	(Barros, Junior, Carlos, & Monteiro, 2020) (Castaman & Szatkoski, 2020)
• Difficulties in adapting to online remote education and accessing digital technologies. • Absence of an environments that endorses remote learning	(Oliveira Almeida, Rosalina, & Alves, 2020)
• Lack of technological structure at schools • Lack of formation for teachers and students	(CANI, Sandrini, SOARES, & SCALZER, 2020), (Castaman & Szatkoski, 2020)
• Difficulty for teachers when it comes to promote remote teaching	(Barreto & Rocha, 2020)

• Lack of interaction between the school, teachers and family – this affects primary schools the most	(Silveira, 2020), (M. de Vieira & Silva, 2020)

Table 2 - Impacts of COVID-19 and ERT on teaching (M. de Vieira & Silva, 2020)

It is still too premature to evaluate the damage, holdups or even progresses or impacts [Table 2 demonstrates some of the impacts of COVID-19 on education] in education after the several ERT states throughout this pandemic period. However, it is clear that there is a need to redesign learning and pedagogy in order to obtain a more adaptive, transformative, inclusive and quality education. (Rosa & Zaboroski, 2020; M. de Vieira & Silva, 2020). If we do not accomplish this, we will not evolve in terms of teaching. In order to redesign these presential teaching processes, technology is, perhaps, one of the most important parts of this new process of teaching and learning, since it is what connects students and teachers and what allows us to manage every single task of teaching online. There are platforms to connect teachers and students online with video and sound such as Microsoft Teams, Zoom, and others. There are also tools to prepare presentations, write essays, design graphics, develop questionnaires, quizzes and everything a teacher may need to teach and evaluate students when in need. Without technology and all the tools described above, or similar, ERT would not be possible.

2.2. Online Evaluation during the ERT

2.2.1. Online evaluation strategies and methods during ERT

To evaluate a student is, in a brief definition, to assign him a grade according to the knowledge he/she acquired during a school period. When moving to an ERT state, the evaluation processes must also be adjusted and organized for the context that is being lived. The classic models of face-to-face student assessment where the main strategy is based on theoretical-practical tests, also must be replaced by new online evaluation models (Robles & Braathen, 2002).

According to Phases of Teaching (2019), there are three different phases of a virtuous teaching process. The first is the *Pre-active Phase* and it represents the Planning. Here is where the

teacher formulates the goal, what he expects his students to achieve, structures the teaching and organizes the teaching methodologies ("Phases of teaching," 2019). The second is the *Interacting phase* which refers to how the teacher manages and leads the activities such as creating a class, connecting to the students and transmitting knowledge in the form of teaching. The third and last phase is the *Post-Active Phase*. It represents, essentially, feedback and consolidation, which is the part where evaluation activities, assigning grades and summarizing what the results were, happen ("Phases of teaching," 2019). If done correctly, it will not only affect the student's grade but also his or her motivation in learning. In addition, it can also enhance the gaps in strategies used and that leads to an opportunity to improve, "an effective evaluation process would indicate what was learned by the students, what would need to be improved, the difficulties encountered and the most appropriate strategies for teaching." (Lima et al., 2020, p127).

Online evaluation can be done in several ways, considering, every time, the students' needs and respecting schools and society characteristics and beliefs. Hence, there should be used adequate, as well as updated evaluation techniques and methodologies throughout the whole evaluation process (Masetto, 2003).

When it comes to online assessment there are infinite strategies and methods [as shown in table 3] there are several common online techniques used by teachers for an online evaluation. Each technique can be categorized either as a quantitative method or qualitative method, considering that qualitative focus on text/narrative and quantitative focus on numbers (Castaman & Szatkoski, 2020).

Examples of online assessments	Source
- Each class has a topic and/or a task and in the end of the class there is a quiz about it. Examples:Day 1: Search 20 websites on LiveBinder; Quiz 1 - Day 2: Brochure on Smore; Quiz 2 - Day 3: Multimedia Program using PowerPoint; Quiz 3 - Day 4: Budget on Word; Quiz 4	(Martin & Ndoye, 2016)
Define the assessment Criteria and its respective form of testing: - Knowledge of the material: Selt-test, Traditional Assignment. - Discovering learning: Writing Assignment; Online Discussion. - Knowledge of the Subject: Assignment, Electronic portfolio - Depth of understanding: One- minute online paper - Degree of understanding/lack of understanding: Content of questions	(Robles & Braathen, 2002;Costa & Miranda, 2017)
- Choose combined assessment tools: - Online Quizzes - Open-Ended/Essay Questions - Drag-And-Drop Activities - Online Interviews - Dialog Simulations - Online Polls - Game-Type Activities - Peer Evaluation and Review - Forum Posts	(Fox, Anne, & Hackerman, 2003; Fox et al., 2003; "9 Proven Ways to Assess Student Learning Online," 2020; Belo, Moro, Martins, Ramos, Costa & Esmerado, 19; Moro, Martins, Ramos, Esmerado, Costa & Almeida, 2020)

Table 3 – Online Evaluation Methods, based on 3 different authors

The idea is that the most efficient assessment is the one that not only aims for students' success but also promotes efficient teaching practices based on students learning outcomes (Fox et al., 2003;Doherty, Riordan, & Roth, 2002). "Assessment of student learning at its best enables students to identify their own strengths and weaknesses and to

determine the kinds of information they need to correct their learning deficiencies and misconceptions. When such evaluation is properly employed, students learn that they can engage in self-assessment and continuous improvement of performance throughout their lives" (Fox et al., 2003, p. 71)

The methodology of evaluating outcomes as a form of measuring students learning and extract knowledge from it, is considered a good strategy for improving evaluation and teaching (Fox et al., 2003; Doherty et al., 2002). "An effective evaluation process would indicate what was actually learned by the students, what would need to be improved, the difficulties encountered and the most appropriate strategies for teaching." (Lima et al., 2020, p127)

2.3. Discovering patterns and extracting knowledge from online questionnaires

When there is a need to solve a question by interviewing subjects to extract knowledge from their responses and create a pattern, there are several different forms to do so, such as, for example, online, face-to-face, telephonic or even, surveys interviews.

2.3.1. Questionnaire as an investigation interview instrument

There are several ways of interviewing participants and there is a need to understand which instrument fits each investigation the best, in the means of gathering more successfully the data needed.

One-on-one, also known as face-to-face, might be the type of interview that consumes more time because of the preparation and treatment of the data, it is common in the educational world to use this type of investigation whether in evaluations or presentations (Levine, 2016). Another type of interview is the *Focus Group*, it usually used when there are a lack of numeric measurements and statistical analysis and consists of a discussion between a small group of participants (Clifford, Cope, & French, 2016). *Telephone* another interviewing technique. Although it can be a fast method to obtain answers to a small and simple group of questions, it requires the interviewer to have a big number of

contacts and there is a risk of the participant not wanting to reply to the questions or, even, not picking up the call (Levine, 2016). *E-mail* is also an interviewing technique, it can be, more informal, used directly with questions written on the body of the email or more formal, referencing to a link to an online questionnaire or survey (Levine, 2016). Last but, clearly not least, the *questionnaire*, also known as a survey, is as well as all the previous, a method to interview participants, it is a way to extract data or information that represents a target population through the use of an instrument that is usually a survey, which is an online instrument (Freitas Henrique, Oliveira Mírian, Z. Saccol Amarolinda, 1999).

Considering that we find ourselves in a pandemic situation and if the goal was to obtain the maximum answers as possible in a small time interval, this technique is, clearly, the best choice: it is fastest, cheapest, most accurate form to obtain several answers in a short period of time, as well as the answer are quick to analyze and the tool is easy to use for participants and researchers and flexible of styling (Levine, 2016;Freitas Henrique, Oliveira Mírian, Z. Saccol Amarolinda, 1999).

2.4. Related Work

This sub-chapter will be focused on exploring different works that have successfully collected and extracted data from online surveys and created a pattern by working the information gathered.

2.4.1. Questionnaire development for assessing in higher education

The development of a questionnaire assessing metacognitive patterns of students majoring in accounting in higher education (Brown, 2006)

This study's ambition was to develop a survey, the Metacognitive Pattern Indicator (MPI), that attempted to identify metacognitive patterns. These patterns can be described as "unconscious thinking or sorting preferences that influence students' behavior at a level above (or 'meta' to) conscious awareness" (N. Brown, 2006, p. 302). The survey was presented to a sample of 862 students and consisted of a list of question items that

would demonstrate the appropriate language and behavior of every single pattern without being explicit. Therefore, the goal of the questionnaire was to discover if a student who has a particular taste for a specific metacognitive pattern can be able to identify that pattern only by reading the description of the question item. The research goal was to understand how in terms of reliability and validity these patterns behave. (Brown, 2006).

In order to analyze the answer from the surveys, a statistical method called exploratory factor analysis was implemented. This method is, usually used to discover a structure or a pattern in a big set of variables.

As a result of this implementation, it was clear that students only guessed correctly, on average, 52% of the patterns, this means that only 52% of the patterns were valid and reliable. It also means that the students have different opinions regarding what pattern corresponds to each question item and leads us to think that the fact of existing several students "with contrasting preferences has important implications for teaching. It would influence how they prefer to be taught." (Brown, 2006).

2.4.2. Patterns of Response from teachers to Surveys

<u>Patterns of Response and Nonresponse from teachers to traditional and web Surveys,(Mertler, 2002):</u>

This research was performed on the Bowling Green State University to find patterns in nonresponses from teachers to web-based surveys as well as to traditional surveys. Therefore, exposing a part of the participants to an online survey and another part of the group of the participants to a traditional survey. Considering that, 25% of the responses to a survey are on average, indeed, nonresponses, this becomes an interesting subject to explore (Mertler, 2002).

The survey was constructed based on four Methodological Issues Related to Web-based Survey Methodology. Firstly, the survey must be accessible and simple to answer for all participants, if the survey is too complex might lead people to quit. Secondly, the design of the survey should be similar the traditional version (paper version). Thirdly, in order to keep data safe there is a security system required. Finally, there are necessary computer skills, in order to, complete the survey successfully from the moment the subject opens

the email with the URL to the moment the "submit survey" button is pressed. (Mertler, 2002).

The main goal of the investigation is to extract knowledge, a nonresponse pattern, from web-based surveys and make it available for educational research. The second main goal is to find patterns in the resultant nonresponses to the survey. (Mertler, 2002).

After the surveys were taken, nonresponses were categorized in five categories, according to the most suitable reasons discovered among the collected data that was extracted (Mertler, 2002):

(1) Refusal to participate.

(2) Lack of time to participate.

(3) Change of some kind in the person's life – It doesn't make sense for the person to answer

(4) Only answers to questionnaires that are part of the business of national organizations.

(5) Disliking the way, the survey is designed or its content.

197 teachers were considered as participants, these represented the most distinct geographic districts, courses, ages and genders. Teachers were split in two testing groups, one of them was tested on a web-survey and another on traditional paper survey. They were questioned about nonresponses and asked opinion and suggestion to improve surveys. (Mertler, 2002).

Based on the results, there was no difference from the web-survey to paper tradition survey and it was clear that there was a pattern regarding the most common answer on the survey, "I simply didn't want to take the time to respond,". The truth is, surveys tend to be extent, complex and take lots of time to answer. This shows that these kind of studies should be a warning to subjects who construct surveys and should help them in the process of improving the way surveys are still being assembled nowadays and make it simpler and accessible (Mertler, 2002).

2.4.3. A Mixed-Method Approach to Extracting the Value of Data

The idea of this study was to create a tool that could ease the use of social media data for the users of a particular department. Firstly, the research needed the acquisition of data regarding users' needs to have a successful New Product Development ("NPD") and, for that to happen, data was collected from social media using online free applications that exist on the World Wide Web and allow users to access this data without any cost. Two months of comments and approximately 86 thousand were extracted from the official Samsung Mobile Facebook page, the tool used to do so was NCapture of Nvivo 10 Software. All comments were analyzed individually, and, because of the internationality of the Samsung, company and website, many comments in different languages were extracted (Chan, Wang, Lacka, & Zhang, 2016). To maintain coherence, only English comments were considered. The extraction resulted in 1654 comments in the English language. The authors chose to develop an algorithm, to ease the analysis of the data extracted, and to facilitate data correlation. A relational analysis was made using a cluster to identify the relationship between the comments extracted. Factors such as categories, products, department, and quality were correlated and showed that, most probably, the majority of Samsung customers were satisfied. This enhances the importance of data analysis, as models can be developed to find similarities that could help finding patterns and helping in decision-making in companies, but this also applies to education (Chan et al., 2016).

Chapter 3 – Methodologies

3.1 Purpose of the study and Research Design

This research aims to analyze teachers' opinions about how students were assessed during the ERT state during the school terms of 2020/2021. It was conducted to study how much the ERT condition impacted evaluation on compulsory education during the COVID19 pandemic and its benefits. It is based on the collaboration of teachers from Portuguese primary and secondary schools.

