Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi

Journal of Measurement and Evaluation in Education and Psychology

ISSN: 1309-6575

Cilt: 14-Sayı: Özel Sayı
Volume: 14-Issue: Special Issue

Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi

Journal of Measurement and Evaluation in Education and Psychology

ISSN: 1309 – 6575

İletişim
e-posta: epodderdergi@gmail.com
Web: https://dergipark.org.tr/tr/pub/epod

Contact
e-mail: epodderdergi@gmail.com
Web: http://dergipark.org.tr/tr/pub/epod

Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi (EPOD) yılda dört kez yayımlanan hakemli uluslararası bir dergidir. Yayımlanan yazıların tüm sorumluluğu ilgili yazarlara aittir.

Journal of Measurement and Evaluation in Education and Psychology (JMEEP) is a international refereed journal that is published four times a year. The responsibility lies with the authors of papers.

Dizinleme / Abstracting & Indexing
Emerging Sources Citation Index (ESCI), DOAJ (Directory of Open Access Journals), SCOPUS, TÜBİTAK TR DIZIN Sosyal ve Beşeri Bilimler Veri Tabanı (ULAKBİM), Tei (Türk Eğitim İndeksi), EBSCO

The Journal of Measurement and Evaluation in Education and Psychology
JMEEP

Foreword to special issue

Dear Colleagues and Dear Readers,

In this special issue, we are excited and honored to address the topic of "***Technology-Assisted Formative Assessment***". This special issue, which focuses on the importance of technology-assisted formative assessment in the fields of education and psychology in order to evaluate and guide students' learning processes, will play an important role in shaping the learning methods of the future.

Today, with the rapid development of technology, education and training processes are undergoing a major transformation. Technology-supported formative assessment tools and methods, which are used to monitor and support students' learning more effectively, have attracted great interest among educators and researchers. In this special issue, seven articles are included to address different aspects of this interesting topic. These articles provide important insights into how formative assessment is supported and enhanced by technology.

The first article is titled **"The Effect of Using E-Portfolios on The Self-Regulation Skills of Students: A Meta-Analysis Study"**, investigates the effect of e-portfolio applications on students' self-regulation skills. This study provides important information in terms of understanding how students develop their self-regulation skills using e-portfolios.

The second article is titled **"The Role of E-portfolios in Formative Assessment: A Systematic Literature Review"**. This literature review analyses the role and use of e-portfolios in formative assessment in depth.

The third article is titled **"Examining Students' Formative Test-Taking Behaviours Using Learning Analytics"**. This study analyses students' formative test-taking behaviours using learning analytics in detail.

The fourth article is titled **"Design and Development of an Interactive Video Player for Supporting Formative Assessment in Online Learning"**, aims to design and develop an interactive video player to support formative assessment in online learning environments.

The fifth article is titled **"Investigation of the Effect of Online (Web-Based) Formative Assessment Applications on Students' Academic Achievement"**, investigates the effect of web-based formative assessment applications on students' academic achievement. This study evaluates the contribution of technology-supported formative assessment to learning outcomes.

The sixth article is titled, "**Raising 21st Century English Language Teachers in Turkish Context: Development of a Technology-Enhanced Measurement Curriculum**". This study explains how a technology-enhanced assessment curriculum was developed for the training of English language teachers with 21st-century skills in the Turkish education system.

The seventh article is titled "**Learning Analytics in Formative Assessment: A Systematic Literature Review**" discusses how learning analytics contributes to formative assessment and analyses the existing research in depth.

The articles in this special issue provide valuable information about technology-supported methods and tools of formative assessment. We recommend that you read these articles with interest, which will contribute to a better understanding of the impact of technology-supported formative assessment on student learning processes and outcomes.

We would like to thank the authors, reviewers, and editorial board members who contributed to the preparation of this special issue. We are grateful for your support for these studies to shape the future of technology-supported formative assessment.

Best wishes,

Special Issue Editors

Ramazan Yilmaz, Bartın University, ramazanyilmaz067@gmail.com

Ke Zhang, Wayne State University, bb2145@wayne.edu

Nuri Dogan, Hacettepe University, nuridogan2004@gmail.com

Journal of Measurement and Evaluation in Education and Psychology

Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi (Ekim 2023, Sayı: 14 – Özel Sayı)

Journal of Measurement and Evaluation in Education and Psychology (October 2023, Issue: 14 – Special Issue)

İÇİNDEKİLER / CONTENTS

The Effect of Using E-Portfolios on The Self-Regulation Skills of Students: A Meta-Analysis Study
Ceren MUTLUER 287

The Role of ePortfolios in Formative Assessment: A Systematic Literature Review

Ezgi GÜN-TOSİK, Bilal ATASOY, Akça Okan YÜKSEL 299

Examining Students' Online Formative Test-Taking Behaviors Using Learning Analytics

Alper BAYAZIT, Denizer YILDIRIM, Gökhan AKÇAPINAR, Hale ILGAZ 320

Design and Development of an Interactive Video Player for Supporting Formative Assessment in Online Learning

Alper BAYAZIT, Gökhan AKÇAPINAR..332

Investigation of the Effect of Online (Web-Based) Formative Assessment Applications on Students' Academic Achievement

Şeref AKPINAR, Bayram ÇETİN 345

Training 21st Century English Language Teachers in Turkish Context: Development of a Technology-Enhanced Measurement Curriculum

Burcu ŞENTÜRK, Beyza AKSU DÜNYA, Mehmet Can DEMİR 359

Learning Analytics in Formative Assessment: A Systematic Literature Review

Ke ZHANG, Ramazan YILMAZ, A. Berk USTUN, F. Gizem KARAOĞLAN YILMAZ 382

ISSN: 1309 – 6575
Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi
Journal of Measurement and Evaluation in Education and Psychology
Research Article; 2023; 14(Special Issue); 287-298

The Effect of Using E-Portfolios on The Self-Regulation Skills of Students: A Meta-Analysis Study

Ceren MUTLUER*

Abstract

The change of self-regulation skills, one of the variables determining success in education, with measurement and evaluation techniques in the education process has been the subject of many studies. However, each research result leads to different results due to the planned situation and limitations. For this reason, e-portfolios used in the education process were examined with a meta-analysis study to evaluate whether they were effective on students' self-regulation skills from a more general point of view. The research is limited to published academic studies written in Turkish and English between 2000 and 2023. When the criteria determined in the research were examined, only 19 studies were found by these criteria. Theoretically, when the effect sizes in each study were examined, a meta-analysis was carried out with the random effects model. The analyzes of the research were made with the CMA version 3 program. As a result of the research, it was concluded that using e-portfolios greatly and significantly affect students' self-regulation skills.

Keywords: e-portfolio, self-regulation, meta-analysis, random effects model

Introduction

The expression 'academic achievement' is commonly used to refer to education quality within the education system. Many studies have reported that academic achievement used in determining the effectiveness of education is positively related to self-regulation strategies (Eom & Reiser, 2000; Pintrich & De Groot, 1990; Trainin & Swanson, 2005; Üredi & Üredi, 2005).

Performance for academic tasks is explained as academic achievement in the education process. Self-regulation skills play an active role in the process of determining the goals in the process of revealing the work done by the individual in academic tasks, the feedback in the process of reaching the goals, and the evaluation process in terms of concrete products. While describing the notion of self-regulation, Zimmerman (1986) defines it as comprehending students as active contributors to an academic task in terms of cognitive, motivational, affective, and behavioral aspects. Therefore, identifying which self-regulated learning strategies are important and which self-regulation strategies facilitate their use is essential to promote academic performance.

Dent & Koenka (2016) focused on determining the relationship between learning and academic achievement according to the self-regulation strategy in their meta-analysis study. In the meta-analysis study, when the overall effect calculated for metacognitive processes and cognitive strategies for self-regulation strategies was examined, it was seen that there was a significant and moderate effect. As an important variable affecting academic success, self-regulation skills are shaped by measurement and evaluation techniques. Measurement and evaluation processes, especially complementary measurement and evaluation techniques, focus on individuals' awareness of their characteristics. Personal choice and self-control are necessary for students to gain and develop self-regulation skills (Zimmerman, 1989).

Gözüyeşil & Tanrıseven (2017) examined the effectiveness of alternative measurement and evaluation techniques in their study. In their study, it was determined that alternative measurement and evaluation

* PhD, Bolu Abant İzzet Baysal University, Faculty of Education, Bolu-Turkiye, e-mail: cmutluer@yandex.com, ORCID ID: 0000-0002-3935-336X

To cite this article:
Mutluer, C. (2023). The effect of using e-portfolios on the self-regulation skills of students: A meta-analysis study. *Journal of Measurement and Evaluation in Education and Psychology, 14*(Special Issue), 287-298. https://doi.org/10.21031/epod.1267809

Received: 20.03.2023
Accepted: 20.06.2023

techniques increased academic achievement. In a detailed examination, the effect of portfolio applications on academic achievement was higher than other measurement and evaluation techniques.

In this process, it is appropriate to use performance-based measurement and evaluation tools. The use of portfolios is preferred to monitor the initial and developmental processes in self-regulation skills. Portfolio applications provide an authentic and meaningful collection of student work and allow students to accurately demonstrate their success or development (Ekbatani & Pierson, 2000).

As a reflection of technology in the educational process, the portfolio technique has also been moved to the electronic environment and names such as e-portfolio and web-based portfolios have been given. Electronic portfolios have also been promoted as the evolution of traditional portfolios. They have many advantages over paper-based portfolios. These are:

- Easier, even more immediate, access to students' work can be provided by a wider audience, including peers, teachers, parents, and others (Barrett, 2006)
- Increases information communication skills (van Wesel & Prop, 2009)
- Fewer cultural barriers (Wanchid & Charoensuk, 2015)
- Allows two-way communication (without time and place restrictions) (Barrett, 2006)
- Faster feedback is received (Wanchid & Charoensuk, 2015).
- Increases a greater sense of potential and a high sense of pride and achievement regarding the permanence of the content (Campbell & Schmidt, 2005).
- Allows individuals to develop self-regulation skills with the support of technology.
- Prevents the manual storage of portfolio materials.

Due to these advantages, e-portfolio is used as one of the measurement and evaluation techniques of the education process. There are electronic portfolios as a technology-assisted formative assessment. The e-portfolios ensure ease of access, data storage opportunities, time-saving, contribution to teaching, and continuous monitoring of teacher performances. It assists in the collection, update, and management of data. It is a guide in terms of the effectiveness of the evaluation process (Polat & Köse, 2013). Apart from being an evaluation tool, it has been decided that an e-portfolio encourages students to learn, increases their motivation, changes their attitudes and perceptions positively, and increases their success in the process (Barış & Tosun, 2013; Chou, 2012; Demirli, 2007; Demirli & Gürol, 2010; Gülbahar & Köse, 2006).

Many benefits can be mentioned when a literature review is done on the educational inputs, processes, and outputs of e-portfolios. Primarily, the tasks are received in feedback as they are shared with teachers and friends. In addition, E-portfolios support students' individual development by shaping learning materials (Kinash et al., 2012). Among other benefits, e-portfolios encourage students to develop their skills using multimedia components and to reveal all individuals' learning achievements and expectations. In the process of creating an e-portfolio, students are provided with the opportunity to reflect on their learning levels, teachers provide feedback and guidance, the continuity of students' development is ensured, participation in collaborative activities is ensured, and socialization, encouragement, and motivation are provided (Demirli & Gürol, 2010; Ghosh, 2003; Lorenzo & Ittelson, 2005). E-portfolios assessment include the features of self-reflection, self-review, self-monitoring, and self-improvement (Bartlett & Sherry, 2006).

In their study, e-portfolios' benefits are listed under five headings by Jenson & Treuer (2014). These five topics are collection, self-regulation, reflection, integration, and collaboration. Jenson & Treuer (2014) explained these concepts as follows, taking into account the skills of the 21st century.

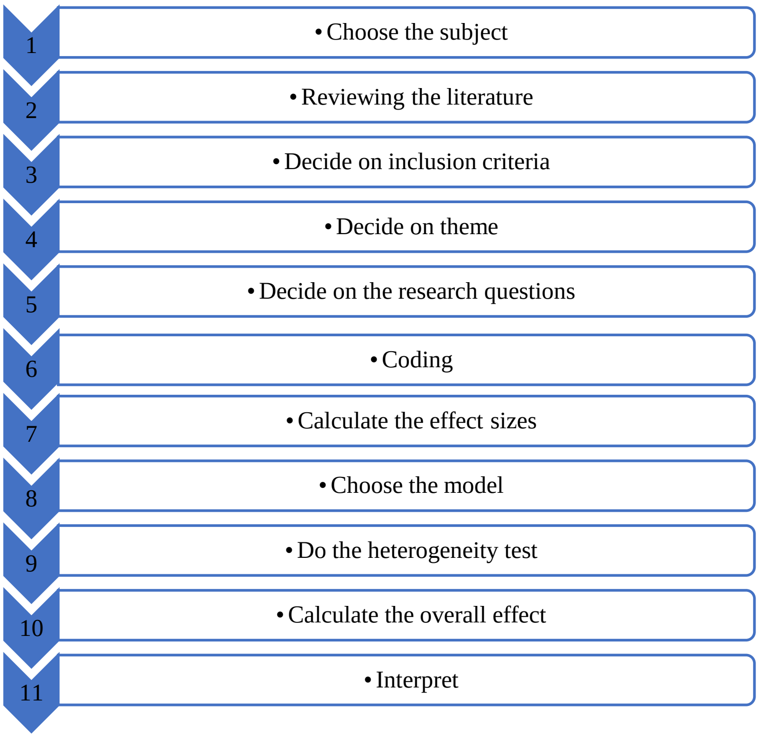
- *Collecting:* Relevant artifacts that demonstrate learning outcomes.
- *Self-Regulating:* Being aware of behavior, students can control and exercise that control for learning.