Considering that there was not much time available and the risk of not collecting all the answers needed for a solid conclusion when the only element of data collection is a questionnaire, the link of the questionnaire was sent to educational facilities email, know professors and social media (Figure 4).
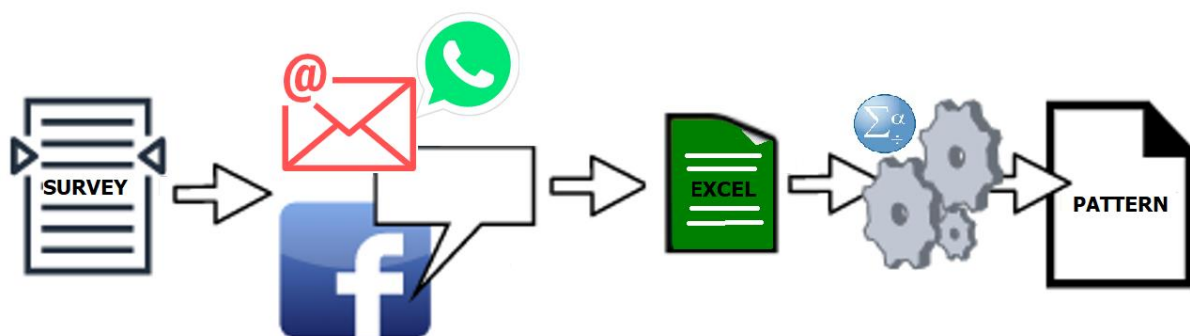
Figure 3 - Illustration of the process of collecting data for the research

An investigation is, in fact, the construction of knowledge according to certain and fixed scientific requirements. For this investigation to be considered scientifically valid, the criteria of coherence, consistency, originality and objectification must be applied (Silva & Menezes, 2001).

According to Silva and Menezes (2001)**,** a scientific research is only successful if:

- There is a specific question that needs to be answered.

- A series of steps is worked out to obtain the desired answer.

- The degree of confidence in the answer received is high.

This research was planned in three phases according to Silva and Menezes (2001), proposal of research planning. The first is the decision phase, which refers to the selection, definition, and delimitation of the topics. The second is the constructive phase, which refers to the creation of a plan for the research and its execution, that is, writing, developing, testing, and delivering the questionnaires to the teachers. The third and final phase is the editorial phase, which involves analyzing the results from the second phase and organizing the ideas for the final report (Silva & Menezes, 2001). The research design is as, briefly, described on Figure 5.

Decision Phase

Topic selection, definition and delimitation

Contructive Phase

Create a reasearch plan. Create, test and deliver surveys

Editorial Phase

Analyse results and organize final report
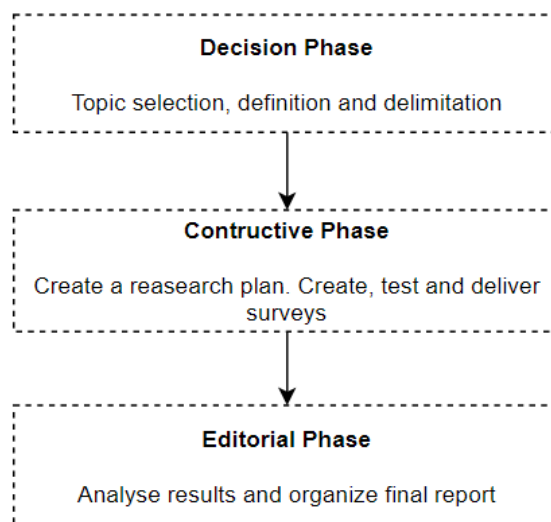
Figure 4- Illustration of the several phases of the research

After completing the narrative literature review, a questionnaire was developed. The main objective in its development was to keep it simple and easy for the respondents to answer. In this way, the constructed instrument allows collecting the needed information without burdening the participant (Mertler, 2002).

Although some online teaching methods are complex for both teachers and students to accept, as the majority have never used these new technological tools, online learning is

now a growing possibility of teaching in the future, whether for a pandemic or other possible situation that may lead to remote education. This research aims to compare and interpret teachers' responses and opinions about online evaluation and gain helpful insight from this interpretation. It is important to analyze teachers' opinions based on their experiences because they are the ones who experience the assessment firsthand. There are many advantages in studies of this type because it allows us to find out more about how students are being evaluated, how are they being evaluated and the difficulty in implementing each technique during ERT. About this specific investigation the following advantages, must be stated:

- Learning about how students are evaluated, and the credibility of online evaluation methods used in online evaluation during ERT;
- Learn about processes of online evaluation during ERT in terms of student understanding all steps of the learning process and delivered evaluation products and what should be improved;
- Create patterns to help teachers to decide the best modes of evaluation in ERT;
- Facilitate future ERT states for teachers and students;
- A well-constructed evaluation process is one of the first steps for the student to grow and learn successfully.

3.2. Methods and Type of research

According to Macdonald & Headlam (2008), unless an appropriate research methodology is used, it is very unlikely that the data collected, as well as its value for science to be able to generate a strong basis for an investigation, study or even for a evaluation (Macdonald & Headlam, 2008). It is important to mention that this research will be mainly based on survey research, which means that it will be based on conducting, correcting, delivering, analyzing, and interpreting the collected data. There are two possible surveys that can be conducted, the cross-sectional survey and the longitudinal survey, according to (Macdonald & Headlam, 2008). The survey used for this research is a Cross-Sectional, where the data collected should be extracted in a single period so that it can be treated and presented to the respondents (Table 4).

Type of questionnaires	Description	Example	Author
Cross-sectional	This type of survey is "used to gather information on a population at a single point in time. A cross-sectional survey questionnaire might try to determine the relationship between two factors, like the impact of a pro - gramme of activity on the level of benefits claims for example."	"a questionnaire that collects data on peoples' experiences of a particular initiative or event."	(Macdonald & Headlam, 2008, p.14)
Longitudinal surveys	This type of survey is used to "gather data over a period. This would allow analysis of changes in the population over time and attempt to describe and/or explain them. The three main types of longitudinal surveys are trend studies, cohort studies, and panel studies (for more details see further reading). A longitudinal study will also seek to determine the relationship between factors, but the difference is that the examination will be of a change in factors over time"	"the relationship between health and employment."	

Table 4 –Cross Sectional and Longitudinal survey differences based on (Macdonald & Headlam, 2008, p.14)

To address this problem, both quantitative and qualitive research were conducted, regarding quantitative research, it can be considered as such, because the study aims to translate the extracted information into numbers using statistical techniques, as well as qualitative research, as there are free answer questions of opinions to be considered. Considering that both approaches will be used, the study can be called a mixed research (Figure 6).

This research will take place in a short period of time, therefore it is classified as transversal in time (Silva & Menezes, 2001). The idea is to analyze the information collected in questionnaires; therefore, it is technically considered a collection of information.

Based: Atif, 2013
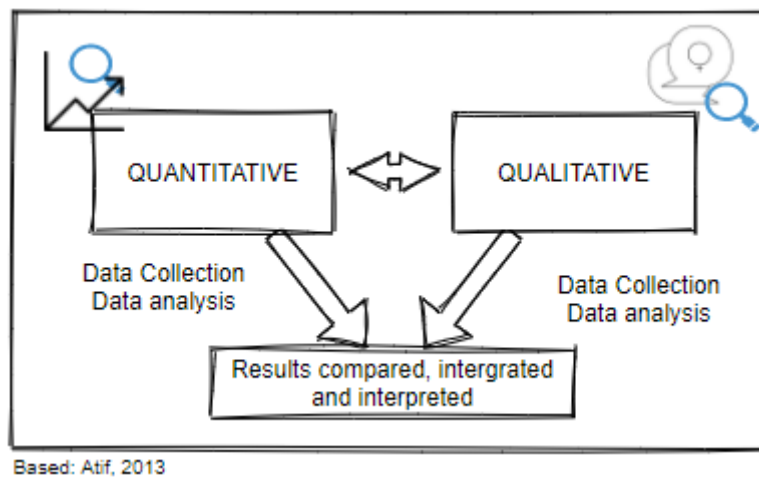
Figure 5– The mixed methods concurrent Triangulation Strategy

3.2.1. Universe and sample

The idea is to analyze the answers and use them for interpretation, based on the confidence interval that we can estimate. One possible definition and interesting point of view of the confidence interval is "an interval that estimates of a population parameter, this is the plus-or-minus figure reported in, for example, newspaper or television opinion poll results. If you use a confidence interval of 4 for example, and 54% percent of your sample picks one answer, you can be "sure" that if you had asked the question of the entire relevant population, between 50% and 58% would have picked that answer (plus or minus 4)" (Macdonald & Headlam, 2008, p. 25). According to Macdonald & Headlam, (2008), three factors can affect the confidence interval: the sample size, the percentage of answers and the size of the population (on Table 5 can be seen the values from this investigation and respective confidence interval).

Sample	*103*	
Population	*102*	
% of respondents answering to all questions	*72%*	
% of respondents that have quit the survey without finishing it	*0%*	
Confidence interval	*0,73*	

Table 5– Factors that affect the confidence interval

Only valid responses were considered among the total of responses from teachers to the questionnaire. To obtain a more robust analysis, responses by age group, gender, and type of teaching (basic or secondary) were considered. Unfinished surveys were also considered.

	Respondents' characteristics	Number Samples (102)	% Samples (100%)
Age	Less than 30	6	5.88%
	30 to 39	16	15,68%
	40 to 49	36	35,29%
	50 to 59	35	34,31%
	More than 60	9	8,82%
Basic Education	Yes	60	58,82%
	No	42	41,18%
Secundary Education	Yes	42	41,18%
	No	60	58,82%
	Basic School - Primary	28	27,45%
	Visual Arts, Visual and Technological Education	16	15,69%
	English	7	6,86%

	Mathematics	5	4,90%
	Informatics	5	4,90%
	Portuguese	6	5.88%
Disciplinary area	Physical Education	5	4,90%
	Philosophy	5	4,90%
	History	2	1,96%
	Geography	1	0,98%
	Foreign Languages	4	3,92%
	Musical Education	1	0,98%
	Physical and chemical	1	0,98%
	Economics and accounting	1	0,98%
	Biology and Geology	1	0,98%
	Others (Law, Professional Courses, Health)	14	13,73%

Table 6 - Teachers sample categorization

The sample collected was of 103 answers. Each of the answers had a category that was basic education, disciplinary area, age and years of experience. One answer was considered invalid and the remaining 102 were considered valid (the distribution of these answers by category can be analyzed in the table above, table 6)

3.2.2. Questionnaire development

The survey was created using Google Forms to save financial and time resources. Not all questions required a response, multiple choices and checkboxes were marked mandatory to simplify the analysis, and all written responses were marked optional to avoid respondents abandoning halfway through. The survey was divided into three parts. The first aims to categorize participants by age, gender, or type of education, considering the various ethical principles of a research [as shown in Figure 7].

Based: Dr. Sohall Bajammal

Figure 6 – Ethnical Principles of Research (based on Dr. Sohall Bajammal)

The second part consisted of questions related to the transition from face-to-face to distance education. The third part consisted of statements that relate to assessment during ERT. Most of the questions were closed as there is an interest in obtaining a pattern. Some questions were asked based on a Likert scale presented as a classification table such as "strongly agree" or "strongly disagree", others were asked based on multiple choices, tick boxes or open responses. The questionnaire can be answered only one time by each participant. The responses were ordered to make it easier for the participants to read. The extraction from social media still needs to be defined on which platforms and with which software this will be done.

3.2.3. Data collection and analysis

The idea was to build a survey with questions well structured so that responses could be exploited in many ways. There are several types of questions and variables that could be considered and explored, but the most important ones for this research are, namely the dichotomous, categorical, and latent ones (Table 7).

Types of varibles/questions	Description	Author
Dichotonomous	It is used to obtain a clear view of the participants	
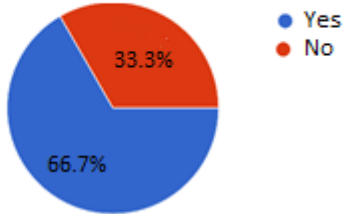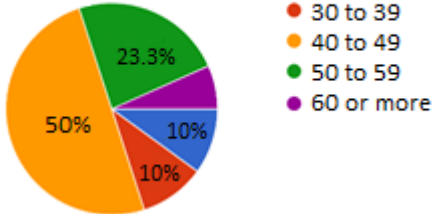
	opinion since it is, normally, obtained by the choice of one from two possible answers to a single question. e.g., Yes/No.	(Jales, 2015)
	Do you teach basic education? 30 responses ● Yes ● No 33.3% 66.7%	
	Figure 7 - Example of a dichtonomous variable in a survey question	
Categorical	It is used to obtain a more descriptive answer without any measurement scale. e.g., educational level	
	Age 30 responses ● Less than 30 ● 30 to 39 ● 40 to 49 ● 50 to 59 ● 60 or more 23.3% 50% 10% 10%	
	Figure 8- Example of a categorical variable in a survey question	
Latent	It is a hidden variable and, consequently, cannot be seen. It is normally extracted from the interpretation of other questions.	(Jales, 2015)

Table 7– Types of the variables which are present in the survey

Not only questions that only have one answer are dichotomic, checkbox questions that allow selecting more than one option and are also considered isolated dichotomous questions. Nevertheless, questions with only one possible answer and no relationship between them should be considered categorical because they only allow descriptive analysis. The questions that allow for more in-depth analysis are those that have scales once they measure data and can be compared to other questions on scales or even to other

dichotomous variables, although this is much more limited.

Some questions will tend to measure a latent variable. In this case, it will be 1) the teachers' opinion of the teaching/learning process at this stage, 2) the instruments they use, and 3) their opinion of the effectiveness of the instruments they use. To measure these latent variables, all questions related to this variable must be of the same type.

Categorical variables are ideal for more descriptive information, anything that can be measured on a scale should be the preferred method. This will facilitate validation and discover new latent variables through the relationships between questions.

Regarding data validation, in 1951, Lee J. Cronchan, invented the Cronbach Coefficient, also known as, Alfa Coefficient. Nowadays, this is a tool often used as a statistical instrument to measure a surveys' consistency. "A Cronbach alpha estimate (often symbolized by the lower-case Greek letter alpha) should be interpreted just like other internal consistency estimates, that is, it estimates the proportion of variance in the test scores that can be attributed to true score variance. Put more simply, Cronbach alpha is used to estimate the proportion of variance that is systematic or consistent in a set of test scores" (J. D. Brown, 2002)

This coefficient measures the extent to which two variables are positively or negatively related. For this purpose, the IT tool SPSS – version 25 was used. The pre-test Alfa Coefficient was 0,773, within the acceptable values that validate questionnaires [as shown on table 7]. This ended up resulting in the elimination of non-compliant questions from the pre-test due to the answers analyzed where it was clear that the respondent did not understand or interpret the question.
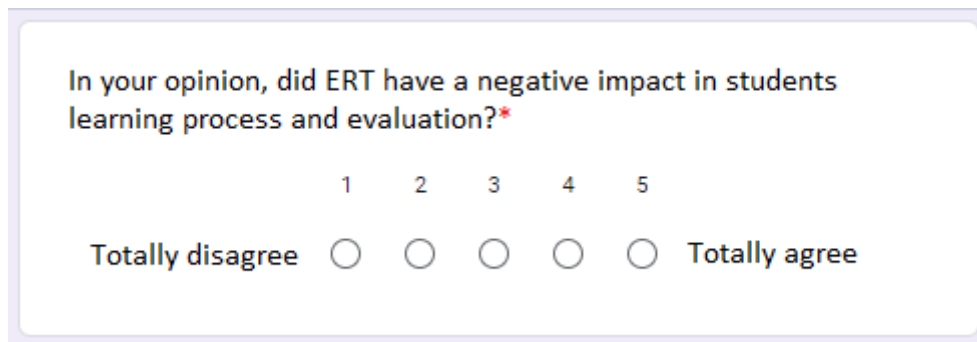
Condition	α Considered acceptable	Author
Individual Forecast Forecast for groups of 25-50 individuals	Above 0.75 Above 0.5	(Davis, 1964, p. 24)
Fundamental research Applied research	0.7-0.8 0.95	(Kaplan & Saccuzzo, 2012, p. 106)

Unacceptable reliability	<0.6	(Murphy & Davidshofer, 2004, p. 89)
Low reliability	0.7	
Moderate to high reliability	0.8-0.9	
High Reliability	>0.9	
Preliminary investigation	0.7	(Nunnally & Bernstein, 1994, pp. 245–246)
Fundamental research	0.8	
Applied research	0.9-0.95	

Table 8– The acceptable Cronbach α, according to several authors for different researches

3.2.3.1 Pretesting overview

The pre-test took place in a single stage over a period of 20 days between April and May 2021. The questionnaire was administered to teachers from primary and secondary schools and 30 responses were considered for this stage. Examples of the questions that were submitted are shown in Figure 7 and 8.
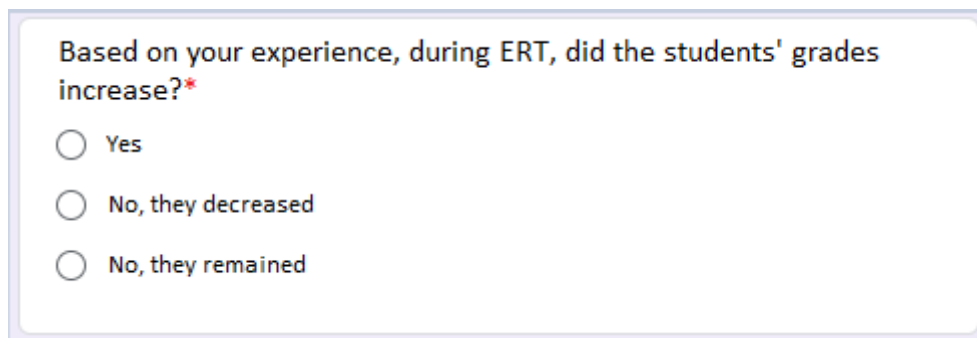
Figure 9- Question about how much did ERE negatively affect students

Figure 10– Question about students grades during ERE

41

After analyzing the pre-test responses, some questions were rephrased to facilitate understanding. In general, the questions were considered relevant and easy to answer. After revising the instrument, the final version was created and prepared for distribution.

Chapter 4 – Discussion of Results

4.1. Data treatment and processing

The questionnaires gathered were exposed to a severe cleaning process that resulted in the exclusion of a few subjects, obtaining a total of 103 subjects for the sample [as can be seen on table 9 below]. Cleansing is nothing more than turning a raw dataset into a better dataset for statistical analysis. Identifying and removing responses from individuals who do not match the target audience as well as inappropriate responses is part of data cleansing. When done successfully, the data resulting from the cleanup allows for a greater degree of success in decision making. When not done or done improperly, it limits and alters the conclusion and discredits the investigation (Aloo, 2021).

The context and nature of the responses must have been considered while deciding their exclusion. According to Aloo, (2021), in order to reduce outliers, there are a few criteria to consider when filtering or excluding a response or respondent:

- Blank questions – decide whether to maintain.
- Respondent characteristics do not match search criteria.
- When the responses are:
  - Incomplete or incomprehensible.
  - Same answer or option on all questions.
  - Unrealistic.
  - Inconsistent.
  - Nonsense answers to open-ended questions,

4.2. Instrument analysis: Validate collected data

It is necessary that the instrument used for analysis can measure what is proposed in the investigation, thus bringing a contribution to science. To validate the data collected the Alpha Coefficient, also known as Cronbach's Alpha, was used. This is a commonly used instrument to measure data reliability (Matthiensen, 2011).

The use of Cronbach's alpha works as follows, from the set of questions in the questionnaire, a unique numerical value is assigned to each answer. (Bland & Altman, 1997). A set of values is assigned to each answer to form a numerical scale, used to find predictive variables for certain and determinate models (Matthiensen, 2011). Cronbach's alpha varies, as previously mentioned in table 7, from 0 to 1. The closer the alpha is of 1, the greater the reliability in the inferred questions, the closer the alpha is to 0, the lower the reliability (Matthiensen, 2011). There is no guarantee of one-dimensionality in a questionnaire using Cronbach's Alpha, but Alfa Cronbach considers its existence. One-dimensionality is defined as a characteristic of a set of data that have something in common (Hair et al, 2005).

The data of this questionnaire was validated during the pretesting phase and Alfa resulted to be greater than 0,773, as shows on table 10, below.

Alfa de Cronbach	N of Items
,773	27

Table 9– Reliability Statistics

According to (João Maroco & Garcia-Marques, 2006), the Cronbach's alpha can be classified as follows: greater than 0,9 is the best possible result, between 0,9 and 0,8 is considered a good result, between 0,8 and 0,7 is considered a reasonable result and results below 0,7 are considered weak.

For the 33 responses analyzed in the pre-test, the value obtained was 0,773, which indicates that it is a reasonable value to ensure reliability in the data collected, to proceed with the investigation and continue to the final questionnaire, in accordance with Table 9.

4.3 Instrument to standardize the evaluation of students

4.3.1. Descriptive Statistical Analysis

4.3.1.1 Qualitative and Quantitative Variables

According to (J. Maroco, 2007) and (Vue, 2008), statistical variables can be classified by qualitative means and quantitative means. About qualitive variables, which are present in this study, the scale of measurement represents only its existence in exhaustive and mutually exclusive categories of discrete classification (J. Maroco, 2007). Qualitative variables can be measured on two different scales, nominal and ordinal. The first described, the nominal scale, represents variables measured in discrete classes, this type of variable usually does not allow sorting or qualification due to its simplistic origin, in this study the only two existing variables of this type are "lectures_basic_school" and "lectures_secundary_school" which are filled with data such as "0" or "1", "yes" or "no" (J. Maroco, 2007). The second scale described, the ordinal scale, on the other hand, represents, as its name indicates, variables that are measured in discrete classes, where it is possible to establish some ordering or qualification. (J. Maroco, 2007). It should be noted that the analysis of these scales is limited to the level of comparison between variables, for example, it is possible to analyze confidence in the percentage of each variable and percentage of people with a certain opinion but it is not possible to measure the confidence between the answers given (Vue, 2008).

Regarding quantitative variables and, considering that this study mostly consists of analyzing this type of variables. These are composed by measurement scales, this means that numbers are assigned to answers according to certain rules (Da Hora, Rego Monteiro, & Arica, 2010). Data can be ordered and the differences between the various variables can be quantified (Costa & Miranda, 2017). These scales are designated by interval and ratio (J. Maroco, 2007). Interval, allows the variables to assume quantitative values within one or more intervals, it also allows the extraction of conclusions from the relationship between two or more variables of the same type (J. Maroco, 2007). Examples of variables where this scale is commonly used are, for example, age ranges and temperature. About the second described scale, the ratio, the variables assume quantitative values and it is possible to establish an exact relationship because this scale has an absolute zero, contrary to the interval scale. Examples of common usage data are weight and height, for example (J. Maroco, 2007).

4.3.1.2 Normal distribution

The normal distribution, also usually known as Gauss Distribution, is furthermost, one of the most communal types of distribution and likely to be used in many different types of important statistical analyses. The standard normal distribution must consider the following paraments: the mean and the standard deviation (Chen & Scott, 2021).

According to Chen & Scott, (2021), for normal distribution to occur, the sample collected has to follow to subsequent criteria in order to be considered successful;

- 68% of the data collected is within +/- one standard deviation of the mean

- 95% of the data collected is within +/- two standard deviations of the mean

- 99.7% of the data collected is within +- three standard deviations of the mean

Figure 11 - Important characteristics of a normal distribution	Figure 12– Example of a normal distribution curve extracted from a variable in this study
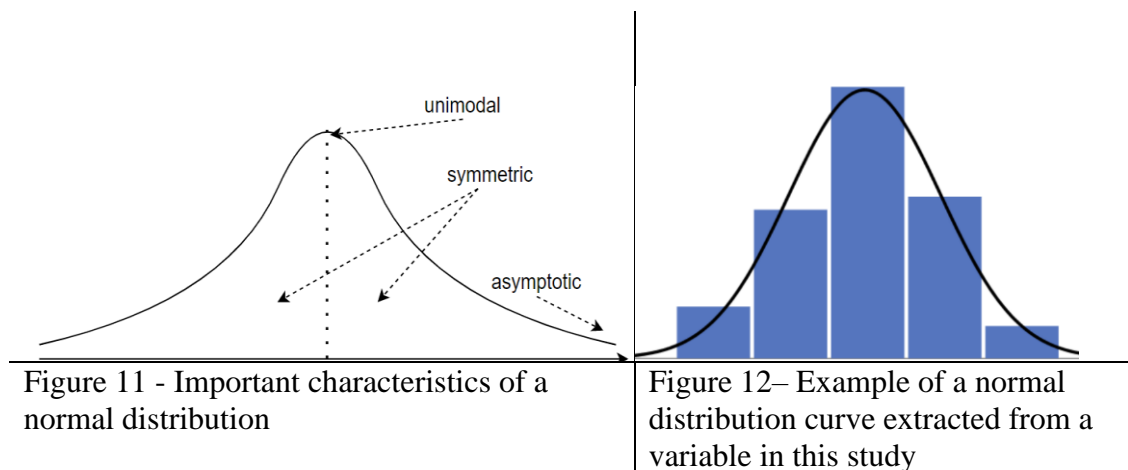
Table 10– Normal distribution curves

As the investigation followed the required criteria, there was a need to recur to normal distribution to understand if the observations when in a distributed environment are probable to fall above or below the mean, because of that, and first of all, a normality test was taken. An example of an ideal distribution curve can be seen on table 11, figure 12 and an example of a distribution curve of one of the registries of the sample analyzed (Q7) can also be seen on table 11 but on figure 13.According to (Boodie, 2019), if the p-value, which represents the most significant value in the normality test, is greater than 0,05, the null hypothesis is confirmed, this means that the sample distribution is normal. If p value is under 0,05 this means that the distribution is not normal and, therefore, the alternative hypothesis is confirmed.