- *Reflecting:* Contextualizing the meaning and significance of learning, consistent with established goals and values.
- *Integrating:* Synthesizing and transferring learning to any number of situations.
- *Collaborating:* Participating in the community to build knowledge and skills based on existing knowledge.

As one of the benefits mentioned above, the e-portfolio application has a great place in acquiring and developing self-regulation skills. Many studies in the literature use e-portfolios in the input, process, and process-oriented evaluation phase of learning environments. However, considering the differences in the number of study groups, their aim at different times, and the different target audiences, studies have yet to be decided to examine the effect of e-portfolios on self-regulation skills with a holistic perspective. Therefore, this study aims to examine the effect of using electronic portfolio (e-portfolio) in the literature between 2010-2023 on self-regulation skills in learning processes from a holistic perspective. In this study, a meta-analysis study was planned over the studies conducted in the literature between the years 2000-2023 using the variables specified in the purpose of this study.

The meta-analysis is one of the first proposals to test the statistical significance of combined results (Hedges, 1992). Furthermore, a meta-analysis was felt to consider the research results holistically and form a common opinion (Mutluer, 2022). According to Borenstein et al. (2010), meta-analysis studies are studies that can be more generalized as a result of integrating the results of studies with the same or related purpose and reach results that many studies have confirmed. There are eleven sequential steps to conducting a quality meta-analysis (Borenstein et al., 2010; Field & Gillett, 2010; Şen & Yıldırım, 2020):

Figure 1
The Steps of the Meta-Analysis Procedure



The meta-analysis stages was interpreted in Figure 1. The overall effect of the meta-analysis should be interpreted in line with the specified stages.

Methods

The statistical process used to reach a more general result than the results of studies conducted for similar purposes at different times is possible with meta-analysis. According to Glass (1976), *meta-analysis* is the statistical analysis of many analyses emerging from individual studies to integrate the findings. The effects of e-portfolios on the self-regulation skills of individuals have been examined with meta-analysis and the steps to be followed in the process are included. This section includes the criteria determined for the meta-analysis, the studies included in the meta-analysis, the process of dealing with publication bias, and the data analysis.

Data Sources and Search Strategies

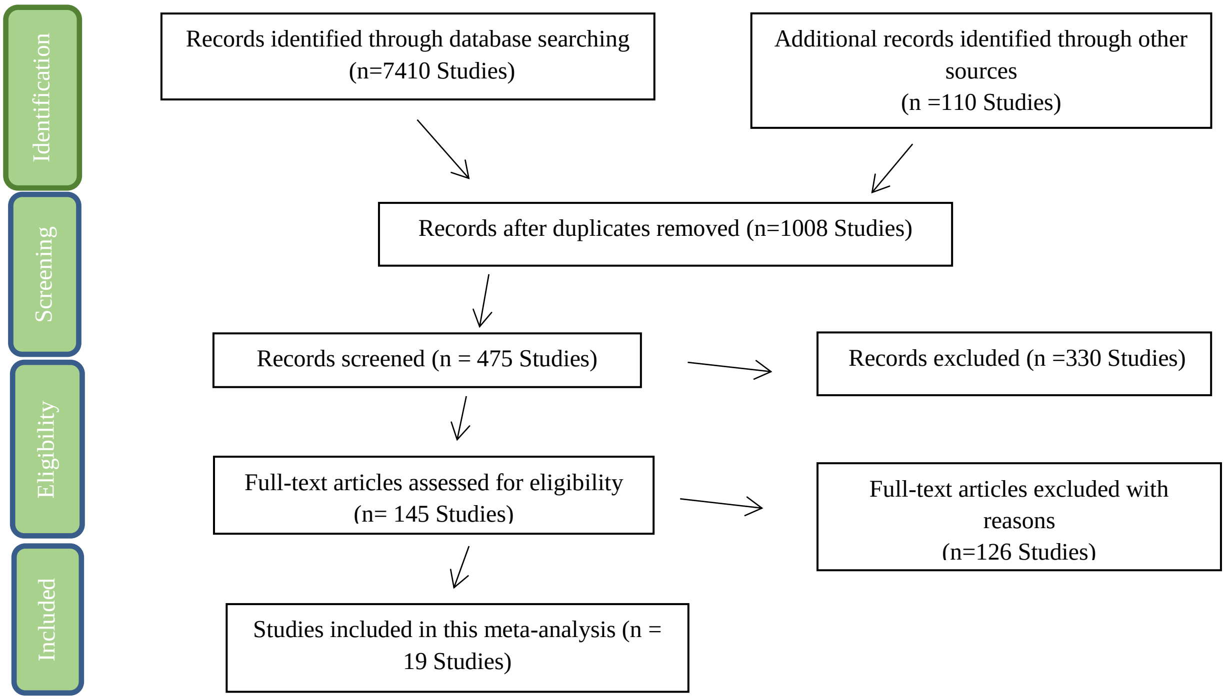
Both published articles and theses were reviewed to obtain a wide range of available resources for meta-analysis. The electronic search consisted of databases including JSTOR, ERIC, ScienceDirect, Wiley Online Library, SAGE Journals, ProQuest Dissertations, YÖK (Council of Higher Education) National Thesis Center, Web of Science, and Google Scholar. The primary search terms were ‘electronic portfolio’, ‘e-portfolio’, ‘digital portfolio’, ‘online portfolio’, ‘web-based portfolio’, ‘self-regulation’, ‘self-regulation skills’, ‘self-regulation strategies’, ‘self-organizing’, and ‘self-reflective’. Considering these keywords, among the studies conducted between 2000-2023, ‘e-portfolios affect the self-regulation skills’. The following criteria were used to select the studies included in the meta-analysis study:

- To be examined the effect of e-portfolio on self-regulation,
- To be published as Master's, doctoral thesis or article,
- To be written only in English and Turkish,
- To be decided the valid studies which were used an experimental design,
- To be included sufficient information (sample size, mean, standard deviation) in studies to calculate the effect size. In the figure below, which studies were selected according to the criteria determined for the meta-analysis are summarized in stages.

Publications were sorted by taking into account the criteria in the meta-analysis process. In the Prisma model in Figure 2, the number of publications in the meta-analysis process and the elimination process according to the criteria are given.

Figure 2

The Flowchart Shows the Selection of Included Studies (PRISMA)



When the literature was examined, the data extraction process started with 7410 studies in which the keyword was mentioned and continued by reducing it to 19 studies. 19 studies were selected by the criteria specified in the study. Information about the studies selected for the research is given in Table 1.

Table 1

Descriptive Explanations for All Included Studies in This Study

Study	Year	Author(s)	Publication Type	Effect Size
1	2007	Cooney	Thesis (Doctoral)	0,569
2	2010	Meyer and others	Article	-0,03
3	2010	Koç	Thesis (Doctoral)	0,9
4	2013	Cheng & Chau	Article	0,497
5	2013	Abrami & others	Article	0,652
6	2013	Alexiou & Paraskeva	Article	0,839
7	2014	Alexiou & Paraskeva	Article	6,322
8	2015	Tseng & Lin	Article	0,512
9	2015	Nguyen & Ikeda	Article	0,006
10	2016	Liang & others	Article	1,638
11	2017	Sasai	Article	3,411
12	2018	Chang *et al.*	Article	0,780
13	2019	Karami *et al.*	Article	0,553
14	2019	Corta	Article	-3,396
15	2019	Alexiou & Paraskeva	Article	1,863
16	2020	Akgün & Şahin Kölemen	Article	0,539
17	2021	Alhitty & Shathawi	Thesis(Master)	1,801
18	2021	Türkkaynağı	Article	0,214
19	2022	Lysenko *et al.*	Article	0,116

The table gives the year, publication type, author information, and effect sizes of 19 studies. When the research was examined, it was found that two studies (%10.5) in 2020, 2012, 2015, 2018, 2019, and 2021 years were related to the subject title. In addition, three of the studies are dissertations. Of these theses, only the work completed by Türkkaynağı in 2021 is a master's thesis. Finally, while only three of these studies were written in Turkish, the remaining 16 studies were written in English.

Publication Bias in the Meta-Analysis Process

Conducting meta-analysis with only studies supporting a certain hypothesis in the research process causes publication bias. Therefore, publication bias should be determined in studies that meet the criteria. In this study, Funnel-Plot, Classic Fail-Safe N analysis, and Egger regression estimation coefficient were used for investigating publication bias. The results of these methods were evaluated together and it was decided whether there was a publication bias or not.

Analysis of Data

Firstly, the effect sizes were calculated. For this research, Comprehensive Meta-Analysis (CMA)-Version 3 program was used to calculate both effect sizes and overall effects in the process.

While investigating the effect of e-portfolios on students' self-regulation skills within the scope of the research, the comparison between pre-test and post-test scores was considered. Therefore, the following equation of sample score means differences will be used for the difference between pre-test and post-test scores.

$$D= \overline{X}_1-\overline{X}_2 \tag{1}$$

$$D= \overline{X}_{Diff}$$

The second equation below was used to calculate the variance for both measurements in the study.

$$V_D=\frac{n_{1-}n_2}{n_1 n_2} S^2_{pooled} \tag{2}$$

In this context, the effect size can be calculated.

$$SE_D=\sqrt{V_D} \tag{3}$$

The Cohen D effect size was then converted to Hedge's g effect size. Hedge's effect size formula will be used to calculate the effect size in the research. In addition, the coefficient developed by Hedges (1982) will be used to calculate individual effect size values:

$$\text{Hedges' } g = (M_1 – M_2) / SD_{pooled} \tag{4}$$

Choosing the appropriate model to determine the overall effect size is important when performing meta-analyses (Srinivasjois, 2021). Where there is no heterogeneity between studies and there is a valid reason to assume that the true effect is constant, it is appropriate to use the fixed effects model (Harrer *et al.*, 2022). However, in cases where the studies will be generalized to a universe with different effects and characteristics, the random effects model should be used based on the assumption that the actual effect may vary from study to study (Hanji, 2017). Therefore, before the analysis, it was decided by the researchers to calculate the overall effect size using the random effects model. In addition, heterogeneity analysis was performed to determine the variability between the effect sizes of the primary studies included in the meta-analysis.

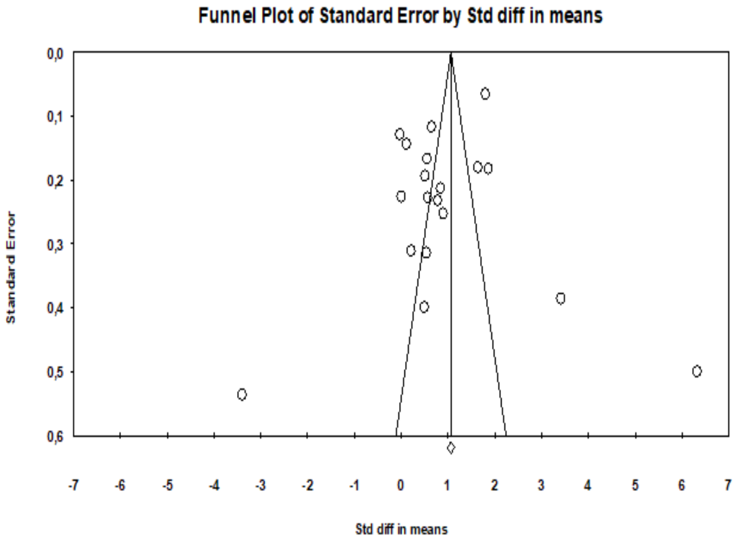
Results

Result of Publication Bias

Including studies that defend and confirm only one hypothesis in the research process causes publication bias. For this reason, the process should be started by examining whether the meta-analysis has publication bias.

In this study, Funnel-Plot, Classic Fail-Safe N analysis, Egger regression estimation coefficient were used to investigate publication bias. With Funnel Plot, the distribution of the effect sizes on the funnel plot is presented. The distributions of the effect sizes of all studies on the Funnel plot are given in Figure 2 below.

Figure 3
The Effect Size Distributions on the Funnel Plot



When the Funnel plot above is examined, it is seen that the distribution around the funnel plot cone is homogeneous, although there is no perfectly symmetrical distribution. This distribution of the effect sizes of the funnel plot indicates the absence of publication bias (Rothstein *et al.*, 2005; Sterne *et al.*, 2011).

As the other analysis bias method, Classical Fail-Safe N value was examined. The overall effect will change significantly if more than 2135 studies are added to the research. Since the number of these studies is quite high, it has been proven again that there is no publication bias. As the last bias analysis, Egger's regression test was used. The regression intercept was insignificant (intercept = -2,664, p = .309). Therefore, the hypothesis was accepted to show that the regression constant did not deviate from zero significantly.