There were 37 quantitative questions that were submitted to the normality test, 82% had a p value superior to 0,05 and the other 18% had a *p* not inferior to 0,01 This means that the study follows, as expected, a normal distribution(Boodie, 2019).. With this validated, the statistics such as the mean and standard deviation of the several quantitative variables were analyzed.

Several quantitative variables collected from the final questionnaire were used to be able to achieve the proposed goals. A Likert scale with five possible answers – Totally disagree, Disagree, Do not agree nor disagree, Agree, Totally agree – was used in a total of 37 questions to obtain more precise answers and facilitate the analysis.

To obtain the most frequent answer to each question the mean and standard deviation were analyzed [as shown in Table 10 below].

Question	Mean (μ)	Stand. Dev. (σ)	N
Students did not cooperate α	2,96≈3	,989	
Students were more distracted α	3,55≈4	,883	
Students were less applied α	3,5≈4	,884	
Technological means were insufficient or uncooperative α	3,02≈3	1,18	
Technological means were hard to use α	2,38≈2	, 971;	
The training of teachers was insufficient in the technological field α	2,92≈3	1,073	103
It is the same, teaching in person or remotely α	2,28≈2	, 901	
Did the ERT negatively impact the learning process and student assessment?	3,48≈4	,989	
Did the evaluation tests allow the teacher to know if the subject was well taught?	2,93≈3	,855	
Did the assessments during the ERT allow the teacher to know if the student acquired the expected knowledge corresponding to that period under assessment?	2,99≈3	,922	101
Were the assessment techniques that you implemented	3,29≈3	,898	

sufficient to identify what is necessary to work on the students?			
Should a student who fails to achieve satisfactory results be retained even during ERT?	3,1≈3	1,035	
Based on your experience, during the ERT, did the retention rate increase?	2,23≈2	,754	
Quizzes β	2,87≈3	1,242	
Presentation (e.g. Power Point, prezi, etc) β	3.05≈3	1,137	
Text Processor (eg Word) β	2,78≈3	1,174	
Oral Discussion (e.g. teams, zoom β	3,69≈4	1,075	
Simulation of Dialogue between students (e.g. teams, zoom) β	2,89≈3	1.347	99
Didactic Games (e.g. drag-and-drop activities, others) β	2,29≈2	1,206	
Work review and online peer review (e.g. teams, zoom)" β	2,7≈3	1,216	
Traditional test (several questions with timeout)	2,68≈3	1,268	

α. Based on your experience during the ERT, did you encounter any of the problems described?

β. How often did you use the instrument described for online assessment

Table 11- Mean, standard deviation of quantitive variables

The analysis that can be rethrown from table 11, above, summing up teacher opinions about teaching during ERT, the most important variables to analyze, as well as the most interesting in terms of analysis, are the variables with whom teachers agreed and disagreed the most, therefore the mean was analyzed

- "Students were more distracted in online classes". this variable was classified with a M=3,55≈4, which means that the majority of teachers', in average, agreed that students were more distracted in online classes.

- "Students were less engaged in the class", this variable resulted in a M=3,5≈4 which means that most teachers', by average, agreed that students were less engaged in online classes than in face-to-face classes.

- "Students were affected negatively as well as their learning and evaluation process" with a mean M= 3,48≈4, this means that the majority of teachers', by average, agreed that students learning process was negatively impacted during ERT.

- "During the ERT, technology was hard to use "during the ERT, technology was hard to use", with a M= 2,38≈2, this means that the majority of teachers', by average, disagreed that there was lack of knowledge towards technology.

- "Lecturing online is the same as face-to-face lectures", this resulted in a M= 2,28≈2 and means that the majority of teachers', by average, disagreed that there distant teaching is the same as face-to-face teaching during lectures.

- "Oral Discussion" is the most used evaluation technique according to teachers' opinions, it resulted in a M=3,69≈4.

- "Educational Games" is the most used evaluation technique according to teachers' opinions, it resulted in a M=2,29≈2.

In a few of the free answers, teachers left it clear that the most positive part of ERT was that they can now use more new and different technology to connect to students as well as to evaluate. Students were also forced to use new technology and present different types of work, which was also important. Teachers have also claimed that they gained experience and skills in the digital area and that ERT opened a new door towards distant teaching for teachers and students when in case of need. Teachers also stated that the most negative part of distance teaching was that they could not manage the teaching as they wished and neither the students learning process, because they simply could not control as they used to in face-to-face classes. This means that, during ERT, it was more difficult to evidence if students were indeed assisting the classes, allegedly some students just leave the session open and when questioned they use the lack of internet as an answer, and the evaluation is as hard for teachers to control due to not being able to check whether the students are cheating or not.

4.3.2. Levene test of the sample

As one of the goals is to know more about the student's assessment, a Levene test and

subsequent analysis was carried out to check if the values of the grouping would have an impact on the assessment of students in various ways and, for this purpose, the test of Levene was used and performed over the sample in SPSS version 25 to equalize variance between variables. For this purpose, the variables presented in the research hypothesis were analyzed.

In order to apply a parametric hypothesis test it is necessary that the population variances, which were previously estimated, to be homogeneous, or, in other words, to be equal (J. Maroco, 2007). Therefore, there was a need to use the Levene test for this purpose. This test is not only one of the most robust to calculate deviations from normality, but also one of the most powerful tools in testing the homogeneity of variables (J. Maroco, 2007). The hypotheses to be tested in Levene's Test are the Null Hypothesis, H0, where the variances are homogeneous, are equal, therefore, they are connected to each other and the Alternative Hypothesis, H1, where variances are not homogeneous, they are different and have no connection with each other (J. Maroco, 2007). If p>0.05 the valid hypothesis is the null hypothesis (H0). If p<0.05 the valid hypothesis is the alternative hypothesis - H1 (J. Maroco, 2007).

Investigation Goals		Levene	Sig.(p)	Technique
Students of primary schools kept their grades during the remote emergency education and if there was an evaluation pattern.	[Basic_School]	5,575	,020	Retention_Rate_Increased_DuringERT
		,964	,329	Quizzes
		1,008	,927	Online_Presentation
		1,204	,275	Text_Processor
		,226	,636	Oral_discussion
		3,436	,067	Dialogue_Simulation
		5,390	,022	Educational_Games
		,188	,665	Work&Peer Review
		,171	,680	Traditional Tests
	[Secundary_School]	4,147	,045	Retention_Rate_Increased_DuringERT
		,149	,701	Quizzes
		,053	,818	Online_Presentation
		,260	,611	Text_Processor
		,062	,805	Oral_discussion
		2,154	,145	Dialogue_Simulation
		1,940	,167	Educational_Games
		2,412	,124	Work&Peer Review

		0,14	,907	Traditional Tests
The difficulty degree in implementing the evaluation at different levels of education was similar among teachers;	[Basic_School]	1,086	,300	Evaluation_difficulty
	[Secundary_School]	,981	,324	
The remote emergency education caused changes in grades;	[Basic_School]	,109	,742	ERT_Changes_in_Grades
	[Secundary_School]	,155	,695	
There was an evolution in learning during the remote emergency education.	[Basic_School]	,354	,553	Learning_Process_Evolution
	[Secundary_School]	,894	,347	

Table 12 - Levene test

As can be seen in table 11 referring to the Levene's test made to the several variables that define the different four research pre-defined objectives, about 92% of the results of the Levene test have a *p-value* superior to 0,05 and, independently of that, the remaining 8.33% even though that the *p-value* of this variables is less than 0.05, they are also greater than 0.01, which means that, although little, there is some chance of homogeneity.

These results prove that these samples are homogeneous and, as such, we conclude that most of the variables analyzed influence our answers and proceed to the last and most important analysis, the cluster analysis.

4.3.3. Cluster Analysis

Cluster analysis is usually a powerful tool when trying to do pattern recognition. Hence the reason why it was used in this analysis. The main goal when clustering is connecting objects or variables and creating groups in a certain data set. Cluster analysis is, basically, a way of representing similar objects in a graphic form.

"The possible methods differ either in how groups are defined or in the algorithm used to create the groups. Generally speaking, group definition is based on within-group measures (e.g. high similarity between observations) or alternatively on between-group

measures (e.g. maximum distance between objects), while clustering algorithms are based on different ways to define proximity, either similarities or dissimilarities" (Tullis & Albert, 2013, para. 1). When it comes to this types of algorithms, there is a large set of choices, and each method has a different shape and "different characteristics in terms of shape, dimension and density, and each different cluster analysis approach is more oriented towards detecting a particular type of cluster rather than others they work better when objects form round, dense clusters, rather than having elongated, overlapping distributions." (Tullis & Albert, 2013, para. 1). For this study, Ward's method was used. After defining the metric, which was Euclidean metrics, which was the default, the distance matrix and the corresponding similarity matrix were calculated. Firstly, SPSS identified the two variables which are more similar. Secondly, these variables that were discovered were linked in a cluster and then the new similarity "The lines which depart from each object are connected according to the degree of similarity at which the linkage between objects or clusters happens, so that it is possible to visualize in a fast way which level of similarity intercourses among the samples" (Tullis & Albert, 2013, para. 1).

The way of interpreting this cluster is, looking at the dendrogram and checking were any pair of lines join, the lines, or variables, with the lowest distance join in first place, as can be seen in the dendrogram showed on Figure 14.
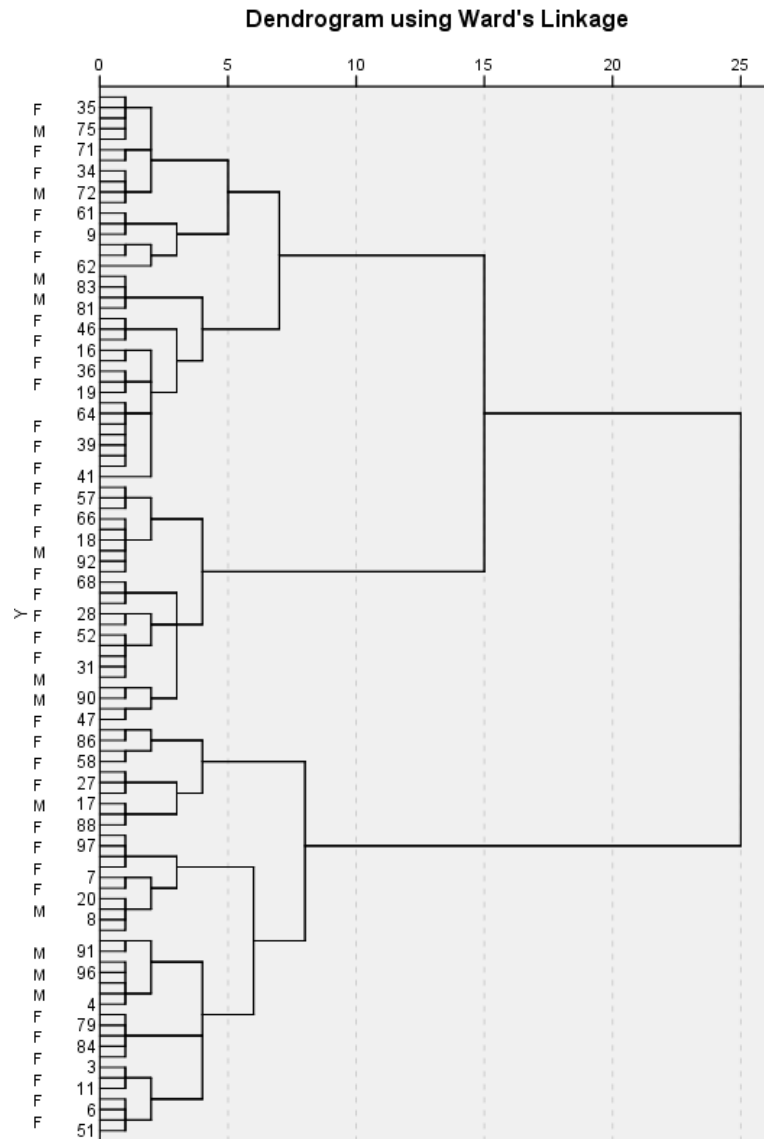
Figure 13– Dendrogram using Ward's Linkage (hierarchical)

A dendrogram is, basically, a diagram produced by a cluster algorithm during a cluster analysis that represents a tree, in this specific case, an hierarchical clustering (Tullis & Albert, 2013). This type of diagram usually represents the arrangement of the clusters produced when corresponding each analysis. It is also a branching strategy that reflects the relationships and differences within a group of entities or variables. A dendrogram is a network structure, constituted by a root node that splits in several other nodes which are connected by branches. The closer the clusters in the diagram are, the more they are related, that is, they influence each other (Tullis & Albert, 2013).

The horizontal axis (X) represents the distance between each cluster after the use of

Ward's method, method chosen due to its potential, the vertical axis (Y) represents the entities regarding the used evaluation techniques used (Tullis & Albert, 2013).

Based on Figure 14, that there are 2 groups of clusters and that become bigger and more heterogeneous as we analyze the graph from a bottom up, which means that there is more variation in what is compared within the cluster, that is, if we choose two responses from the same group, they will be more similar than if we choose one from each cluster. This can also be checked in Figure 15, which shows that the results are clear and the variables in cluster two are far more similar than in cluster one, since the majority of the variables distance is superior in cluster one than in cluster two (Tullis & Albert, 2013).