Testing for heterogeneity

Although heterogeneity is a process that needs to be analyzed statistically, it is a process that needs to be decided based on the actual literature. It is not a correct approach to determine a model based on only statistical results. To treat the heterogeneity situation hypothetically, it should be decided whether the effect in the universe differs according to the situations in the study.

According to Borenstein *et al.* (2021), a fixed effects model was proposed if the real effect is unique in the universe and has a constant feature in all studies. In this case, homogeneity is achieved. According to the fixed effects model, each study's effect on the universe is the same as the real effect. Furthermore, it is stated that a single source of error in the fixed effects model is sampling error. As an alternative to the assumptions of the fixed effects model, the fact that the effect in the studies is not equal to the general effect in the universe is explained with the random effects model. The effect in the universe may not be the same in all studies and there may be different subgroups. In this case, different effect sizes are mentioned. The actual effect varies from study to study. For example, the fact that a drug has different effects at different age levels causes different effect sizes to be calculated in the research results. In this model, it is desired to estimate the mean of the distribution of effects. It is not correct to accept a single effect size in the random effects model. According to this model, there are multiple effects in the universe.

It is incorrect to test for heterogeneity in research and choose a model in this context. However, if it is known that there is more than one effect size in the study and the distribution of these effects is theoretically supported, the random effects model is recommended. In this study, the random effects model was chosen because it was theoretically explained that the effect in the universe was different in all studies. Although the literature-supported model was chosen in the process, the heterogeneity tests in the research process are given in the Table 2 below as second evidence.

Table 2

Overall Effect Size and Confidence Intervals for Heterogeneity by Random Effect Model

Number of the Studies	Overall Effect Size	df	Q	Se	I^2	Tau square (τ^2)	Effect size and %95 confidence	
							Lower limit	Upper Limit
19	0,925	18	559,201	0,233	96,781	0,955	0,469	1,381

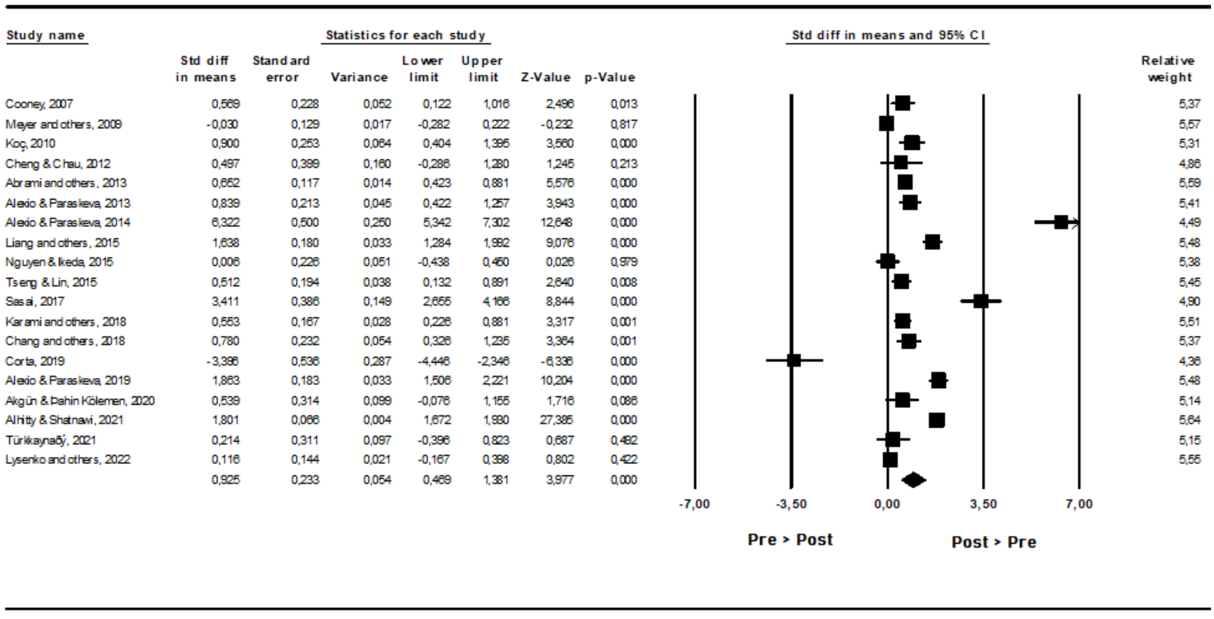
Although it has been stated that the effect in the studies differs in theory, the results are supported by looking at the statistical heterogeneity test for model selection. According to the heterogeneity table, the general effect is mentioned first. The forest-plot table also interprets the overall effect obtained from all effects. Interpreting the Q statistic alone is impractical. A value of Q-df > 0 was obtained. This proves the heterogeneity. Unfortunately, the number of studies (Borenstein *et al.*,2010; Borenstein *et al.*, 2021) affects the Q value. The value of τ^2 is calculated by faulting Q. But independent of the Q value, the heterogeneity value is unaffected by the number of studies. It also allows estimation on the effect size scale (Borenstein *et al.*, 2021; Hedges & Vevea, 1998). The examined τ^2 supports heterogeneity. The I^2 value is interpreted as the explained variance. 96.781% of the observed distribution of the self-regulation variable is explained through the effectiveness of the e-portfolio application considered for the research. According to the random effects model, the effect sizes vary between 0.469 and 1.382. The above heterogeneity analysis results, which are given as a second proof of the theoretically decided random effects model, also argue that the model selection is appropriate.

Result of the meta-analysis

All the effect sizes discussed within the scope of the research were analyzed over the CMA-Version 3 program and the distribution of the effect sizes was given in the forest plot below. The overall effect size was obtained from 19 studies selected in accordance with the criteria in the study.

Figure 4

Forest Plot For All Included Studies



While the 0 point in the forest plot above shows that e-portfolios do not affect self-regulation, it is stated that the effect sizes on the left are significantly higher than the pre-test results of the post-test results and there are effect sizes of the studies showing that the e-portfolio application does not work. The effect sizes of the studies conducted by Meyer *et al.* (2010) and Corta (2019) are included in this section. The part to the right of the reference point summarizes that the post-test results show a significant change from the pre-test results. For the effect sizes on the right, it is seen that e-portfolios have significant and positive effects on self-regulation. Although the effect sizes of Nguyen & Ikeda (2015) were very close to 0 in their study, the effect sizes of 17 of 19 studies were in this area.

It is thought that choosing studies that have effect sizes were higher than the reference point, that the independent variable causes the same effect in terms of the dependent variable. It will also cause publication bias. In this research, the inclusion of studies that are very close to the reference point, and studies that e-portfolios have no effect on self-regulation and have an effect are the strengths of the study.

As considered in the heterogeneity decision part of the study, there wasn't only one effect size was in the studies. This decision was supported by the literature and estimation was made according to the random effects model. As a result of the analysis, the overall effect of e-portfolios used in the education process on self-regulation is 0,925. When this value is compared with the effect size criterion values (Wide effect = 0,75 ≤ Effect size ≤1,10) suggested by Thalheimer & Cook (2002), it is seen that it has a very large effect.

The overall effect value shows that the independent variable in the study has a significant effect on the dependent variable. And this effect size indicates that the e-portfolio application has a very large effect (Hedge's g = 0.925) on the self-regulation variable, which is determined as the dependent variable.

Discussion and Conclusion

There are many studies in which e-portfolio, one of the complementary measurement and evaluation techniques, is discussed. Among these studies, the focus was on studies conducted to examine the effectiveness of e-portfolios on self-regulation skills. Some of these studies reported that e-portfolio did not affect self-regulation, while others reported that it had positive effects. In line with the results of these studies carried out for the same purpose, the quantitative ratio of positive-negative or ineffective results will not give an accurate result. Therefore, Meta-analysis was used to reconsider the statistical results obtained from studies conducted under the same purpose and to interpret the overall effect. When examined with the criteria given in the method section for the research, 19 studies were found between the years 2000-2023 by these criteria. The meta-analysis conducted with 19 studies resulted in an overall effect size of +0.925. When this overall effect value is compared with the effect size scale, it is seen that it has a very large effect.

The fact that the e-portfolio application has been found to be significantly effective, like many variables that may affect self-regulation skills, is in line with the result of the study conducted by Railean (2008). Railean (2008) explains this situation provided and activities suggested to help the learning develop meta-cognitive abilities. These are awareness and regulation of cognition (which includes planning, monitoring, and self-evaluation of learning). Romero et al. (2019) remarked that, especially in higher education, individuals adapt to online tasks regarding 21st-century skills and being active in self-assessment and evaluation will increase the effect on self-regulation skills at a positive level.

Van der Gulden et al. (2020) examined the effects of e-portfolios on the components of the concept of self-regulation in their studies. In line with the results of their research, the researchers positively affected the concepts of self-assessment, reflection, feedback, goal setting, planning, and monitoring. levels were found to be affected. Each of these concepts represents the basic characteristics of self-regulation skills. The result obtained is consistent with the study of Van der Gulden et al. (2020).

Unfortunately, researchers only want to report their publications if the treatment, stated in studies based on a strategy, method or a technique trial, does not make a difference in post-tests. This study tested publication bias with very few studies stating that e-portfolios were ineffective on self-regulation skills. The study can be renewed with a larger sample number by changing the criteria and adding the remaining studies to the drawer. Especially in this age where education and training activities in the online environment gain meaning as one of the 21st century skills, the positive or negative effects of new measurement and evaluation techniques on the education process can be a new research topic.

Declarations

Ethical Approval: Secondary data were used in this study. Therefore, ethical approval is not required.

Author Contribution: Since there is only one author in this publication, the single author has 100% contribution in conceptualization, research, methodology, data curation, supervision, writing – review, editing, original draft, formal analysis and visualization.

Funding: The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Consent to Participate: All authors have given their consent to participate in submitting this manuscript to this journal.

Consent to Publish: Written consent was sought from each author to publish the manuscript.

References

Abrami, P. C., Venkatesh, V., Meyer, E. J., & Wade, C. A. (2013). Using electronic portfolios to foster literacy and self-regulated learning skills in elementary students. *Journal of Educational Psychology, 105*(4), 1188. https://doi.org/10.1037/a0032448

Akgün, E., & Kölemen, C. Ş. (2020). E-portfolyo oluşturmanin öğrencinin akademik öz düzenleme becerisine etkisi. *Milli Eğitim Dergisi, 49*(227), 117-140.

Alexiou, A., & Paraskeva, F. (2013). *Exploiting motivation and self-efficacy through the implementation of self-regulated oriented ePortfolio.* In International Conference on E-Learning in the Workplace, NY, USA.

Alexiou, A., & Paraskeva, F. (2014). *Implementing a Self-Regulated Oriented ePortfolio: The Design of an Affective Goal-Setting Plugin*. In 2014 IEEE 14th International Conference on Advanced Learning Technologies (pp. 474-476). IEEE. http://doi.org/ 10.1109/ICALT.2014.140

Alexiou, A., & Paraskeva, F. (2019). Examining self-regulated learning through a social networking ePortfolio in higher education. *International Journal of Learning Technology, 14*(2), 162-192. https://doi.org/10.1504/IJLT.2019.101849

Alhitty, A., & Shatnawi, S. (2021). *Using E-portfolios for Writing to Promote Students' Self-Regulation*. In Proceedings of the AUBH E-Learning Conference.

Barış, M. F. ve Tosun, N. (2013). Sosyal ağ ve e-portfolyo entegrasyonu: Facebook Örneği. *Eğitim ve Öğretim Araştırmaları Dergisi, 2*(2), 122-129.

Barrett, H. (2006). Using electronic portfolios for classroom assessment. *Connected Newsletter, 14(*2), 4-6.

Bartlett, A., & Sherry, A. (2006). Tow views of electronic portfolios in teacher education: Non-tech- nology undergraduates and technology graduate students. *International Journal of Instructional Media, 33*(3), 245–253.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research synthesis methods*, *1*(2), 97-111. https://doi.org/10.1002/jrsm.12

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2021). *Introduction to meta-analysis*. John Wiley & Sons.

Campbell, M. I., & Schmidt, K. J. (2005). Polaris: An undergraduate online portfolio system that encourages personal reflection and career planning. *International Journal of Engineering Education*, *21*(5), 931-942.

Cheng, G., & Chau, J. (2013). Exploring the relationship between students' self-regulated learning ability and their ePortfolio achievement. *The Internet and Higher Education*, *17*, 9-15. https://doi.org/10.1016/j.iheduc.2012.09.005

Chang, C. C., Liang, C., Chou, P. N., & Liao, Y. M. (2018). Using e-portfolio for learning goal setting to facilitate self-regulated learning of high school students. *Behaviour & Information Technology, 37*(12), 1237-1251. https://doi.org/10.1080/0144929X.2018.1496275

Chou, C. M. (2012). Influence of teachers' perceived e-portfolio acceptance on teacher evaluation effectiveness in Taiwan. *Australasian Journal of Educational Technology, 28*(4). https://doi.org/10.14742/ajet.1382

Cooney, F. E. (2007). *Adolescent self-regulation skills, working portfolios, and explicit instruction: A mixed methods study* (Doctoral dissertation, Walden University).

Corta, D. (2019). *Use of an e-portfolio to increase academic self-regulation in an undergraduate academic research course*. (Unpublished doctoral thesis, Private University in Peru).