	Cluster	
	1	2
Quiz	3,27	2,48
Online Presentation	3,31	2,80
Text Processor	3,04	2,52
Oral Discution	4,24	3,14
Dialogue Simulation	3,88	1,92
Educational Games	2,82	1,78
Work and Peer Review	3,49	1,92
Traditional test	2,65	2,70

Figure 14- Final cluster centers

The final cluster centers, shown in Figure 15, are calculated based on the mean for each variable on each final clusters (Malang, n.d.). This is interesting to analyze because these results usually reproduce the characteristics of the typical case for each cluster.

- Teachers in cluster 1 tend to agree more with the use of the online evaluation tools described in Figure 15. Overall, there is a high tendency and preference with Oral Discussion and Dialogue simulation and a clear disbelief for Traditional tests and Educational Games.

- Teachers in cluster 2 tend to disagree more with the use of the online evaluation tools described in Figure 15. Overall, there is a greater preference for Oral Discussion and Dialogue simulation and a greater disbelief for Educational Games, Dialogue Simulation and Work and Peer Review.

4.4. Research hypothesis: Assessment Patterns during ERT

To validate this study, the previously stated hypothesis and determine whether the results were indeed statistically significant, we needed to compare the differences between the means and compare the p-value according to its significance level (J. Maroco, 2007). For this, two one-way ANOVA was performed on the data, once for teachers from basic school and another for teachers from secondary schools. This means that the p-value needs to be analyzed according to each variable significance level and assess the null or the alternative hypothesis. For p-values above 0,05, the alternative hypothesis is rejected . For values under 0,05 the null hypothesis is rejected. A significance level of 0,05 indicates a 5% risk of concluding that a difference exists when there is no actual difference (Duffy, Gilbert, Kennedy, & Kwong, 2002).

Regarding the ANOVA made to data collected from primary school teachers (as shown in Figure 16), since a large percentage of the various p-values is greater than 0.05, we reject the alternative hypothesis, which represents the inequality of means for any level of significance. Thus, ANOVA allowed us to conclude that, except for educational games and the traditional test, the means of the various groups are all similar for any level of significance. This result means that, regarding the usage of each technique, there are no significant differences between answers and opinions about the evaluative techniques according to their category (primary and secondary teachers).

Demographic data such as disciplinary area, age and years of experience impacted these results.

ANOVA

	Sum of Squares	Medium Square	Z	Sig.
Quizzes_usage	3,720	3,720	2,445	,121
Quizzes_easy_to_implement	,208	,208	,323	,571
Online_Presentation_usage	,146	,146	,112	,739
Online_Presentation_easy_to_implement	,167	,167	,218	,642
Text_Processor_usage	,459	,459	,331	,566
Text_Processor_easy_to_implement	2,155	2,155	3,237	,076
Oral_Discussion_usage	,030	,030	,025	,873
Oral_Discussion_easy_to_implement	,059	,059	,070	,793
Dialogue_Simulation_usage	,225	,225	,123	,727
Dialogue_Simulation_easy_to_implement	,072	,072	,081	,777
Educational_Games_usage	4,390	4,390	3,083	,082
Educational_Games_easy_to_implement	3,889	3,889	3,999	,050
Peer_and_Work_Review_usage	,924	,924	,623	,432
Peer_and_Work_Review_easy_to_implement	1,451	1,451	1,756	,190
Traditional_Test_usage	5,541	5,541	3,534	,063
Traditional_test_easy_to_implement	1,284	1,284	1,498	,225

Figure 15– First One-way ANOVA table– cluster 1

ANOVA

	Sum of Squares	Medium Square	Z	Sig.
Quizzes_usage	2,317	2,317	1,508	,222
Quizzes_easy_to_implement	,208	,208	,323	,571
Online_Presentation_usage	,032	,032	,024	,876
Online_Presentation_easy_to_implement	,098	,098	,127	,722
Text_Processor_usage	1,176	1,176	,852	,358
Text_Processor_easy_to_implement	,520	,520	,757	,387
Oral_Discussion_usage	,191	,191	,164	,686
Oral_Discussion_easy_to_implement	,000	,000	,001	,982
Dialogue_Simulation_usage	,074	,074	,040	,842
Dialogue_Simulation_easy_to_implement	,040	,040	,044	,834
Educational Games_usage	12,381	12,381	9,229	,003
Educational Games_easy_to_implement	6,395	6,395	6,900	,011
Work_and_Peer_Review_usage	3,150	3,150	2,155	,145
Work_and_Peer_Review_easy_to_implement	2,245	2,245	2,755	,102
Tradicional_Test_usage	7,621	7,621	4,927	,029
Tradicional_Test_easy_to_implement	1,744	1,744	2,051	,157

Figure 16– First One-way ANOVA table — cluster 2

The number of observations in each group is equal, so the ANOVA is robust to the violation of the assumption of equality of variances (Maroco, 2007). Concerning the ANOVA made to both clusters (shown on Figure 17 and 18) it was validated that for all variables, without exception, are statistically significant.

Concerning the ANOVA regarding opinions of secondary school teachers (shown on Figure 17) to the ANOVA about the data collected from primary school teachers, it was validated that for all variables, without exception, the means were equal or similar. Significance level, also known as p-values, are above 0.05 and never less than 0,01 what means that we reject ANOVA alternative hypothesis and it is very probable that the means analyses are similar. Also, as the results of the ANOVA for cluster 1 are very similar to the results of an ANOVA to basic school and as the results of the ANOVA for cluster 2 are very similar to the results of an ANOVA of secondary school. This leads to the belief that cluster one represents teachers from basic schools and that cluster two represents teachers from secondary schools.

To measure the degree of association between two variables, the Spearman Correlation, also known as The Spearman rank-order correlation coefficient, was used. This is a nonparametric measure of association between two variables measured (J. Maroco, 2007). As it can be seen in Figure 18, most of the data collected is perfectly associated to the teacher's personal data, such as age, years of of experience and disciplinary area which means that these variables were well chosen for the purpose of the dissertation.

Spearman Correlations

		Lectures_ Basic_Education	Lectures_ Secundary_Education	Disciplinary_ area	Years_of_ Experience	Age
Quizzes_usage	Correlation coefficient	,109	-,152	-,054	,059	-,034
	Sig	,285	,134	,596	,560	,737
Quizzes_easy_to_implement	Correlation coefficient	-,084	,084	,090	-,096	-,207
	Sig	,480	,480	,451	,425	,081
Online_Presentation_usage	Correlation coefficient	-,009	,033	-,140	-,074	-,021
	Sig	,927	,749	,167	,465	,833
Online_Presentation_easy_to_implement	Correlation coefficient	-,075	,093	-,031	-,076	,009
	Sig	,509	,414	,783	,507	,936
Text_Processor_usage	Correlation coefficient	-,090	,058	-,026	-,027	,017
	Sig	,374	,567	,795	,792	,867
Text_Processor_easy_to_implement	Correlation coefficient	-,060	,180	,024	,009	-,008
	Sig	,604	,119	,837	,935	,946
Oral_Discussion_usage	Correlation coefficient	-,045	,005	,141	-,168	-,111
	Sig	,657	,958	,163	,096	,274
Oral_Discussion_easy_to_implement	Correlation coefficient	,013	-,047	,144	-,143	-,135
	Sig	,904	,657	,176	,177	,205
Dialogue_Simulation_usage	Correlation coefficient	,035	-,049	,020	-,249	-,154
	Sig	,729	,629	,844	,013	,129
Dialogue_Simulation_easy_to_implement	Correlation coefficient	,094	-,105	-,046	-,219	-,232
	Sig	,431	,380	,700	,065	,050
Educational_Games	Correlation coefficient	,284	-,164	-,087	,075	-,003
	Sig	,004	,105	,391	,460	,976
Educational_Games_easy_to_implement	Correlation coefficient	,319	-,264	-,014	-,074	-,179
	Sig	,016	,047	,920	,586	,184
Work_and_Peer_Review	Correlation coefficient	-,147	,085	,007	-,156	-,156
	Sig	,148	,405	,948	,124	,123
Work_and_Peer_Review_easy_to_implement	Correlation coefficient	-,226	,179	,139	-,114	-,211
	Sig	,058	,136	,248	,342	,077
Traditional_Test	Correlation coefficient	-,215	,186	,150	-,024	-,159
	Sig	,032	,065	,137	,815	,116
Traditional_Test_easy_to_implement	Correlation coefficient	-,148	,131	,078	-,055	-,164
	Sig	,215	,272	,514	,645	,169

Figure 17- Spearman Correlations

While looking at Figure 18 the first thing that comes to sight is the demographic data, which is described on the columns of the table and are, respectively, "Lectures_Basic_Education", "Lectures_Secundary_Education", "Disciplinary_area"," Years_of_Experience" and "Age". These are the variables to check if they impact the data collected. As for all the p-values above 0.05 the null hypothesis is validated and that means that one of the variables depends and varies based on the other. It can be concluded the following about Spearman's correlation and the analyzed online evaluation tools:

- Quizzes are, clearly, more used in Basic Schools than in secondary, although they are quite used in both.
- Online Presentation, is as well, more used in Basic Schools than in secondary but, as well, quite used in both.

57

- Text Processor is more used in Secondary Schools, as expected, but also used in Basic Schools.

- Oral discussion is, as well, more used in Secondary Schools but also used in Basic Schools in resemblance with the above.

- Dialogue Simulation is, as well, more used in Secondary Schools but also used in Basic Schools.

- Work and Peer Review is much used in Secondary but less used in Basic, as expected.

- The lesser used tools are Educational Games and Traditional test.

All techniques seemed to be easy to implement, except for Educational Games and Work and Peer Review. The variable that had less significance, and because of that, less impact in the implementation of techniques, was Age. All the other were classified as having a great deal of importance.

- **This evaluation pattern resulted in good grades and was easy to implement.**

   According to Figure 17, the most adequate and less difficult evaluation techniques to implement that resulted in good grades were Oral Discussion, Dialogue Simulation and Online Presentation, all these techniques reject the null hypothesis and have a p-value superior to 0,722 and under 0,982, and this means that these variables are statistically significant. Oral Discussion was classified by teachers as the technique which resulted in better results as well as being the easier to implement.

- **This evaluation pattern resulted in average grades and was medium-hard to   implement.**

   According to Figure 17, and since the best and worst techniques were identified, the remaining, which have a p-value above 0,011 and below 0,722 and do reject the alternative hypothesis are the ones that are not as statistically significant as Oral Discussion, Dialogue Simulation and Online Presentation but are still quite significant. Results with average grades that were medium-hard to implement

showed that these were the predominant tools: Traditional Tests, Work and Peer Review, Text Processors and Quizzes.

- **This evaluation pattern resulted in bad grades and was hard to implement.**

  According to Figure 17, the most difficult evaluation technique which was not adequate and resulted in bad grades was Educational Games, this had a p-value bellow 0,05 which means that it rejects the null hypothesis and that this variable is not statistically significant, and, so, this technique did not have an impact in this study.

- **The evaluation pattern was adequate.**

  According to table Figure 16 as well as Figure 17, all techniques rejected the null hypothesis but the Educational Games. This is quite normal once that this is an easier tool to implement in presential classes than in online classes.

- **No specific evaluation pattern was used.**

  According to the table Figure 16, there was an evaluation pattern, in general, both in basic and secondary schools, being that the most three significant evaluation techniques were firstly Dialogue Simulation, with a p-value of 0,95, secondly Work and Peer Review, with a p-value of 0,826, and thirdly the Quizzes, with a p-value of 0,720. All techniques rejected the null hypothesis but the Educational Games, as expected, which had a p-value below, as can be seen on table 17. It can be stated that there were several evaluation patterns used being that the most used ones were Dialogue Simulation, Work and Peer Review and Quizzes.

Chapter 5 – Final Considerations

The biggest motivation in this investigation was to be able to scientifically contribute for education, which is and will always be a part of the daily lives of each and every single person.

The truth is that this pandemic accelerated a process that was already at a course that was the passage of many components of education from presential to online. As a pandemic, this huge neither a remote emergency state had ever happened, there had never been the necessity of passing from 100% face-to-face to 100% online. Even though this brought many social-economic problems, it opened an opportunity for studies such as this one, which was to find evaluation patterns during the ERT.

This dissertation's main goal was to analyze how the online evaluation in mandatory education, basic and secondary, was implemented during the emergency remote education state. Some of the specific goals looked to answer simple questions such as, if grades increased or decreased during ERT, what was the best online evaluation technique, what was the difficulty while implementing the technique if there was an evolution in learning if the ERT caused changes in grades, and so. A larger sample would certainly have a more consistent analysis. However, we came to a very solid conclusion. In addition to the lack of a larger sample to consider for analysis, no gaps were identified that could alter or even discredit the analysis.

The data analysis and treatment was performed mostly between Excel and SPSS – version 25. The plan of analysis to pursuit was: Validation of data using Normal distribution, analysis of Mean and Standard Deviation to extract simple conclusions form the data collected, Levene Test to validate homogeneity and follow through to the correlation and association of the data using Cluster Analysis, ANOVA and Spearman Correlation.