Demirli, C. (2007). *Elektronik portfolyo öğretim sürecinin öğrenen tutumlarına ve öğrenme algılarına etkisi*. (Yayımlanmamış doktora tezi, Fırat Üniversitesi Sosyal Bilimler Enstitüsü, Elazığ).

Demirli, C., & Gurol, M. (2010, March). Teacher candidates' attitudes and perceptions towards electronic portfolio process. *In Society for Information Technology & Teacher Education International Conference (pp. 1412-1417).* Association for the Advancement of Computing in Education (AACE).

Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review, 28*(3), 425-474. http://doi.org/10.1007/s10648-015-9320-8

Ekbatani, G., & Pierson, H. (2000). *Learner-directed assessment in ESL.*Mahwah, NJ: Erlbaum.

Eom, W. (1999). *The effects of self-regulated learning strategy on academic achievement in a computer networked hypertext/hypermedia learning environment* (Unpublished doctoral dissertation). Florida State University, Florida.

Eom, W., & Reiser, R. A. (2000). The effects of self-regulation and instructional control on performance and motivation in computer-based instruction. *International Journal of Instructional Media, 27*(3), 247.

Field, A. P., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology, 63*(3), 665-694. https://doi.org/10.1348/000711010X502733

Jenson, J. D., & Treuer, P. (2014). Defining the E-Portfolio: What it is and why it matters. *Change: The Magazine of Higher Learning*, *46*(2), 50-57. https://doi.org/10.1080/00091383.2014.897192

Ghosh, R. (2003). Development and evaluation of a portfolio assessment package for year fve local studies [Unpublished doctoral thesis]. University of Malaya, Kuala Lumpur, Malaysia. http://studentsrepo. um.edu.my/3050/

Glass, G. V. (1976). Primary, secondary and meta-analysis of research. *Educational Researcher, 5*, 3–8.

Gozuyesil, E., & Tanriseven, I. (2017). A meta-analysis of the effectiveness of alternative assessment techniques. *Eurasian Journal of Educational Research, 17*(70), 37-56. http://dx.doi.org/10.14689/ejer.2017.70.3

Gülbahar, Y. ve Köse, F. (2006). Öğretmen adaylarının değerlendirme için elektronik portfolyo kullanımına ilişkin görüşleri. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi, 39*(2), 75-93.

Hanji, M. B. (2017*). Meta-analysis in psychiatry research: Fundamental and advanced methods.* CRC Press.

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2022). *Doing meta-analysis with R: A hands-on guide*. CRC Press.

Hedges, L. V. (1982). Estimation of effect size from a series of independent experiments. *Psychological Bulletin, 92*(2), 490–499. https://doi.org/10.1037/0033-2909.92.2.490.

Hedges, L. V. (1992). Meta-analysis. Journal of Educational Statistics, 17(4), 279-296.

Hedges, L. V., & Vevea, J. L. (1998). Fixed-and random-effects models in meta-analysis. *Psychological methods*, *3*(4), 486.

Karami, S., Sadighi, F., Bagheri, M. S., & Riasati, M. J. (2019). The Impact of Application of Electronic Portfolio on Undergraduate English Majors' Writing Proficiency and Their Self-Regulated Learning. *International Journal of Instruction, 12*(1), 1319-1334.

Kinash, S., Wood, K., & McLean, M. (2012). The whys and why nots of e-portfolios. Learning and teaching papers. http://epublications.bond. edu.au/tls/44.

Koç, A. (2010). *Bilgisayar okuryazarlığı becerisi ediniminde e-portfolyo sürecinin öğrenen performansına ve tutumlarına etkisi.* (Yayımlanmamış doktora tezi, Fırat Üniversitesi Sosyal Bilimler Enstitüsü, Elazığ).

Liang, C., Chang, C. C., Shu, K. M., Tseng, J. S., & Lin, C. Y. (2016). Online reflective writing mechanisms and its effects on self-regulated learning: A case of web-based portfolio assessment system. *Interactive Learning Environments, 24*(7), 1647-1664. https://doi.org/10.1080/10494820.2015.1041403

Lorenzo, G., & Ittelson, J. (2005). An overview of e-portfolios. Educause Learning Initiative, Paper 1:2005. http://electronicportfolio.pbworks. com/f/reading04overview.pdf.

Lysenko, L., Wade, C. A., Abrami, P. C., Iminza, R., & Kiforo, E. (2022). Self-Regulated Learning in Kenyan Classrooms: A Test of ePEARL, a Process e-Portfolio. *International Journal of Instruction, 15(*3), 63-82.

Meyer, E., Abrami, P. C., Wade, C. A., Aslan, O., & Deault, L. (2010). Improving literacy and metacognition with electronic portfolios: Teaching and learning with ePEARL. *Computers & Education, 55*(1), 84-9. https://doi.org/10.1016/j.compedu.2009.12.005

Mutluer, C. (2022). Tamamlayıcı Ölçme ve Değerlendirme Tekniklerinin Akademik Başarıya Etkisi: Bir Meta-Analiz Çalışması. *Ankara Üniversitesi Sosyal Bilimler Dergisi*, *13*(1), 15-27. http://doi.org/ 10.33537/sobild.2022.13.1.2

Nguyen, L. T., & Ikeda, M. (2015). The effects of ePortfolio-based learning model on student self-regulated learning. *Active Learning in Higher Education*, *16*(3), 197-209. https://doi.org/10.1177/1469787415589532

Pintrich, P. R., & DeGroot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*, 33–40. http://doi.org/10.1037/0022-0663.82.1.33.

Polat, M. ve Köse, Y. (2013). Okullarda bir performans değerlendirme aracı olarak e-portfolyo kullanımına yönelik ilköğretim öğretmenlerinin görüşleri. *Bilgisayar ve Eğitim Araştırmaları Dergisi, 1*(1), 57-82.

Railean, E. (2008). Electronic textbooks in electronic portfolio: a new approach for the self regulated learning. *In Proceedings of 9th International Conference on Development and Application Systems* (pp. 138-141).

Romero, L., Saucedo, C., Caliusco, M. L., & Gutiérrez, M. (2019). Supporting self-regulated learning and personalization using ePortfolios: a semantic approach based on learning paths. *International Journal of Educational Technology in Higher Education*, *16*, 1-16. http://doi.org/ 10.1186/s41239-019-0146-1

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005). Publication bias in meta-analysis. *Publication bias in meta-analysis: Prevention, assessment and adjustments*, 1-7. http://doi.org/ 10.1002/0470870168

Sasai, L. (2017). Self-regulated learning and the use of online portfolios: A social cognitive perspective. *Journal of Educational and Social Research, 7*(2), 55.

Srinivasjois, R. (2021). Fixed and random-effects models for meta-analysis. In *Principles and practice of systematic reviews and meta-analysis* (pp. 73-78). Cham: Springer International Publishing.

Sterne, J. A., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., Higgins, J. P. T. (2011). Recommendations for examining and ınterpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ, 343*, 1-8. https://doi.org/10.1136/bmj.d4002

Şen, S. & Yıldırım, İ. (2020). *CMA ile meta-analiz uygulamaları.* Anı.

Thalheimer, W., & Cook, S. (2002). How To Calculate Effect Sizes from Published Research: A Simplified Methodology. *Work-Learning Research, 1*, 1-9.

Trainin, G., & Swanson, H. L. (2005). Cognition, metacognition, and achievement of college students with learning disabilities. *Learning Disability Quarterly*, *28*(4), 261-272.

Tseng, K. H., & Lin, C. Y. (2015). Does using e-portfolios for reflective writing enhance high school students' self-regulated learning?. Technology, Pedagogy and Education, 25(3), 317-336. https://doi.org/10.1080/1475939X.2015.1042907

Türkkaynağı, U. (2021). *Investigating the use of speaking E-Portfolios in an EFL context and its effects on students' perceived speaking self-efficacy* (Master's thesis, Pamukkale Üniversitesi Eğitim Bilimleri Enstitüsü).

Üredi, İ., & Üredi, L. (2005). İlköğretim 8. sınıf öğrencilerinin öz-düzenleme stratejileri ve motivasyonel inançlarının matematik başarısını yordama gücü. *Mersin Üniversitesi Eğitim Fakültesi Dergisi, 1*(2), 250-260.

Van Der Gulden, R., Heeneman, S., Kramer, A. W. M., Laan, R. F. J. M., Scherpbier-de Haan, N. D., & Thoonen, B. P. A. (2020). How is self-regulated learning documented in e-portfolios of trainees? A content analysis. *BMC medical education*, *20*(1), 1-13. http://doi.org /10.1186/s12909-020-02114-4

Van Wesel, M., & Prop, A. (2009). Comparing students' perceptions of paper-based and electronic portfolios. *Canadian Journal of Learning and Technology/La revue canadienne de l'apprentissage et de la technologie*, *34*(3).

Wanchid, R., & Charoensuk, V. (2015). The Effects of Paper-Based Portfolios and Weblog-Based Electronic Portfolios on Limited English Proficiency Students in Writing for Service Industry Course. *English Language Teaching*, *8*(9), 131-145. http://doi.org/ :10.5539/elt.v8n9p131

Zimmerman, B. J. (1986). Becoming a self-regulated learner: which are the key subprocesses? *Contemporary Educational Psychology, 11*, 307–313.

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology, 81*, 329–33.

ISSN: 1309 – 6575
Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi
Journal of Measurement and Evaluation in Education and Psychology
Research Article; 2023; 14(Special Issue);299-319

The Role of ePortfolios in Formative Assessment: A Systematic Literature Review

Ezgi GÜN-TOSİK* Bilal ATASOY** Akça Okan YÜKSEL***

Abstract

In educational settings, formative assessment is used to determine the status of students and foster their development. Task, process, and feedback serve as the basis of formative assessments. Typically, tasks are introduced to lessons to facilitate student feedback. Due to their technological potential, ePortfolios are often used to help with assigning tasks, managing processes, and getting feedback. It is necessary to incorporate the findings from environments where ePortfolios are used for formative assessment into the literature. A systematic literature review was employed for this purpose. Three researchers reviewed 33 studies utilizing content analysis. The Fleiss kappa was used to determine inter-rater reliability. Quantitative analysis was performed on the articles' dependent or related variables, environments, research method, implementation period, evaluators, ePortfolio type, activity format, evaluation instruments, education level, and course categories. Among the qualitative findings of the studies were the prominent concepts of the use of ePortfolios, as well as the themes and codes of advantages and challenges. The findings indicate that the integration of ePortfolios into the formative assessment process, despite its challenges, provides significant advantages. It is expected that the study's findings will be useful for researchers as well as practitioners who intend to use ePortfolios for formative assessment processes.

Keywords: ePortfolio, formative assessment, feedback, task, systematic literature review

Introduction

Assessment is a crucial component of education. Sometimes summative assessment is used to determine a student's grade, while sometimes formative assessment is used to support learning by disclosing the student's status in the learning process (Wiliam & Black, 1996; Bennet 2011; Van der Kleij, Vermeulen, Schildkamp, & Eggen, 2015; Box, 2019). Formative assessment can enlighten the learner, the teacher, and other educational stakeholders about the learner's developmental progress (Black & Wiliam, 1998; Chappuis & Stiggins, 2002; Black & Wiliam, 2010; Bennett, 2011; Fuller & Dawson, 2017). With the acquired information, gaps can be identified as the process continues (Dunn & Mulvenon, 2009) and student-specific feedback and guidance can be provided (Bennett, 2011). Rather than evaluating the student or the learning outcome, the purpose here is to provide information about the student's circumstance and to assist in improving the quality of the learning process (Stobart, 2008). However, formative assessment requires more time and effort from teachers and students to implement (Black & Wiliam, 2010). For instance, assigning tasks to students, determining the assessment tools/criteria for these tasks, providing information on the use of these tools, establishing a feedback approach (self-peer-tutor-mentor, etc.), and providing feedback on the tasks/products at specific intervals can result in a substantial amount of work.

* Res. Assist. Dr., Ondokuz Mayıs University, Faculty of Education, Samsun-Turkiye, ezgi.tosik@gmail.com, ORCID ID: 0000-0001-7747-1917
** Assoc. Prof. Dr., Gazi University, Faculty of Education, Ankara-Turkiye, atasoybilal@gmail.com, ORCID ID: 0000-0001-6894-0646
*** Corresponding Author, Dr. Lecturer, Middle East Technical University, Presidency Office, Ankara-Turkiye, akca@metu.edu.tr, ORCID ID: 0000-0002-5430-0821

To cite this article:
Gün-Tosik, E., Atasoy, B., & Yüksel, A.O. (2023). The role of eportfolios in formative assessment: A systematic literature review, *Journal of Measurement and Evaluation in Education and Psychology, 14*(Special Issue), 299-319. https://doi.org/10.21031/epod.1276063

Received: 3.04.2023
Accepted: 28.08.2023

Technological advancements can contribute to this process and provide alternate solutions to the challenges encountered in both the administration of the formative assessment process and its constituent parts (such as sharing, feedback, etc.) (Irving, 2015). The ePortfolio approach, which enables both summative and formative assessment, is regarded as a key technology and paradigm for facilitating this challenging endeavor (Barrett, 2010; Vogel, 2018; Lam, 2022). ePortfolio systems facilitate the process of observing, monitoring, sharing, reflecting, and providing feedback on the learning processes and experiences of students (Barrett, 2006; Chang, Chou, & Liang, 2018). ePortfolio platforms also facilitate the creation, storage, and accessibility of formative assessment resources. Supporting the process of sharing these resources and providing feedback on development is another contribution (Fathi & Rahimi, 2022). ePortfolio systems also provide an opportunity to share and receive feedback with the creator of the artifacts for as long as desired (Bennett, 2011).