The results were as conclusive as possible. The analysis proved that 68% of the data collected was within one standard deviation of the mean, 95% was within two standard deviations of the mean and 99,7% was within three standard deviations of the mean. This allowed to follow through with the investigation in the beginning.

Regarding the analysis itself, there were 37 quantitative questions that were submitted to the normality test. 82% of them had a p-value higher to 0,05 and the other 18% had a p-value not lower than 0,01. This means that the study follows, as expected, a normal distribution. With the normal distribution validated, the statistics such as mean and standard deviation of the several quantitative variables were analyzed. According to the mean and standard deviation, teachers alleged the following:

- Students were more distracted in online classes

- Students were less engaged in the class

- Students were affected negatively as well as their learning and evaluation process

- Technology was not hard to use for teachers

- Lecturing online is not the same as face-to-face lectures

- Oral Discussion is the most used evaluation technique according to teachers' opinions - basic and secondary education.

- Educational Games is the less used evaluation technique according to teachers' opinions from basic and secondary education.

The Leven test results proved that these samples are homogeneous and, as such, we conclude that most of the variables analyzed influence our answers and proceed to the last and most important analysis, the cluster analysis.

Based on Figure 14, there are two groups of clusters and they become bigger and more heterogeneous as we analyze the graphic from a bottom up, the following was concluded:

- Teachers in cluster 1 tend to agree more with the use of the online evaluation tools described in Figure 15, over-all, there is a big tendency and preference with Oral Discussion and Dialogue simulation and a clear disbelief for Traditional test and Educational Games.

- Teachers in cluster 2 tend to disagree more with the use of the online evaluation tools described in Figure 15, over-all, there is a greater preference for Oral Discussion

and Dialogue simulation and a greater disbelief for Educational Games, Dialogue Simulation and Work and Peer Review.

The conclusions about the patterns discovered are that in general, all techniques seemed to be used and were medium-easy challeging to implement. The less used for both Basic and Secondary schools are Educational Games that, consensually, was the variable that had less significance.

The most adequate and less difficult evaluation techniques to implement that resulted in good grades and were easy to implement were Oral Discussion, Dialogue Simulation and Online Presentation, being that Oral Discussion was classified by secondary teachers as the technique which resulted in better results as well as being the easier to implement. The ones that resulted in average grades and were medium-hard to implement were Traditional Tests, Work and Peer Review, Text Processors and Quizzes, either for basic and secondary schools, and the most difficult evaluation technique which was not adequate and resulted in bad grades was Educational Games, as said, this technique did not have an impact in this study.

All the evaluation techniques stated above were adequate but the Educational Games. This is quite normal given that traditional tests are more commonly used in presential teaching than in remote teaching once that it is more difficult to implement remotely. It can be stated there was an evaluation pattern, in general, both in basic and secondary schools, being that the most three significant evaluation techniques were firstly Dialogue Simulation, secondly Work and Peer Review, and thirdly the Quizzes.

There might be a possibility to follow through and escalate this study, perhaps with the use of social media as well as a deeper analysis of the qualitive data (Appendices B), on the future and it may be equally interesting to use longitudinal surveys as a way of understanding how assessment develops over time in schools, whether in ERT or outside of it.

5.1. Research Limitations

No investigation is entirely free of error, or at least none that includes all possible

scenarios. Therefore, this investigation is no exception. According to Price (2004), there are two most common important groups of limitations. These are represented by threats to internal validity and threats to external validity (Price, 2004). Both can affect the outcome of the research. Internal validity threats can affect obtaining more accurate results, e.g., unfinished surveys, surveys answered by people who are not the target respondent. External validity threats are those that compromise sample results, e.g., sample size (Price, 2004). These threats must be considered because they can affect, indeed, the research quality and outcome. The quality and reliability of the answers are very important issues, perhaps the most important. It is better to have fewer honestly answered responses than to have many surveys answered by non-targets or not concluded. This may be one of the limitations of the inquiry.

The fact that this is a study on a relatively new topic. Distance learning during ERT carries risks and means that not that much research has been done on the topic, and this also means that there is not much information available to recapitulate. So, the information was gathered from different platforms with articles from several different authors and it was analyzed and considered to gather more concise data.

Bibliographic References

9 Proven Ways to Assess Student Learning Online. (2020). Retrieved November 6, 2020, from https://www.ispringsolutions.com/blog/8-ways-to-assess-online-student-learning

A. Kirschner, P. . N., & Mirjam. (2020). Emergency Online Teaching ≠ Online Learning – 3-Star learning experiences. Retrieved January 4, 2021, from https://3starlearningexperiences.wordpress.com/2020/11/05/emergency-online-teaching-≠-online-learning/

Abdul Rahim, A. F. (2020). Guidelines for Online Assessment in Emergency Remote Teaching during the COVID-19 Pandemic. Education in Medicine Journal, 12(2), 59–68. https://doi.org/10.21315/eimj2020.12.2.6

Aloo. (2021, March 24). Processo limpeza de dados. Retrieved from https://blog.aloo.com.br/limpeza-de-dados/

Aras Bozkurt, & Ramesh Sharma. (2020). Emergency remote teaching in a time of global crisis due to CoronaVirus pandemic. Asian Journal of Distance Education, 15(1), 1–6.

Barreto, A. C. F., & Rocha, D. S. (2020). COVID 19 E EDUCAÇÃO: RESISTÊNCIAS, DESAFIOS E (IM)POSSIBILIDADES. Revista Encantar, 2(1), 01–11. https://doi.org/10.46375/encantar.v2.0010

Barros, V., Junior, S., Carlos, J., & Monteiro, S. (2020). EDUCAÇÃO E COVID-19: AS TECNOLOGIAS DIGITAIS MEDIANDO A APRENDIZAGEM EM TEMPOS DE PANDEMIA EDUCATION AND COVID-19: DIGITAL TECHNOLOGIES MEDIATING THE LEARNING PROCESS IN TIMES OF A PANDEMIC EDUCACIÓN Y COVID-19: LAS TECNOLOGÍAS DIGITALES MEDIANDO EL APRENDIZAJE EN TIEMPOS DE PANDEMIA. Revista Encantar - Educação, Cultura e Sociedade, 2, 01–15. https://doi.org/10.46375/encantar.v2.0011

Bernard, R. M., Abrami, P. C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., … Huang, B. (2004). How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. Review of Educational Research. SAGE Publications Inc. https://doi.org/10.3102/00346543074003379

Bland, J. M., & Altman, D. G. (1997). Statistics notes: Cronbach's alpha. BMJ, 314(7080), 572. https://doi.org/10.1136/bmj.314.7080.572

Boodie, K. (2019). Normal Distribution of Data: Examples, Definition & Characteristics. Retrieved from https://study.com/academy/lesson/normal-distribution-of-data-examples-definition-characteristics.html

Brown, J. D. (2002). The Cronbach alpha reliability estimate The Cronbach alpha reliability estimate How should we interpret Cronbach alpha? JALT Testing & Evaluation SIG Newsletter (Vol. 6). Retrieved from http://jalt.org/test/bro_3.htm

Brown, N. (2006). The development of a questionnaire assessing metacognitive patterns of students majoring in accounting in higher education The Development of a Questionnaire Assessing Metacognitive Patterns of Students Majoring in Accounting in Higher Education. https://doi.org/10.1080/09639280600850810

CANI, J. B., Sandrini, E. G. C., SOARES, G. M., & SCALZER, K. (2020). EDUCAÇÃO E COVID-19: A ARTE DE REINVENTAR A ESCOLA MEDIANDO A APRENDIZAGEM "prioritariamente" PELAS TDIC. Revista Ifes Ciência, 6(1), 23–39. https://doi.org/10.36524/ric.v6i1.713

Castaman, A. S., & Szatkoski, E. (2020). Educação a distância no contexto da educação profissional e tecnológica: considerações em tempos de pandemia. Research, Society and Development, 9(7), e491974399. https://doi.org/10.33448/rsd-v9i7.4399

Chan, H. K., Wang, X., Lacka, E., & Zhang, M. (2016). A Mixed-Method Approach to Extracting the Value of Social Media Data. Production and Operations Management, 25(3), 568–583. https://doi.org/10.1111/poms.12390

Chen, J., & Scott, G. (2021). Normal Distribution. Retrieved from https://www.investopedia.com/terms/n/normaldistribution.asp

Clifford, N., Cope, M., & French, S. (2016). Key Methods in Geography. Retrieved from https://books.google.pt/books?hl=pt-PT&lr=&id=7hcFDAAAQBAJ&oi=fnd&pg=PA143&dq=types+of+online+interviews&ots=TDLMtl3UhC&sig=3qeMoW2hWkqYRsU6_4MJ0ZCM2Vc&redir_esc=y#v=onepage&q=types of online interviews&f=false

Davis, F. B. (1964). Educational measurements and their interpretation: Davis, Frederick B: Amazon.com: Books. Wadsworth Pub. Co. Retrieved from https://www.amazon.com/Educational-measurements-their-interpretation-Frederick/dp/B0007DEJCK

Doherty, A., Riordan, T., & Roth, J. (2002). Student Learning: A Central Focus for Institutions of Higher Education.

Duffy, T., Gilbert, I., Kennedy, D., & Kwong, P. W. (2002). Comparing distance education and conventional education: Observations from a comparative study of post-registration nurses. ALT-J, 10(1), 70–82. https://doi.org/10.1080/0968776020100110

e Oliveira, J. B. A., Gomes, M., & Barcellos, T. (2020). Covid-19 and back to school: Listening to evidence. Ensaio, 28(108), 555–578. https://doi.org/10.1590/S0104-40362020002802885

Ebner, C., & Gegenfurtner, A. (2019). Learning and Satisfaction in Webinar, Online, and Face-to-Face Instruction: A Meta-Analysis. Frontiers in Education, 4, 92. https://doi.org/10.3389/feduc.2019.00092

Ferri, F., Grifoni, P., & Guzzo, T. (2020). Online Learning and Emergency Remote Teaching: Opportunities and Challenges in Emergency Situations. Societies, 10(4), 86. https://doi.org/10.3390/soc10040086

Fojtik, R. (2015). Comparison of Full-Time and Distance Learning. Procedia - Social and Behavioral Sciences, 182, 402–407. https://doi.org/10.1016/j.sbspro.2015.04.804

Fox, Anne, M., & Hackerman. (2003). Evaluating and Improving Undergraduate Teaching in Science, Technology, Engineering, and Mathematics. Evaluating and Improving Undergraduate Teaching in Science, Technology, Engineering, and Mathematics. National Academies Press. https://doi.org/10.17226/10024

Freitas Henrique, Oliveira Mírian, Z. Saccol Amarolinda, M. J. (1999). O método de pesquisa survey.

Gonzalez, T., De la Rubia, M. A., Hincz, K. P., Comas-Lopez, M., Subirats, L., Fort, S., & Sacha, G. M. (2020a). Influence of COVID-19 confinement on students' performance in higher education. PLoS ONE, 15(10 October). https://doi.org/10.1371/journal.pone.0239490

Gonzalez, T., De la Rubia, M. A., Hincz, K. P., Comas-Lopez, M., Subirats, L., Fort, S., & Sacha, G. M. (2020b). Influence of COVID-19 confinement on students' performance in higher education. PLoS ONE, 15(10 October), 1–23. https://doi.org/10.1371/journal.pone.0239490

Hair et al. (2005). Análise Multivariada de dados. Retrieved from
https://pt.slideshare.net/ngsouza/livro-analise-multivariada-de-dados-hair-et-al

INE, P. (2020). PORDATA - Agregados domésticos privados com computador, com ligação à
Internet e com ligação à Internet através de banda larga (%). Retrieved January 9, 2021,
from
https://www.pordata.pt/Portugal/Agregados+domésticos+privados+com+computador++c
om+ligação+à+Internet+e+com+ligação+à+Internet+através+de+banda+larga+(percenta
gem)-1158-9361

Jales, M. (2015). Variáveis Estatísticas. Retrieved April 2, 2021, from
http://petdocs.ufc.br/index_artigo_id_475_desc_Bioestatística_pagina__subtopico_13_b
usca_

Kaplan, R. M., & Saccuzzo, D. P. (2012). Psychological Testing: Principles, Applications,
and Issues: Kaplan, Robert M., Saccuzzo, Dennis P.: 9781133492016: Amazon.com:
Books (8th ed.). Retrieved from https://www.amazon.com/Psychological-Testing-
Principles-Applications-Issues/dp/1133492010

Khvilon, E., & Patru, M. (2002). Information and communication technologies in teacher
education: a planning guide - UNESCO Digital Library. Retrieved from
https://unesdoc.unesco.org/ark:/48223/pf0000129533

Levine, A. (2016). Research tools: Interviews and questionnaires - Research Methodology in
Education. Retrieved February 14, 2021, from https://lled500.trubox.ca/2016/225

Lima, M. M., Oliveira, A. M. De, Lima, C. S., Teixeira, C., Bernardes, V., Sugita, D. M., …
Arruda, J. T. (2020). Estratégias Avaliativas No Ensino Remoto : Avaliação Terminal Ou
Contínua ?, 126–132.