Despite numerous studies in the literature on the educational effects of the ePortfolio frequently used in formative assessment processes (Barrett, 2006; Nicolaidou, 2013; Ebil, Salleh, & Shahrill, 2020; Beckers, Dolmans, & Van Merrinboer, 2022), ePortfolio use in formative assessment processes has only been discovered in a systematic study (Burner, 2014). The focus of the related research was limited to foreign language writing skills. In light of this, there is a need for literature review works that will broaden the scope and obtain recent studies. Furthermore, it is hypothesized that this research, which will additionally examine the outcomes of experimental investigations on the topic, would provide valuable insights for scholars and professionals in relation to the utilization of portfolios within the formative assessment procedure.

Aim of the Study

This study examines quantitative analyses that employ formative assessment in an ePortfolio environment in terms of dependent or related variables, environments, research methodology, implementation period, evaluators, ePortfolio type, activity type, evaluation tools, education level, and courses. Furthermore, based on the results derived from the qualitative study of the ePortfolio applications utilized during the formative assessment procedure, the prominent concepts and challenges encountered were presented in the existing literature.

Background

Formative Assessment

According to a frequently cited definition, formative assessment is the organization of learning and teaching activities with feedback from teachers and/or students (Black & Wiliam, 2010; Asamoah, Shahrill, & Abdul Latif, 2022). Similarly, according to Ramsey and Duffy (2016), formative assessment is a process that supports instruction by providing both educators and learners with continuous and real-time feedback. Pachler, Daly, Mor, and Mellar (2010) assert that formative assessment is distinguished from summative assessment by the presence of feedback. On the positive effects of feedback on learning, there is an extensive body of research, despite the existence of a few controversial studies with contradictory findings (Wiliam, 2011; Lui & Andrade, 2022). The objective of the feedback should be to reduce the distance between the current and desired state of the learners (Andrade, 2010). All stakeholders in education, including teachers, students, parents, and mentors can provide feedback. Some information regarding the learning process is required for effective feedback and guidance (Schildkamp, Van der Kleij, F, Heitink, Kippers, & Veldkamp, 2020). According to Klenowski (2009), those required information can be gathered via assignments, tests, mini-examinations, dialogues, observations, and discussions. In addition to these, data obtained by individuals with tools such as structured criteria lists, guides, and rubrics (Kutlu, Doğan, & Karakaya, 2017) and data obtained automatically by computer systems (Karaoğlan Yılmaz, Yılmaz, & Öztürk, 2020) can be used to provide constructive feedback. Formative assessment differs from summative assessment in that it focuses on the learning process and the quality of this process rather than on the results (Stobart, 2008; Gezer, Wang, Polly, Martin, Pugalee, & Lambert, 2021). The objective is to ensure that all stakeholders are aware of the process, rather than to evaluate it, and to provide educational environments with data so that enhancements can be made as needed.

ePortfolio

Portfolios are physical files in which the learning process is recorded systematically, including what students learn, their thinking, questioning, analyzing, and producing abilities, as well as their interactions with teachers and peers (Gibson, 2006; Stefani, Mason, & Pegler, 2007; Yancey, 2023). ePortfolios are electronic collections of every kind of data associated with the learning process (Barrett, 2010; Törmala, 2021). There are challenges with storing, accessing, and updating physical portfolios (Heath, 2005). In addition to overcoming these issues, ePortfolios provide additional advantages, such as the ability for instructors and peers to provide immediate feedback and the offering of rich multimedia content (Barrett, 2006). In addition, ePortfolios have been reported to improve research skills (Demir & Kutlu, 2016), writing performance (Nicolaidou, 2013), vocabulary learning (Sharifi, Soleimani, & Jafarigohar 2017), self-directed learning skills (Beckers et. al., 2022), and engagement and self efficacy (Mason, Pegler, & Weller, 2004; López-Crespo, Blanco-Gandía, Valdivia-Salas, Fidalgo, & Sánchez-Pérez, 2021).

The literature classifies ePortfolios in a variety of ways. Himpsl and Baumgartner (2009) identified three kinds of portfolios: reflection, development, and presentation, whereas the Global Learning Consortium (2005) listed six types: assessment, presentation, learning, personal development, multiple-owner, and working. While classified in different categories, ePortfolios frequently emphasize process, reflection, assessment, and development.

In the beginning, ePortfolios were stored on CDs and DVDs, but with the widespread use of the Internet and the development of Web 2.0 tools, they migrated to online environments. In addition, by adding ePortfolio components to learning management systems (Demir & Kutlu, 2016; Beckers et. al., 2022), existing blogs (Huang & Hung, 2010; Chang, Liang Tseng, & Tseng, 2014), Facebook (Kabilan, 2016), other social media platforms (Shepherd & Bolliger, 2011; López-Crespo et. al., 2021), or ePortfolio-specific platforms (Chuang, 2010; Garrett, 2011; Yang, Tai, & Lim , 2016; López-Crespo et. al., 2021) ePortfolio has been developed and its use has become widespread. Platforms such as Mahara and Elgg, which support all the expected ePortfolio features, are frequently used for this purpose (Balaban & Bubas, 2010; Mgarbi, Chkouri, & Tahiri, 2022). The use of various forms of ePortfolios may be favored depending on the educational or instructional purpose. This situation may also result in the selection of various assessment instruments. In the literature, it is frequently found that assessment types such as rubrics (Barbera, 2009; Chau & Cheng, 2010; Nicolaidou, 2013), checklists (Sánchez Gómez, Ostos, Solano, & Salado, 2013), questionnaires (Hung, 2012; Beckers et. al., 2022), content analysis (Huang & Hung 2010; Kabilan, 2016), and observation forms are frequently used in ePortfolios (Kutlu et. al., 2017). ePortfolios provide academic support for both assessment and the establishment of effective learning environments (Barrett, 2006; Chang & Kabilan, 2022). This support is enabled not only by a technological environment but also by its structure, which relies on the student-centered, active participation, and constructivist learning theories, encourages sharing, and facilitates the process of collaboration (Gülbahar, 2009).

ePortfolio for Formative Assessment

In the portfolio approach, not only the outcome or product but also the development and process that contribute to this improvement are significant (Barker, 2006; Kerr, 2007; Bennett, Knight, & Rowley, 2020; Beckers et. al., 2022). Similarly, formative assessment does not emphasize product or evaluation but rather focuses on process awareness and development (Nitko & Brookhart, 2014; Morris, Perry, & Wardle, 2021; Mashauri, 2023;). ePortfolio broadens the perspective of a particular final product, thereby contributing to the development of process-quality knowledge (Barrett, 2003). This is an essential component of both the educational setting and the formative assessment procedure. The ePortfolio approach enables students to make connections between "what they have learned" and "how they have learned" while creating artifacts and recognizing their development (Hallam & Creagh, 2010; Blaschke & Marin, 2020).

Learners can access ePortfolios from anywhere, at any time, using a variety of devices, including computers, tablets, and mobile phones. In addition, the inclusion of social media components in ePortfolio platforms makes them familiar to users and facilitates their use (Oh, Chan, & Kim, 2020). Students can use these platforms, which require few technical skills, to share their portfolios with whomever they want and receive as much feedback as they require (Heinrich, Bhattacharya, & Rayudu, 2007; Hegarty & Thompson, 2019). They can access the information in their shared ePortfolios whenever and with any devices they prefer. The core components of the formative assessment procedure are the mentioned sharing and feedback features (Black & William, 1998; Leighton, 2019). This process can be disrupted in non-digital or classroom settings due to factors that involve lack of time and access (Gamlem & Kari Smith, 2013). The ePortfolio approach can contribute to the perpetuation of the formative assessment process by offering numerous solutions to these challenges.

Method

In this study, a systematic literature review method was used. This method was chosen to acquire a structured and comprehensive synthesis of studies conducted on a particular topic according to specific criteria (Kitchenham, 2004). Systematic research is necessary because it enables researchers to disclose the similarity and diversity of studies on a topic and reveals the general trend to researchers who study or wish to study in the area of interest (Cohen, Manion, & Morrison, 2007).

Search Strategies and Sampling

The purpose of this study is to carry out a literature review on the use of ePortfolio environments for formative assessment. For this purpose, the following search was performed in the Web of Science database. This search was performed last on 15.10.2022.

“Electronic portfolio (Title) or e portfolio (Title) or e-portfolio (Title) and "formative assessment" (Topic) and Article (Document Types) and Education Educational Research or Education Scientific Disciplines (Web of Science Categories) and English (Languages)”

A total of 272 studies were identified at the end of the query. As there is so much research on the topic, it needs to set some criteria for the selection of publications (Heitink, Van der Kleij,Veldkamp, Schildkamp, & Kippers, 2016). For this reason, three researchers analyzed 17 articles via online sessions to create a consistent framework for reviewing the studies, and they defined the inclusion and exclusion criteria in Table 1. It was decided to exclude 187 studies whose full texts were inaccessible, not indexed in SSCI, reviews, model and system development, book chapters and proceedings. 85 studies remained to be reviewed after the studies that were excluded. The researchers reconsidered 10 papers collectively during a second online discussion on the themes to be used in the data analysis. In these reviews, the researchers decided to exclude 52 studies that did not use ePortfolios for formative assessment or did not provide information on how they used them. 33 studies were ready to be analyzed after this exclusion criteria was followed.

Table 1

Inclusion and exclusion criteria

Inclusion Criteria	Exclusion Criteria
Using ePortfolio for formative assessment process	Book chapters, reviews
Full text was accessible	model/system development and integration studies, proceedings
Indexed in SSCI	not provide information/details on how tasks and feedback are carried out in the ePortfolio assisted formative assessment process

Analysis, Validity, and Reliability

The data underwent content analysis. Content analysis is a systematic method that transforms texts into categories based on some criteria (Stemler, 2000). Before conducting a content analysis, three researchers analyzed ten studies in collaboration to identify potential themes.To ensure consistency between the codings, they then collaboratively reviewed the studies according to these themes. Next, the researchers independently reviewed nine studies chosen at random. To determine the agreement between researchers, the Fleiss kappa coefficient was calculated as 0.93. This indicates a very high level of agreement among the experts according to Landis and Koch (1977). Contradictory instances encountered by the researchers during the analyses were discussed, and the themes were finalized collaboratively during online sessions. All procedures were clarified in detail to ensure the external validity (Guba & Lincoln, 1982).

Findings

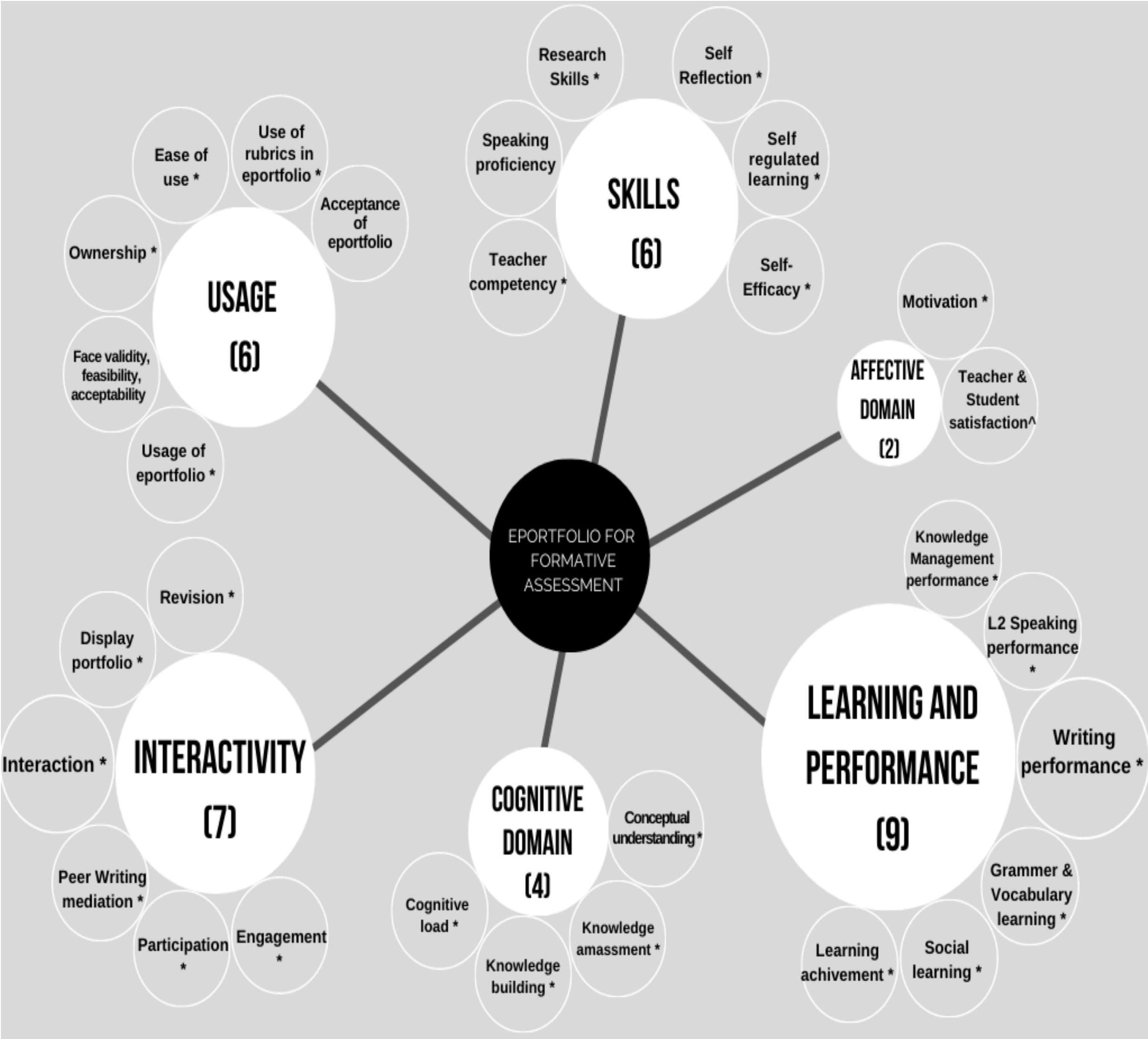
The included studies were analyzed according to the following topics. These are dependent or related variables, environments, research method, implementation period, evaluators, ePortfolio type, activity format, evaluation instruments, education level and courses, prominent concepts, advantages, and challenges.

Dependent or Related Variables

The themes and codes that emerged from the dependent or related variables obtained from the quantitative analyses were created by taking into account only the studies that collected data with valid and reliable measurement instruments, and statistical analyses were conducted on these data. The themes are skills, learning and performance, cognitive domain, interactivity, usage and affective domain. It is observed that the codes related to these themes represent a diverse spectrum, and that a remarkably high proportion of significant results have been obtained in the context of these analyzed variables (Figure 1). Themes indicate that more study has been conducted specifically on the topic of learning performance. Additionally, a similar amount of studies have been conducted on the topics of skills, interactivity, and usage. Writing performance (8.8%) and interaction (5.88%) are the most prominent categories.

Figure 1

Dependent or related variables



**studies with significant results, ^ increase but not significant*

Environments

The environments/platforms utilized in the studies were analyzed and classified into four categories: social media/network, developed by researcher or organization, learning management system, and others. The social media/network topped the list among the environment theme to the analyzed studies (f=16), approximately 48%. Ten categories are included in the social media / network as Blog, Mahara, e pass, Elgg, Edmodo, PebblePad, Facebook, Google Sites, Google Groups, QQ. Following the social media / network, the most frequent theme is developed by researchers or organizations. It corresponds to a total of 28.57% (f=8). Learning Management System and other themes are included in two studies. The environment could not be determined in five studies. In addition, in 15.2% of the studies (f=5), the development platforms were not specified.

Table 2

Environments

Theme	Environment/platform	f	%
Social Media / Network	Blog (Blogger, Wretch, wordpress based weblog	5	15.2
	Mahara	3	9.1
	e-pass	1	3.0
	Elgg	1	3.0
	Edmodo	1	3.0
	PebblePad	1	3.0
	Facebook	1	3.0
	Google Sites	1	3.0
	Google Groups	1	3.0
	QQ	1	3.0
Developed by researcher or organization	Private system	8	24.2
Learning Management System	BlackBoard	1	3.0
	Angel LMS	1	3.0
Other	PDA	1	3.0
	Microsoft Office OneNote Class Notebook	1	3.0
Unspecified	-	5	15.2
Total		33	100

Research Methods

When the research methods are examined, it is seen that mixed methods are the most common (f= 16, 48.5%). In addition, the studies also contain quantitative (f=9, 27.3%) and qualitative (f=8, 24.2%) methods.

Table 3

Research Methods

Method	f	%
Mixed	16	45.8

Table 3

Research Methods (Continued)

Quantitative	9	27.3
Qualitative	8	24.2
Total	33	100

Implementation Period

The implementation periods of the studies are classified under 3 themes (Table 4). Analyzing the article's implementation periods reveals that most implementations lasted one semester (f=20, 60.6%). Seven studies (21.2) were implemented for two semesters. The implementation periods of the 6 studies (18.2%) varied between 6 and 9 weeks.

Table 4

Implementation Period

Week/Semester	f	%
1 semester (14 weeks)	20	60.6
2 semesters	7	21.2
6-9 weeks	6	18.2
Total	33	100

Evaluators

The majority of studies used a combination of self, peer, and teacher evaluation (45.5%) (Table 5). There were six studies (18.2%) in which both the student and the teacher became co-evaluators. The teacher was the single evaluator in three (9.1%), and the peer was the single evaluator in two studies (6.1%). Also, peers and teachers were in the evaluator role together in three studies. There is only one study in which both the self and a peer served as the evaluator, and there is only one study that utilizes systems as an evaluator.

Table 5

Evaluators

Theme	f	%
Self-peer-teacher	15	45.5
Self-teacher	6	18.2
Teacher	3	9.1

Table 5

Evaluators (Continued)

Theme	f	%
Peer-teacher	3	9.1
Peer	2	6.1
Self-peer-teacher-mentor	1	3.0
Self-teacher-system	1	3.0
Self-peer-teacher-parent	1	3.0
Self-peer	1	3.0
Self	-	-
Total	33	100

Also, only three studies examined the relationship between evaluators. Two of them analyzed the relationship between self-teacher, and one analyzed the relationship between self-peer-teacher.

ePortfolio Type

The distribution of ePortfolio types in studies is represented in Table 6. The most frequent type of portfolio is development (process) with 16 studies (48.5%). The assessment portfolio is the following type of portfolio that is also quite widespread (f=8, 24.2%). In addition, both types of development and assessment portfolios were used in 5 studies (15.2%). Even though they are rare, some studies also used the showcase (f=2, 6.1%) and reflective (f=2, 6.1%) types of ePortfolio.

Table 6

ePortfolio type

Type	f	%
Development (Process)	16	48.5
Assessment	8	24.2
Development + assessment	5	15.2
Showcase	2	6.1
Reflective	2	6.1
Total	33	100

Activity Format

The majority of studies (f=13, 39.4%) employ a task and reflection combination, the most common activity type (Table 7). The studies that use tasks (f=9, 27.3%) are in second rank. Reflection was used as the activity format in four studies (12.1%). In addition, there are two studies that applied [task+discussion] and [task+project+discussion+reflection+presentation] activity types. Also, [Task+reflection+discussion], [Task+group project+reflection] and [Task+group project+reflection+presentation] activity types were each used in a seperate study.

Table 7

Activity Format

Format	f	%
Task + reflection	13	39.4
Task	9	27.3
Reflection	4	12.1
Task + discussion	2	6.1
Task + project + discussion + reflection + presentation	2	6.1
Task + reflection + discussion	1	3.0
Task + group project + reflection	1	3.0
Task + group project + reflection + presentation	1	3.0
Total	33	100

Assessment Tools

It is understood that mainly (f=12, 32.4%) rubrics are used as assessment tools (Table 8). Quizzes are the second most frequently employed assessment tool (f=4, 10.8%). In a small number of studies, survey (f=2, 5.4%), criteria list (f=2, 5.4%), rubric and survey (f=2, 5.4%), and content analysis (f=2, 5.4%) were preferred. It is surprising that 29.7% (f=11) of studies did not explain how electronic portfolios were evaluated. Standardized language testing and system-assessed tests were used in only one of them (f=1, 2.7%).

Table 8

Assessment tools

Tools	f	%
Rubric	12	32.4
Quiz	4	10.8
Criteria list	2	5.4
Content analysis	2	5.4

Table 8

Assessment tools (Continued)

Tools	f	%
Survey	2	5.4
Rubric + Survey	2	5.4
Standardized language testing	1	2.7
Student reactions (System)	1	2.7
Unspecified	11	29.7
Total	37	100

Education Levels and Courses

Analysis reveals that ePortfolios are used for formative assessment at various educational levels and in a variety of course settings, including undergraduate, K-12, and graduate programs (Table 9). Undergraduate is the most preferred level of education (f =21, 58.3%). The number of studies conducted at K-12 is 8 (22.2%). In seven studies (19.4%), the education level is graduate or postgraduate. Despite the fact that the majority of studies are conducted in language education (f=13, 39.4%) and educational sciences (f=9, 27.3%), it is evident that they are conducted in numerous other fields, such as information and communication technology (ICT) and medical.

Table 9

Education levels and Courses

Education Levels	Courses	f	%
Undergraduate	Language Education	8	22.2
	Educational Science	7	19.4
	Medical Education	2	5.6
	ICT	2	5.6
	Technical and Vocational Education	2	5.6
Sub-Total		21	58.3

Table 9

Education levels and Courses (Continued)

Education Levels	Courses	f	%
	Language Education	4	11.1
	Science	2	5.6
K12	Research Skills	1	2.8
	ICT	1	2.8
Sub-Total		8	22.2
	Medical Education	2	5.6
Graduate and Post- Graduate	Educational Science	2	5.6
	Analysis on Game Industry	1	2.8
	Language Education	1	2.8
	Science	1	2.8
Sub-Total		7	19.4
Total		36	100

Prominent Concepts in Qualitative Analyses

The concept of feedback (f = 9) is most prevalent in the qualitative results of the studies reviewed. At the end of the qualitative analyses, self-assessment (f=6) emerged as a prominent concept. The studies also emphasize that if ePortfolio environments are designed by considering real learning tasks, a community of practice and interaction, they can make positive contributions in components such as confidence, ownership, self-improvement, peer assessment, attitude, professional development, social learning, motivation, new experience, content knowledge/content enrichment, monitoring of progress, peer support, teacher assessment, and reflection.

Figure 2

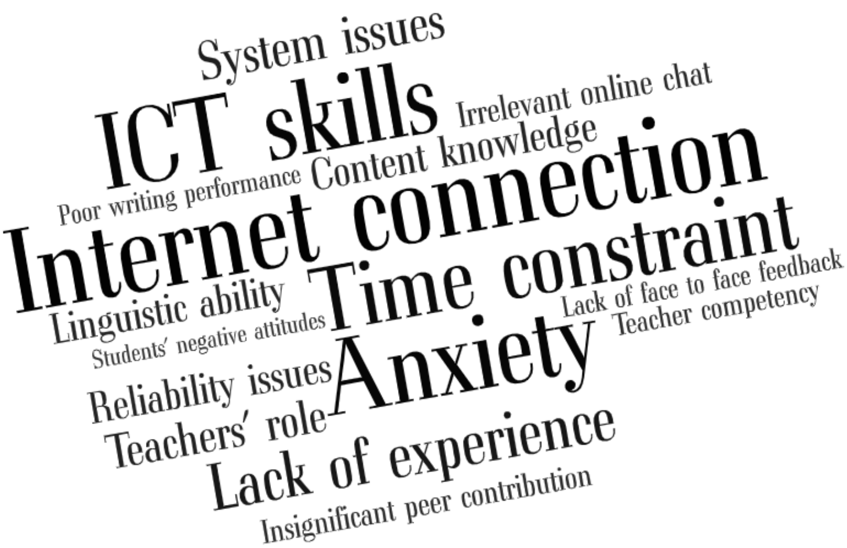
Prominent concepts

Challenges

In the qualitative findings of the studies, the challenges associated with the use of electronic portfolios in formative assessment processes were also examined. Internet connection was the most common difficulty encountered by the users (f=5), followed by anxiety caused by users following each other's work (f=4), the need to have ICT skills (f=4), time constraints (f=3), lack of experience, and system problems (f=2).

Figure 3

Challenges



Conclusions and Discussions

It is expected that studies addressing crucial educational variables, such as perception, attitude, motivation, and satisfaction, will contribute to the body of knowledge. In addition, there have been studies conducted on a variety of important variables in the field of education. This result indicates that ePortfolios are perceived as environments for managing formative assessment in the context of various variables. Remarkably, there has been a limited number of studies have been conducted in the affective domain. Considering the significance of the affective dimension in education (Vankúš, 2021; Pierre & Oughton, 2007), it is anticipated that further studies in the subject will fill this gap in the literature.

Since developing a new platform is costly, time-consuming, and labor-intensive, the consensus generally leans towards favoring existing systems for the majority of ePortfolio environments. However, there are only a few studies conducted using open-source and free platforms developed specifically for this purpose, such as Mahara and Elgg. This could be due to unfamiliarity with these platforms, the time-consuming process of setting up and modifying these environments, or the need for knowledge, skills, and time to maintain them. The positive results of the usability and satisfaction analysis of the studies (Garrett, 2011; McLeod & Vasinda, 2009) indicate the reasons behind the preference for these environments. Another possible explanation is that researchers choose their existing or familiar systems and are unwilling to learn new ones. Investigating the reasons for the limited utilization of these free environments explicitly designed for ePortfolios could contribute to their wider adoption.

Another result reveals that a majority of the analyzed studies employ mixed methods. In the social sciences, it is recommended to diversify analysis by combining quantitative and qualitative insights and opinions (Creswell, 2012). The results of analyzing different kinds of data can be used to enrich the literature.