Macdonald, S., & Headlam, N. (2008). Introductory guide to research methods for social
research. Retrieved from www.cles.org.uk

Malang, B. (n.d.). K-Means Cluster Analysis.

Maroco, J. (2007). Testes paramétricos para comparar populações a partir de amostras
independentes. Analise Estatistica Com SPSS.

Maroco, João, & Garcia-Marques, T. (2006). Qual a fiabilidade do alfa de Cronbach?
Questões antigas e soluções modernas? (Vol. 4).

Martin, F., & Ndoye, A. (2016). Using Learning Analytics to Assess Student Learning in Online Courses. Journal of University Teaching & Learning Practice (Vol. 13). Retrieved from http://ro.uow.edu.au/jutlphttp://ro.uow.edu.au/jutlp/vol13/iss3/7

Matthiensen, A. (2011). Uso do Coeficiente Alfa de Cronbach em Avaliações por Questionários. Retrieved from www.cpafrr.embrapa.br

Meeter, M., Bele, T., Hartogh, C. den, Bakker, T., Vries, R. E. de, & Plak, S. (2020). College students' motivation and study results after COVID-19 stay-at-home orders. https://doi.org/10.31234/OSF.IO/KN6V9

Mertler, C. A. (2002). Patterns of Response and Nonresponse from Teachers to Traditional and Web Surveys. Practical Assessment, Research, and Evaluation, 8(22). https://doi.org/10.7275/2kdf-g675

Mohmmed, A. O., Khidhir, B. A., Nazeer, A., & Vijayan, V. J. (2020). Emergency remote teaching during Coronavirus pandemic: the current trend and future directive at Middle East College Oman. Innovative Infrastructure Solutions, 5(3), 1–11. https://doi.org/10.1007/s41062-020-00326-7

Moraes, H. L. B., Farias, M. A. de F., Nascimento, S. M. do, & Júnior, G. P. dos S. (2020). De ensino presencial para o remoto emergencial: adaptações, desafios e impactos na pós-graduação. Interfaces Científicas - Educação, 10, 180–193. https://doi.org/10.17564/2316-3828.2020v10n1p180-193

Murphy, K. R., & Davidshofer, C. O. (2004). Amazon.com: Psychological Testing: Principles and Applications (6th Edition) (9780131891722): Murphy, Kevin R., Davidshofer, Charles O.: Books. Retrieved from https://www.amazon.com/Psychological-Testing-Principles-Applications-6th/dp/0131891723

Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric Theory: 9780070478497: Medicine & Health Science Books @ Amazon.com. Retrieved from https://www.amazon.com/Psychometric-Theory-Jum-C-Nunnally/dp/007047849X

Odriozola-González, P., Planchuelo-Gómez, Á., Irurtia, M. J., & de Luis-García, R. (2020). Psychological effects of the COVID-19 outbreak and lockdown among students and workers of a Spanish university. Psychiatry Research, 290. https://doi.org/10.1016/j.psychres.2020.113108

Oliveira Almeida, B., Rosalina, L., & Alves, G. (2020). LETRAMENTO DIGITAL EM

TEMPOS DE COVID-19: UMA ANÁLISE DA EDUCAÇÃO NO CONTEXTO
ATUAL RESUMO. Debates Em Educação, 12(28), 1–18. https://doi.org/10.28998/2175-6600.2020v12n28p1-18

Oliveira, R. M. De. (n.d.). ENSINO REMOTO EMERGENCIAL EM TEMPOS DE COVID-19: FORMAÇÃO DOCENTE E TECNOLOGIAS DIGITAIS, 1–18.

Pereira, A. D. J., Narduchi, F., & Miranda, M. G. De. (2020). BIOPOLÍTICA E EDUCAÇÃO: os impactos da pandemia do covid-19 nas escolas públicas. Revista Augustus, 25(51), 219–236. https://doi.org/10.15202/1981896.2020v25n51p219

Phases of teaching. (2019). Retrieved January 11, 2021, from https://physicscatalyst.com/graduation/phases-of-teaching/

Price, J. H. (2004). Research Limitations and the Necessity of Reporting Them. Retrieved from https://search.proquest.com/openview/b6991f124333fca111dfbc6ef96d080c/1?pq-origsite=gscholar&cbl=44607

Rasmitadila, Aliyyah, R. R., Rachmadtullah, R., Samsudin, A., Syaodih, E., Nurtanto, M., & Tambunan, A. R. S. (2020). The perceptions of primary school teachers of online learning during the covid-19 pandemic period: A case study in Indonesia. Journal of Ethnic and Cultural Studies, 7(2), 90–109. https://doi.org/10.29333/ejecs/388

Robles, M., & Braathen, S. (2002). Online Assessment Techniques.

Rosa, J., & Zaboroski, A. (2020). ENSINO REMOTO E PANDEMIA COVID-19 : DESAFIOS E, 57(55), 41–57.

Seoane-Pardo, A. M., & García-Peñalvo, F. J. (2014). Pedagogical patterns and online teaching. In Online Tutor 2.0: Methodologies and Case Studies for Successful Learning (pp. 298–316). IGI Global. https://doi.org/10.4018/978-1-4666-5832-5.ch015

Silva, E., & Menezes, E. (2001). Metodologia da Pesquisa e Elaboração de Dissertação. Portal, 29(1), 121. Retrieved from http://www.mendeley.com/research/metodologia-da-pesquisa-e-elaborao-de-dissertao-4a-edio-revisada-e-atualizada/

Silveira, S. (2020). O Papel dos licenciados em computação no apoio ao ensino remoto em tempos de isolamento social devido à pandemia da COVID-19. Retrieved from https://www.researchgate.net/publication/342431622_O_Papel_dos_licenciados_em_computacao_no_apoio_ao_ensino_remoto_em_tempos_de_isolamento_social_devido_a_p

andemia_da_COVID-19

Tullis, T., & Albert, B. (2013). Hierarchical Cluster Analysis - an overview. In Measyring the User Experience (2nd ed.). Retrieved from https://www.sciencedirect.com/topics/computer-science/hierarchical-cluster-analysis

Vieira, M. de, & Silva, C. M. S. da. (2020). A Educação no contexto da pandemia de COVID-19 : uma revisão sistemática de literatura, 1013–1031. https://doi.org/10.5753/RBIE.2020.28.0.1013

Vieira, G. S. (2020). ENSINO REMOTO DE EMERGÊNCIA : reconhecimento e usos dos ambientes virtuais de aprendizagem pelos professores do curso de Direito / Ceres, 11–17.

Vue, D. C. (2008). The three eras of Hmong educational history in Laos: French colonial, Laotian independence, and USAID. Retrieved from https://www.proquest.com/openview/0e3d32e4830b701a78e7e3d33171066c/1?cbl=18750&pq-origsite=gscholar

Appendices A – Survey presented to teachers

A avaliação dos alunos no ensino obrigatório (básico e secundário) durante o Ensino Remoto de Emergência

Estimado(a) Professor(a),

Estamos a desenvolver um estudo sobre a existência de padrões na avaliação no ensino obrigatório português (básico e secundário) no Ensino Remoto de Emergência (ERE) entre os meses de Março a Junho de 2020 no ano letivo 2019/2020. Agradecemos, desde já, por despender parte do seu tempo no preenchimento deste questionário.

Este questionário destina-se a professores do Ensino Básico e Secundário que exercem funções em Portugal. O principal objetivo é analisar como foi feita a avaliação neste tipo de ensino durante os vários ERE que foram existindo com base na experiência e opinião dos professores de forma a conseguirmos padronizar os melhores métodos de avaliação durante as várias fases avaliativas do ERE. Estes dados permitirão compreender os desafios que os professores enfrentam durante a tomada de decisão de avaliação do aluno, pelo que a sua colaboração é imprescindível. Asseguramos o seu anonimato e a confidencialidade das suas respostas.

O preenchimento do questionário terá uma duração aproximada de 7 minutos. Todas as informações recolhidas são absolutamente anónimas e confidenciais, não sendo recolhidas informações que permitam a sua identificação individual. Todos os relatórios e publicações deste estudo utilizarão dados agregados, não sendo identificadas opiniões individuais. A sua participação é voluntária podendo abandonar o estudo em qualquer momento, sem qualquer prejuízo pessoal ou profissional.

Agradecemos antecipadamente a sua colaboração.

Pedido de autorização *

○ Sim, concordo em participar neste estudo

○ Não, não concordo em participar neste estudo

Dados do professor

Descrição (opcional)

Género

○ Feminino

○ Masculino

○ Prefiro não responder

Idade

○ Menos de 30

○ 30 a 39

○ 40 a 49

○ 50 a 59

○ 60 ou mais

Leciona Ensino Básico? *

◯ Sim

◯ Não

Leciona Ensino Secundário? *

◯ Sim

◯ Não

Área Disciplinar *

Texto de resposta curta

Anos de experiência *

◯ Menos de 5

◯ 5 a 10

◯ 10 a 20

◯ 20 a 30

◯ Mais de 30

Secção 3 de 6

Título da secção (opcional)

Passagem de presencial para remoto durante o Ensino Remoto de Emergência (ERE)

Com base na sua experiência durante o ERE, encontrou algum dos problemas descritos? *
Assinale, para cada afirmação, a opção mais adequada entre 1 (Discordo Totalmente) e 5 (Concordo Totalmente).

	1	2	3	4	5
Os alunos não …	◯	◯	◯	◯	◯
Os alunos esta…	◯	◯	◯	◯	◯
Os alunos esta…	◯	◯	◯	◯	◯
Os meios tecno…	◯	◯	◯	◯	◯
Os meios tecno…	◯	◯	◯	◯	◯
A formação do…	◯	◯	◯	◯	◯
Sem problemas…	◯	◯	◯	◯	◯

Na sua opinião, o ERE impactou negativamente o processo de aprendizagem e a avaliação dos *
alunos?

	1	2	3	4	5	
Discordo totalmente	○	○	○	○	○	Concordo totalmente

Se respondeu que concorda com a pergunta anterior, em que medida acha que o ERE afetou
negativamente o processo de aprendizagem e a avaliação dos alunos? (Opcional)

Texto de resposta longa

Na sua opinão, durante o ERE, as provas avaliativas permitiram ao professor saber se a matéria *
foi bem leccionada?

	1	2	3	4	5	
Discordo totalmente	○	○	○	○	○	Concordo totalmente

Secção 4 de 6

Untitled Section

A avaliação durante o ERE

Na sua opinião, durante o ERE, as avaliações permitiram identificar se: *

☐ os alunos que precisam de ajuda suplementar

☐ os objetivos e conteúdos programáticos foram alcançados

☐ os alunos dominam os conteúdos ensinados

☐ é necessário avançar ou retomar matéria

☐ é necessário trabalhar alguns pontos em particular com os alunos

☐ Todas as anteriores

☐ Nenhuma das anteriores

Na sua opinião, as avaliações durante o ERE permitiram ao professor saber se o aluno adquiriu
os conhecimentos previstos correspondentes aquele período em avaliação?

	1	2	3	4	5	
Discordo totalmente	○	○	○	○	○	Concordo totalmente

Na sua opinião, durante o ERE, as técnicas de avaliação que implementou foram suficientes para
identificar o que é necessário trabalhar nos alunos?

	1	2	3	4	5	
Discordo totalmente	○	○	○	○	○	Concordo totalmente

Na sua opinião, durante o ERE, a avaliação deve ter em conta todo o processo de desenvolvimento do aluno, assim como, a avaliação do mesmo durante as provas avaliativas online, ou, as provas avaliativas online bastam para atribuir uma nota ao aluno?

○ A avaliação deve ter em conta todo o processo de desenvolvimento do aluno, assim como a avaliação oni

○ A avaliação deve ter em conta apenas as provas avaliativas online

○ Todas as anteriores

○ Nenhuma das anteriores

Uma vez que os métodos de avaliação presencial são diferentes dos remotos, pedimos que, se possível, descreva um exemplo de avaliação presencial que tenha tido que transformar/adaptar a remota devido ao ERT (Opcional)

Texto de resposta longa

Na sua opinião, um aluno que não cumpra resultados satisfatórios deve ser retido mesmo durante ERE?