The effective implementation of formative assessment through ePortfolio depends on users' ability to adapt to the process (Hallam & Creagh, 2010). Considering the related tasks, measurement and feedback systems, and the process of becoming familiar with a platform, it is obvious that this will take time. It is clear that the analyzed studies complied with the required period of time. Segaran and Hasim (2021), in their meta-analysis study on ePortfolio and self-regulated learning, found similar findings regarding the utmost duration of the implementation.

In six studies, both the student and the teacher took part as co-evaluators. This finding is significant in demonstrating that stakeholders are involved in the formative assessment and feedback mechanism. Since feedback and evaluation in ePortfolio environments require continuity, using alternative evaluators may be useful. There is only a single study that utilizes systems as an evaluator. This type of use may help to reduce the duties of teachers. It needs to be determined whether this is due to a lack of trust in student evaluations or other factors. Self and peer assessment and feedback may reduce a teacher's workload, especially in cases where the teacher may find it difficult to monitor the process, such as in massive online courses or large classes. Considering the developments in software, it is surprising to discover only a single study in which a computer system serves as the evaluator. The use of an automated system may reduce the workload of teachers. By integrating technologies like artificial intelligence, data mining, adaptive hypermedia, and intervention systems into educational platforms, evaluation and feedback performance can be enhanced. Although the positive effects of involving parents in educational settings are known (Hoover-Dempsey & Sandler, 1995), there is only one study in which parents were involved (Meyer, Abrami, Wade, Aslan, & Deault, 2010). When compared to the graduate and undergraduate levels, it makes sense for K–12 parents to participate in the evaluation process. In light of the fact that 22% of the reviewed studies were at the K–12 level, it would be useful to investigate the reasons why parents participated in only one study's evaluation process. Analyzing the relationships between evaluators in ePortfolios used for formative assessment can contribute to the literature on evaluator compliance and objectivity.

It is an expected finding that the developmental portfolio was frequently chosen as the ePortfolio type because, in studies using formative assessment, portfolio types that focus on the process are mostly preferred. Our research includes only those studies that implement the formative assessment in terms of both keywords and scope. The use of assessment ePortfolio is also quite widespread. Studies that focus on development aim to collect data about the process and provide appropriate feedback (Tillema, 2001) and guidance (Chetcuti, Buhagiar, & Cardona, 2011). Although not common, some studies also used the showcase type of ePortfolio. Despite the fact that showcase portfolios are product-oriented, process-oriented practices were also observed in these studies by selecting the best of the products developed during the process to create showcase portfolios.

Regarding the activity format, a small proportion of studies employ all activity categories including task, group and individual project, reflection, discussion, and presentation. This may be due to the fact that as the types of activities increase, so does the effort required by students to execute these activities and by teachers to manage them. Despite the fact that these numerous types of activities provide students with richer experiences and a multifaceted view of the studied context, they increase the workload for teachers and students. In addition, there are a lot of studies that use interactions such as group projects or group discussions. This preference is understandable, especially in educational settings designed in accordance with social, constructivist, and connectivist theories (Andrade et. al., 2023; Bryant & Bates, 2015; Zhang, Olfman, Ractham, & Firpo, 2009). ePortfolio environments support educational environments by providing discussion forums, interactive whiteboards, synchronous and asynchronous tools for such group activities. It is thought that the main purpose of selecting these activities is to trigger the feedback process, which is an important part of formative assessment. Asking for reflections on the tasks leads to two products that can be given feedback. One of these feedback is related to the students' products, while the other is related to their reflections. Both feedbacks can be used to support the development of students during the formative assessment process (Wade, Abrami, & Sclater, 2005).

According to examined research, the majority of assessment tools are rubrics which frequently used instruments to measure development based on graded criteria (Contreras-Higuera, Martínez-Olmo, Rubio-Hurtado, & Vilà-Baños, (2016). Rubrics are also preferred for providing and analyzing self-peer-teacher assessments, as well as examining their relationships (De Grez, Valcke, & Roozen, 2012; Barbera, 2009). The personal and structured nature of the rubrics can enhance the objectivity of the feedback. In a small number of studies, a criteria list and content analysis were chosen. It is surprising that some studies did not explain how electronic portfolios were evaluated. Another assessment strategy is to use tests at regular intervals to determine the status of students or to support their learning. In particular, the fact that multiple-choice tests can be evaluated easily by the system will increase the speed with which students receive feedback and reduce the teacher's workload.

Another predictable finding is that formative assessment through ePortfolio environments are applicable to all levels and branches of the educational system (Sweet,1993). It is understandable that alternative tools, such as portfolios, reflections, and so on, are increasingly used to measure higher-level outcomes, particularly at the graduate and postgraduate levels. Given that ePortfolios facilitate reflection in numerous ways (Slepcevic-Zach & Stock, 2018), their prevalence at this level of education is not surprising. In addition, there are numerous opinions in the literature regarding the use of alternative assessment and evaluation tools, such as ePortfolios, at the K-12 level (Gülbahar, 2009; Meyer, Abrami, Wade, & Scherzer, 2011; Mitchell, Campbell, Somerville, Cardell, & Williams, 2021). Surprisingly, only a few of the studies were conducted in grades K-12. The fact that the number of research at the undergraduate level is higher than that of the graduate level can be explained by the number of students at these levels, but considering the number of schools and students at the K-12, it is confusing that the number of studies at this level. According to Barrett (2010), many ePortfolio applications are conducted at the undergraduate level. It is common knowledge that university-based academicians conduct the vast majority of academic research. It is believed that the prevalence of undergraduate-level research is due to the researchers' use of a convenient sampling method. Although the main concentration of study at each educational level is in the fields of language education and educational sciences, there are also studies in some other disciplines as well. ePortfolios are used extensively for reflection and process monitoring in language education and educational sciences. The lack of research in the social sciences and career development, where ePortfolios are commonly utilized, is notable. Thus, the use of ePortfolios for formative assessment in these disciplines can contribute to the literature.

The notion of feedback is predominantly observed in the qualitative findings of the studies examined. For students' success in the learning process, quality support and mentoring are essential, and feedback is an integral component of this (Peacock, Scott, Murray, & Morss, 2012). Peer feedback, for example, has been reported to reduce the need for teacher support in online environments (Shepherd & Bolliger, 2011). In a similar vein, Gardner and Aleksejuniene (2008) assert that formative feedback reduces the workload of both teachers and students and increases the possibility of producing high-quality assignments and outputs. It is apparent that user opinions reflect the benefits of feedback used in ePortfolio environments.

As a result of the content analysis, it is revealed that other prominent concepts are self-assessment and self-reflection. Providing users with a chance to evaluate themselves and to follow and monitor their processes is shown as the benefit of ePortfolio environments that support these concepts (Ebil et. al., 2020). ePortfolio environments facilitate self-assessment of self-directed learning skills (Beckers et. al., 2022). In addition, these environments also provide users with the opportunity for self-reflection (Wang & Jeffrey, 2017). The studies also emphasize that if ePortfolio environments are designed by considering real learning tasks, community of practice, and interaction, they can make positive contributions in components such as confidence, ownership, self-improvement, peer assessment, attitude, professional development, social learning, motivation, new experience, content knowledge/content enrichment, monitoring of progress, peer support, teacher assessment, and reflection. In addition, it becomes apparent that ICT skills and anxiety levels must be taken into account for the useful administration of these environments.

Studies identified slow Internet connections, lack of Internet access, installation problems, and system maintenance as major obstacles (Fathi & Rahimi, 2022; Hung, 2012). It is remarkable that these problems persist today. To avoid these issues, it is recommended to carefully organize the tasks/authorisations, the construction of the technical infrastructure, and the system's sustainability before starting the ePortfolio applications. Basic ICT skills are one of the competencies that must be considered for a successful ePortfolio process. Hung (2012) states that although technology plays an important role in the portfolio development process, it can create frustration for some users.Wang and Jeffrey (2017) state that users with low ICT skills show negative attitudes towards ePortfolio. Therefore, it would be beneficial to provide support in the form of fundamental ICT skills in order to reduce user anxiety and facilitate their participation in these processes. According to Kabilan and Khan (2012), despite the fact that students find the use of ePortfolios beneficial, they experience time constraints as a result of their responsibilities. It demonstrates that in the weeks that follow, students adopt negative attitudes, such as rephrasing the same comment, repeating what others say, and paraphrasing. Consequently, workload and time management are essential factors to consider in these environments (Stefani et. al., 2007).

Despite some challenges, it is clear that the use of ePortfolios in formative assessment provides a very high rate of positive results. Therefore, it can be assumed that their use in educational environments will be beneficial. Adapting techniques such as AI-supported feedback and evaluation to these settings can enhance the ePortfolio's support for formative assessment by shortening the response time and reducing the teacher's workload.

Declarations

Conflict of Interest: No potential conflict of interest was reported by the author.

Ethical Approval: Secondary data were used in this study. Therefore, ethical approval is not required.

References

* indicates the study included in systematic literature review.

Andrade, H. L. (2010). Summing up and moving forward: key challenges and future directions for research and development in formative assessment. *In Andrade, H. L. & Cizek, G. J. (Eds), Handbook of formative assessment (344- 351)*. Routlage.

Andrade, R., van Riper, C. J., Goodson, D. J., Johnson, D. N., Stewart, W., López-Rodríguez, M. D., ... & Raymond, C. M. (2023). Values shift in response to social learning through deliberation about protected areas. *Global Environmental Change, 78*, 102630. https://doi.org/10.1016/j.gloenvcha.2022.102630

Asamoah, D., Shahrill, M., & Abdul Latif, S. (2022). A review of formative assessment techniques in higher education during COVID-19. *The Qualitative Report*, *27*(2), 475-487. https://doi.org/10.46743/2160-3715/2022.5145

Balaban, I., & Bubas, G. (2010, June). Educational potentials of ePortfolio systems: Student evaluations of Mahara and Elgg. In *Proceedings of the ITI 2010, 32nd International Conference on Information Technology Interfaces* (pp. 329-336). IEEE.

* Barbera, E. (2009). Mutual feedback in e-portfolio assessment: an approach to the netfolio system. *British Journal of Educational Technology, 40*(2), 342-357. https://doi.org/10.1111/j.1467-8535.2007.00803.x

Barker, K. C. (2006). *Environmental scan: Overview of the ePortfolio in general and in the workplace specifically*. Retrieved from https://www.futured.com/documents/LIfIABackgroundePforSkilledImmigrants.pdf

Barrett, H. (2003). Symposium on Electronic Portfolios. In *Society for Information Technology & Teacher Education International Conference* (pp. 19-21). Association for the Advancement of Computing in Education (AACE). Retrieved from https://www.learntechlib.org/primary/p/17817/.

Barrett, H. (2006). Using electronic portfolios for formative/classroom-based assessment. *Classroom Connect Connected Newsletter*, *13*(2), 4-6. https://electronicportfolios.org/portfolios/ConnectedNewsletter.pdf

Barrett, H. (2010). Balancing the two faces of ePortfolios. *Educ. Form. & Tecnol 3*(1), 6-14. Retrieved from https://eft-edu.com/index.php/eft/article/view/70/54

* Baturay, M. H., & Daloğlu, A. (2010). E-portfolio assessment in an online English language course. *Computer Assisted Language Learning, 23*(5), 413-428. https://doi.org/10.1080/09588221.2010.520671

* Beckers, J., Dolmans, D., & Van Merriënboer, J. (2022). Student, direct thyself! Facilitating self-directed learning skills and motivation with an electronic development portfolio. *Journal of Research on Technology in Education, 54*(4), 617-634. https://doi.org/10.1080/15391523.2021.1906363

Bennett, D., Knight, E., & Rowley, J. (2020). The role of hybrid learning spaces in enhancing higher education students' employability. *British Journal of Educational Technology, 51*(4), 1188-1202. https://doi.org/10.1111/bjet.12931

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education Principles Policy and Practice, 18,* 5–25. https://doi.org/10.1080/0969594X.