	1	2	3	4	5	
Discordo totalmente	○	○	○	○	○	Concordo totalmente

Com base na sua experiência, durante o ERE, a taxa de retenção aumentou? *

	1	2	3	4	5	
Discordo totalmente	○	○	○	○	○	Concordo totalmente

Com base na sua experiência, durante o ERE, as notas dos alunos aumentaram? *

○ Sim

○ Não, diminuíram

○ Não, mantiveram-se

Untitled Section

Instrumentos de avaliação usados durante o ERE

Com que frequência utlizou Quizzes para avaliação online? *

	1	2	3	4	5	
Não utilizei	○	○	○	○	○	Utilizei em todas as avaliações

Se respondeu 2 ou mais na questão anterior indique em que medida considerou o instrumento adequado para avaliação online

	1	2	3	4	5	
Nada adequado	○	○	○	○	○	Muito adequado

Com que frequência utlizou Apresentação (ex.Power Point, prezi, etc) para avaliação online? *

	1	2	3	4	5	
Não utilizei	○	○	○	○	○	Utilizei em todas as avaliações

Se respondeu 2 ou mais na questão anterior indique em que medida considerou o instrumento adequado para avaliação online

	1	2	3	4	5	
Nada adequado	○	○	○	○	○	Muito adequado

Com que frequência utilizou um processador de texto (ex. Word) para avaliação online? *

	1	2	3	4	5	
Não utilizei	○	○	○	○	○	Utilizei em todas as avaliações

Se respondeu 2 ou mais na questão anterior indique em que medida considerou o instrumento adequado para avaliação online

	1	2	3	4	5	
Nada adequado	○	○	○	○	○	Muito adequado

Com que frequência utilizou Discussão Oral (ex: teams, zoom) para avaliação online *

	1	2	3	4	5	
Não utilizei	○	○	○	○	○	Utilizei em todas as avaliações

Se respondeu 2 ou mais na questão anterior indique em que medida considerou o instrumento adequado para avaliação online

	1	2	3	4	5	
Nada adequado	○	○	○	○	○	Muito adequado

Com que frequência utilizou Simulação de Diálogo entre alunos (ex: teams, zoom) para avaliação online *

	1	2	3	4	5	
Não utilizei	○	○	○	○	○	Utilizei em todas as avaliações

Se respondeu 2 ou mais na questão anterior indique em que medida considerou o instrumento adequado para avaliação online

	1	2	3	4	5	
Nada adequado	○	○	○	○	○	Muito adequado

Com que frequência utilizou Jogos Didáticos (ex: atividades de arrastar objetos/drag-and-drop, outros) para avaliação online *

	1	2	3	4	5	
Não utilizei	○	○	○	○	○	Utilizei em todas as avaliações

Se respondeu 2 ou mais na questão anterior indique em que medida considerou o instrumento adequado para avaliação online

	1	2	3	4	5	
Nada adequado	○	○	○	○	○	Muito adequado

Com que frequência utilizou Revisão de trabalho e avaliação a pares online (ex. teams, zoom) *
para avaliação online

	1	2	3	4	5	
Não utilizei	○	○	○	○	○	Utilizei em todas as avaliações

Se respondeu 2 ou mais na questão anterior indique em que medida considerou o instrumento
adequado para avaliação online

	1	2	3	4	5	
Nada adequado	○	○	○	○	○	Muito adequado

Com que frequência utilizou Teste tradicional (várias perguntas com tempo limite de término) *
para avaliação online

	1	2	3	4	5	
Não utilizei	○	○	○	○	○	Utilizei em todas as avaliações

Se respondeu 2 ou mais na questão anterior indique em que medida considerou o instrumento
adequado para avaliação online

	1	2	3	4	5	
Nada adequado	○	○	○	○	○	Muito adequado

Relativamente às questões anteriores, existe alguma técnica que tenha utizado e não tenha sido
mencionada? Se sim, por favor, mencione qual/quais.

Texto de resposta longa

Se respondeu que sim na questão anterior indique em que medida considerou o instrumento
adequado para avaliação online

	1	2	3	4	5	
Nada adequado	○	○	○	○	○	Muito adequado

Das técnicas avaliativas abaixo, escolha/mencione a/as que obtiveram melhores resultados: *

☐ Quiz

☐ Apresentação (ex: PowerPoint, ...)

☐ Composição escrita sobre certo tema (ex:Word)

☐ Discussão oral (virtual)

☐ Simulação de Diálogo

☐ Jogos ditáticos (ex:Atividades de arrastar objetos (drag-and-drop))

☐ Revisão de trabalho e avaliação a pares online

☐ Teste tradicional (várias perguntas com tempo limite de término)

Das técnicas avaliativas abaixo, escolha/mencione a/as que obtiveram piores resultados: *

☐ Quiz

☐ Apresentação (ex: PowerPoint, ...)

☐ Composição escrita sobre certo tema (ex:Word)

☐ Discussão oral (virtual)

☐ Simulação de Diálogo

☐ Jogos ditáticos (ex:Atividades de arrastar objetos (drag-and-drop))

☐ Revisão de trabalho e avaliação a pares online

☐ Teste tradicional (várias perguntas com tempo limite de término)

Appendices B – Suggestions, difficulties described on the Free Answer Questions

Se respondeu que concorda com a pergunta anterior, em que medida acha que o ERE afetou negativamente o processo de aprendizagem e a avaliação dos alunos? (Opcional)

45 respostas

- Distração
- Nem todos os alunos tinham as tecnologias necessárias para o ERE, logo o processo de aprendizagem não foi positivo
- Menor esforço e aprendizagem.
- A aprendizagem, atenção é menor. Necessitando de um maior incentivo pelos alunos.
- Apreensão de conteúdos pela falta de meios opcionais levou a uma desmotivação generalizada dos alunos mas sobretudo dos docentes
- Duscriminaçao negativa em relação aos alunos do ensino inclusivo
- No sentido que não há proximidade, os alunos distraem-se facilmente e não adotam a postura correta
- Os alunos não sistematizaram os conteúdos em ERE. Não adquiriram nem metade das aprendizagens
- A comunicação fica muito limitada
- Falta de presença do professor
- Menor concentração
- Participação em aula muito menor face ao presencial; Possibilidade de consulta em todos os exercicios; Avaliação dos mesmos não demonstra os conhecimentos adquiridos.
- Aquisição e aplicação dos conhecimentos
- Dificuldade na aquisição de conhecimentos
- As aprendizagens não aconteceram
- Os meios digitais não permitem reter tanta informação
- Os alunos com mais dificuldades, raramente colaboravam sem ser solicitados e alguns nem o faziam
- Os alunos não conseguem assimilar
- Os alunos não estavam atentos nem sabiam o tema da aula.
- Os ajustes aos critérios de avaliação acabaram por valorizar mais uma componente teórica subvalorizado a prática (sendo esta fundamental em aulas práticas).
- As aprendizagens em termos de atividades fisicas sÃo claramente prejudicadas devido à falta de meios e possibilidades de situações de exercicio para os alunos se desenvolverem no ERE
- Os alunos não aprenderam muito conteudo
- Não se verificou verdadeira aprendizagem, a avaliacao nao é fidedigna.

- alunos que obtinham ajuda a realizar elementos avaliativos; justificações de falha de net constantes;

- Nao tiraram duvidas e tinham dificuldades em compreender.

- Os alunos nao se envolveram nas aulas; nao houve verdadeira aprendizagem

- Mais Distração; Pouca participaçao

- Maior a distração, maior desmotivação

- Muito dificil para acompanhar os alunos com dificuldades de aprendizagem

- Aprofundamento de materias diminuiu.

- Os alunos aplicaram-se menos.

- Os alunos perderam ritmo

- A maioria dos professores nao adaptou o metodo de ensino, nem a avaliacao, nas aulas sincronas, desmotivando os alunos e colocando os docentes em desespero total.

- Falhas na rede de internet, distracoes

- empenho e concentracao dos alunos com mais dificuldades.

- Embora me pareça que mais a aprendizagem do que a avaliação, muito do que negativamente ocorreu nada mais foi do que pôr a nu uma realidade que se vive nas escolas e no sistema educativo, há duas décadas, pelo menos. Só que a farsa que se tem assegurado dentro das salas de aula não foi possível de manter à distância. Porque os professores não estavam no espaço deles e nada podiam controlar.

- Os alunos com mais dificuldades, com menos autonomia, com menos acompanhamento parental, saíram prejudicados

- Sendo mais dificil dar a aula é mais dificil determinar a aprendizagem dos alunos e como tal mais dificil avaliar.

- Sendo defensor do ensino híbrido, o que notei foi que os procedimentos definidos pela escola não se adequavam às exigências do ensino híbrido e que os alunos entenderam o ERE como uma oportunidade para não estudar.

- Menos concentração.

- Os meios tecnológicos.

- Por falta da presença e orientação do professor.

- Afetou parcialmente.

- falta de apoio presencial na realização de trabalhos práticos.

- Em parte.

Uma vez que os métodos de avaliação presencial são diferentes dos remotos, pedimos que, se possível, descreva um exemplo de avaliação presencial que tenha tido que transformar/adaptar a remota devido ao ERT (Opcional)

53 respostas

- Tipo de testes de avaliação.
- Simulação prática em grupo
- Esquematização da matéria no momento
- Avaliação de conteúdos presencial
- teste de avaliação escrito presencial foi efectuado na plataforma Moodle. A vigilância do mesmo dependia inteiramente das câmeras, sendo por vezesbdifícil perceber se os alunos teriam cábulas ou se estariam a copiar.
- Apresentação de trabalhos em grupo
- Questões aula
- Na minha área disciplinar, todas
- Realização de tarefas em aula pnline
- aplicar em cada aula uma tarefa, para enviar no proprio dia
- Momentos avaliativos pelo Google forms.
- Avaliação prática
- Apresentações/Testes orais
- Apresentação/testes orais
- Recorrer avaliação formativa
- Trabalhos individuais com os conteudos abordados.
- Fichas de avaliação
- Testes e fichas
- avaliar o desempenho fisico do aluno por, apenas, teoria (fundamental em Educação Física.
- Realização de circuitos de Condição Fisica por parte dos alunos em suas casas
- Forma de demonstraco do conteúdo
- Forms durante a aula online
- Avaliacoes semanais curtas
- Trabalhos de grupo, avaliacao oral
- Avaliaçao formativa
- Teste sumativo
- teste
- Preenchimento de formulário em aula
- Avaliacao da execuçao e resolusao de praticas
- testes escritos
- Avaliar a atencao... Tive que ter duas camaras, uma para ver os alunos outra para projetar a aula
- questionÃ¡rio FORMS que permita aos alunos e responder e submeter o trabalho
- Houve necessidade de reformular a tipologia de atividades propostas num teste escrito, assim como a forma de avaliar a oralidade do aluno.
- questionarios online
- Verificao das atividades desenvolvidas, e a possibilidade de orientar o aluno para melhorar a sua atividade.

- Utilizavao de aplicacoes online em detrimento de programas instalados nos computadores.
- Avaliacao da interpretacao oral de um texto, recorrendo a um formulario de escolha multipla.
- teste de avaliação
- Atividades praticas
- Escrita digital
- Fichas de Avaliação
- A questão do Feedback presencial/ online
- Testes de avaliação
- A ausência de investimento na Educação Artística condiciona aquilo que o docente pode desenvolver com os alunos em sala de aula (continuamos com os lápis de cor, as canetas de feltro, os lápis de cera e, cada vez menos, os guaches). Ao passarmos para o ERT, nada disso fazia sentido, porque não se ensina a modelação da motricidade fina, a crianças e adolescentes, à distância, mediados por um teclado, um rato e um ecrã. Como os alunos e a escola não dispõem de meios digitaisn(software e hardware) para as técnicas possíveis no ERT, nada mais houve, com seriedade, a fazer do que apostar na reflexão sobre o fenómeno artístico e a construção do discurso crítico em Artes Visuais... o que não é a mesma coisa, convenhamos.
- Trabalho de pares, de grupo, de investigação
- Utilização de grelhas de observação. Presencial basta circular pela turma para a preencher, online exige que os alunos partilhem as suas telas o que é muito dificil individualmente
- Não creio que essa diferença seja assim tão evidente e não possa ser colocada em causa... Na minha opinião, o que sucedeu foi que melhorei a forma como trabalhalho com os alunos. Por exemplo, passei a discriminar rubricas de avaliação em vez de apenas dar indicações sobre a realização dos trabalhos.
- Fichas de avaliação
- Teste de consulta.
- Avaliação direta, por observação.
- Apresentação de conteúdos.
- Experiência práticas, por simulações teóricas.
- Teste escrito para avaliação de conhecimentos.

> Relativamente às questões anteriores, existe alguma técnica que tenha utlizado e não tenha sido mencionada? Se sim, por favor, mencione qual/quais.
> 25 respostas

- Visualização de vídeos e filmes.
- Visualização de vídeos de apoio
- Apresentação de trabalhos partilhada com a turma
- Comentários criticos/ visualização de filmes c/ guiao de trabalho
- Recursos da plataforma Khan Academy, Kahoot, Forms .

- Auto e hetero avaliação no google forms
- Envio de fotografia da execução de pequenos exercicios no caderno, com tempo definido
- Apresentacao Oral individual
- Apresentacoes orais, criacao de Sites, elaboracao de videos
- Testes online moodle
- Kahoot, Forms
- Google Forms
- Formulário como teste (forms)
- peardeck
- Trabalhos a pares e em grupo
- Kahoot, google forms
- Nao
- Resolução de exercicios, com partilha de ecran
- Desenhar, fotografar, partilhar, reflectir, discutir.
- Fichas formativas para posterior envio vía classroom, com correção nas aulas online
- Não.
- Avaliação de cada conteúdo, trabalho prático.
- As mencionadas.
- Gravação.
- Trabalhos práticos realizadas na aula e enviados para avaliação da participação.