Black, P., & Wiliam, D. (1998). *Assessment and classroom learning. Assessment in Education Principles Policy and Practice, 5*, 7–74. https://doi.org/10.1080/

Black, P., & Wiliam, D. (2010). Inside the black box: Raising standards through classroom assessment. *Phi delta kappan, 92*(1), 81-90. https://doi.org/10.1177/003172171009200119

Blaschke, L. M., & Marin, V. (2020). Applications of heutagogy in the educational use of e-portfolios. *Revista de Educación a Distancia (RED), 20*(64). https://doi.org/10.6018/red.407831

Box, C. (2019). The Power of Formative Assessment. In: Formative Assessment in United States Classrooms. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-03092-6_2

* Bramley, A., Forsyth, A., & McKenna, L. (2021). Design, implementation and evaluation of novel work-based clinical assessment tool: an e-portfolio with embedded entrustable professional activities. *Nurse Education Today, 107*, 105101. https://doi.org/10.1016/j.nedt.2021.105101

Bryant, J., & Bates, A. J. (2015). Creating a constructivist online instructional environment. *TechTrends, 59*(2), 17–22. https://doi.org/10.1007/s11528-015-0834-1

Burner, T. (2014). The potential formative benefits of portfolio assessment in second and foreign language writing contexts: A review of the literature. *Studies in Educational Evaluation, 43*, 139-149. https://doi.org/10.1016/j.stueduc.2014.03.002

Chang, C. C., Chou, P. N., & Liang, C. (2018). Using ePortfolio-based learning approach to facilitate knowledge sharing and creation of college students. *Australasian Journal of Educational Technology, 34*(1). https://doi.org/10.14742/ajet.2687

* Chang, C. C., Liang, C., Tseng, K. H., & Tseng, J. S. (2014). Using e-portfolios to elevate knowledge amassment among university students. *Computers & Education, 72*, 187-195. https://doi.org/10.1016/j.compedu.2013.10.015

* Chang, C. C., Tseng, K. H., Liang, C., & Chen, T. Y. (2013). Using e-portfolios to facilitate university students' knowledge management performance: E-portfolio vs. non-portfolio. *Computers & Education, 69*, 216-224. https://doi.org/10.1016/j.compedu.2013.07.017

Chang, S. L., & Kabilan, M. K. (2022). Using social media as e-Portfolios to support learning in higher education: a literature analysis. *Journal of Computing in Higher Education,* 1-28. https://doi.org/10.1007/s12528-022-09344-z

Chappuis, S., & Stiggins, R. J. (2002). Classroom assessment for learning. *Educational Leadership, 60*(1), 40-43. Retrieved from https://www.ascd.org/el/articles/classroom-assessment-for-learning

* Chau, J., & Cheng, G. (2010). Towards understanding the potential of e-portfolios for independent learning: A qualitative study. *Australasian Journal of Educational Technology, 26*(7). 932-950. https://doi.org/10.14742/ajet.1026

Chetcuti, D., Buhagiar, M. A., & Cardona, A. (2011). The professional development portfolio: Learning through reflection in the first year of teaching. *Reflective practice*, *12*(1), 61-72. https://doi.org/10.1080/14623943.2011.541095

* Chuang, H. H. (2010). Weblog-based electronic portfolios for student teachers in Taiwan. *Educational Technology Research and Development, 58,* 211-227. https://doi.org/10.1007/s11423-008-9098-1

Cohen, L. M., Manion, L., & Morrison, K. (2007). *Research methods in education.* Routledge.

* Contreras-Higuera, W. E., Martínez-Olmo, F., José Rubio-Hurtado, M., & Vilà-Baños, R. (2016). University students' perceptions of e-portfolios and rubrics as combined assessment tools in education courses. *Journal of Educational Computing Research*, *54*(1), 85-107. https://doi.org/10.1177/0735633115612784

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Pearson.

De Grez, L., Valcke, M., & Roozen, I. (2012). How effective are self-and peer assessment of oral presentation skills compared with teachers' assessments?. *Active Learning in Higher Education*, *13*(2), 129-142. https://doi.org/10.1177/1469787412441284

* Demir, B. P., & Kutlu, Ö. (2016). The Effect of Electronic Portfolio Applications on 6th Graders' Research Skills. *Education & Science/Egitim ve Bilim, 41*(188). http://dx.doi.org/10.15390/EB.2016.6724

* De Swardt, M., Jenkins, L. S., Von Pressentin, K. B., & Mash, R. (2019). Implementing and evaluating an e-portfolio for postgraduate family medicine training in the Western Cape, South Africa. *BMC Medical Education, 19,* 1-13. https://doi.org/10.1186/s12909-019-1692-x

Dunn, K. E., & Mulvenon, S. W. (2009). A critical review of research on formative assessments: The limited scientific evidence of the impact of formative assessments in education. *Practical Assessment, Research, and Evaluation, 14*(1), 7. https://doi.org/10.4324/9780203462041

* Ebil, S., Salleh, S. M., & Shahrill, M. (2020). The use of E-portfolio for self-reflection to promote learning: A case of TVET students. *Education and Information Technologies*, *25*(6), 5797-5814. https://doi.org/10.1007/s10639-020-10248-7

* Fathi, J., & Rahimi, M. (2022). Electronic writing portfolio in a collaborative writing environment: its impact on EFL students' writing performance. *Computer Assisted Language Learning*, 1-39. https://doi.org/10.1080/09588221.2022.2097697

Fuller, J. S., & Dawson, K. M. (2017). Student response systems for formative assessment: Literature-based strategies and findings from a middle school implementation. *Contemporary Educational Technology,* 8(4), 370-389. https://doi.org/10.30935/cedtech/6206

Gamlem, S., M. & Smith, K. (2013). Student perceptions of classroom feedback. *Assessment in Education: Principles, Policy & Practice, 20*(2), 150-169. https://doi.org/10.1080/0969594X.2012.749212

Gardner, K. M., & Aleksejuniene, J. (2008). Quantitative and qualitative analysis of student feedback on ePortfolio learning. *Journal of Dental Education*, *72*(11), 1324-1332. https://doi.org/10.1002/j.0022-0337.2008.72.11.tb04615.x

* Garrett, N. (2011). An e-portfolio design supporting ownership, social learning, and ease of use. *Journal of Educational Technology & Society*, *14*(1), 187-202. Retrieved from https://www.learntechlib.org/p/75362/.

Gezer, T., Wang, C., Polly, A., Martin, C., Pugalee, D., & Lambert, R. (2021). The relationship between formative assessment and summative assessment in Primary grade mathematics classrooms. *International Electronic Journal of Elementary Education, 13*(5), 673-685. Retrieved from https://www.iejee.com/index.php/IEJEE/article/view/1508

Gibson, D. (2006). ePortfolio Decisions and Dilemas. Jafari, A. & Kaufman, C(Eds). In *Handbook of research on ePortfolios* (pp. 135-145). IGI Global.

Guba, E. G., & Lincoln, Y. S. (1982). The place of values in needs assessment. *Educational Evaluation and Policy Analysis*, *4*(3), 311-320. https://doi.org/10.2307/1164061

Gulbahar, Y. (2009). Usage of Electronic Portfolios for Assessment. In *Handbook of Research on New Media Literacy at the K-12 Level: Issues and Challenges* (pp. 702-719). IGI Global. https://www.igi-global.com/gateway/chapter/35946

Hallam, G., & Creagh, T. (2010). ePortfolio use by university students in Australia: A review of the Australian ePortfolio Project. *Higher Education Research & Development*, *29*(2), 179-193. https://doi.org/10.1080/07294360903510582

Heath, M. (2005). Are You Ready to Go Digital?: The Pros and Cons of Electronic Portfolio Development. *Library Media Connection, 23*(7), 66-70. Retrieved from https://eric.ed.gov/?id=EJ717645

Hegarty, B., & Thompson, M. (2019). A teacher's influence on student engagement: Using smartphones for creating vocational assessment ePortfolios. *Journal of Information Technology Education Research, 18*, 113. https://doi.org/10.28945/4244

Heinrich, E., Bhattacharya, M. & Rayudu, R. (2007). Preparation for lifelong learning using ePortfolios. *European Journal of Engineering Education, 32(*6), 653–663. https://doi.org/10.1080/03043790701520602

Heitink, M. C., Van der Kleij, F. M., Veldkamp, B. P., Schildkamp, K., & Kippers, W. B. (2016). A systematic review of prerequisites for implementing assessment for learning in classroom practice. *Educational research review, 17*, 50-62. https://doi.org/10.1016/j.edurev.2015.12.002

Himpsl, K., & Baumgartner, P. (2009). Evaluation of e-portfolio software. *International Journal of Emerging Technologies in Learning (iJET), 4*(1), 16-22. https://doi.org/10.3991/ijet.v4i1.831

Hoover-Dempsey, K. V., & Sandler, H. M. (1995). Parental involvement in children's education: Why does it make a difference?. *Teachers college record*, *97*(2), 310-331. Retrieved from http://futured.com/documents/LIfIABackgroundePforSkilledImmigrants.pdf

* Huang, H. T. D., & Hung, S. T. A. (2010). Implementing electronic speaking portfolios: Perceptions of EFL students. *British Journal of Educational Technology, 41*(5), E84-E88. https://doi.org/10.1111/j.1467-8535.2009.00996.x

* Hung, S. T. A. (2012). A washback study on e-portfolio assessment in an English as a Foreign Language teacher preparation program. *Computer Assisted Language Learning, 25*(1), 21-36. https://doi.org/10.1080/09588221.2010.551756

Irving, K. (2015). Technology-assisted formative assessment. *In M. J. Urban and D. A. Falvo (Eds.), Improving K-12 STEM education outcomes through technological integration* (pp. 380-398). IGI Global.

* Jan van Aalst & K. K. Chan (2007) Student-Directed assessment of knowledge building using electronic portfolios. *Journal of the Learning Sciences, 16*(2), 175-220, https://doi.org/10.1080/10508400701193697

* Kabilan, M. K. (2016). Using Facebook as an e-portfolio in enhancing pre-service teachers' professional development. *Australasian Journal of Educational Technology*, *32*(1). https://doi.org/10.14742/ajet.2052

* Kabilan, M. K., & Khan, M. A. (2012). Assessing pre-service English language teachers' learning using e-portfolios: Benefits, challenges and competencies gained. *Computers & Education, 58(4)*, 1007-1020. https://doi.org/10.1016/j.compedu.2011.11.011

Karaoğlan Yılmaz, F. G., Yılmaz, R., & Öztürk, H. T. (2020). Examining University Students' Acceptance of Web-based Formative Assessment System. *Journal of Hacettepe University Education Faculty. 35*(Special Issue), 25-36. https://dx.doi.org/10.16986/HUJE.2020063671

Kerr, W. (2007). *Service user and staff perspectives on recovery in psychosis: Talking the same language and research portfolio.* University of Glasgow (United Kingdom). Retrived from https://www.proquest.com/docview/2023796799?pq-origsite=gscholar&fromopenview=true

* Khodi, A., Khezerlou, H., & Sahraei, H. (2022). Dependability and utility of using e-portfolios in assessing EFL learners' speaking proficiency. *Computer Assisted Language Learning*, 1-23. https://doi.org/10.1080/09588221.2022.2093379

Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele, UK, Keele University*, *33*(2004), 1-26.

Klenowski, V. (2009). Australian Indigenous students: Addressing equity issues in assessment. *Teaching Education*, *20*(1), 77-93. https://doi.org/10.1080/10476210802681741

Kutlu, Ö., Doğan, C. D., & Karakaya, İ. (2017). *Öğrenci Başarısının Belirlenmesi- Performansa ve Portfolyoya Dayalı Durum Belirleme*. Pegem Publishing.

Lam, R. (2022). Assessing creative writing formatively and summatively with e-Portfolios: A case study in Hong Kong. Chamcharatsri, B. & Iida, A. (Eds). *In International Perspectives on Creative Writing in Second Language Education* (pp. 171-188). Routledge. https://doi.org/10.4324/9781003183914-12

Landis, J. R. & Koch, G. G. (1977) The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174. https://doi.org/10.2307/2529310

Leighton, J. P. (2019). Students' interpretation of formative assessment feedback: Three claims for why we know so little about something so important. *Journal of Educational Measurement, 56*(4), 793-814. https://doi.org/10.1111/jedm.12237

* López-Crespo, G., Blanco-Gandía, M. C., Valdivia-Salas, S., Fidalgo, C., & Sánchez-Pérez, N. (2021). The educational e-portfolio: preliminary evidence of its relationship with student's self-efficacy and engagement. *Education and Information Technologies*, 1-16. https://doi.org/10.1007/s10639-021-10827-2

Lui, A. M., & Andrade, H. L. (2022, February). The next black box of formative assessment: A model of the internal mechanisms of feedback processing. *In Frontiers in Education* (Vol. 7, p. 20). Frontiers. https://doi.org/10.3389/feduc.2022.751548

Mashauri, M. M. (2023). Formative Assessment in EFL Classrooms in the Democratic Republic of the Congo: Analysis of Teachers' Judgements at Oicha Senior School. *Open Access Library Journal, 10*(5), 1-11. https://doi.org/10.4236/oalib.1110210

Mason, R., Pegler, C., & Weller, M. (2004). E-portfolios: an assessment tool for online courses. *British Journal of Educational Technology, 35*(6), 717-727. https://doi.org/10.1111/j.1467-8535.2004.00429.x

* McLeod, J. K., & Vasinda, S. (2009). Electronic portfolios: Perspectives of students, teachers and parents. *Education and Information Technologies, 14*, 29-38. https://doi.org/10.1007/s10639-008-9077-5

* Meyer, E., Abrami, P. C., Wade, C. A., Aslan, O., & Deault, L. (2010). Improving literacy and metacognition with electronic portfolios: Teaching and learning with ePEARL. *Computers & Education, 55(*1), 84-91. https://doi.org/10.1016/j.compedu.2009.12.005

Meyer, E. J., Abrami, P. C., Wade, A., & Scherzer, R. (2011). Electronic portfolios in the classroom: Factors impacting teachers' integration of new technologies and new pedagogies. *Technology, Pedagogy and Education*, *20*(2), 191-207. https://doi.org/10.1080/1475939X.2011.588415

Mgarbi, H., Chkouri, M., & Tahiri, A. (2022). Purpose and Design of a Digital Environment for the Professional Insertion of Students Based on the E-portfolio Approach. International *Journal of Emerging Technologies in Learning (iJET)*, *17*(1), 36-48. https://doi.org/10.3991/ijet.v17i01.26703

Mitchell, L., Campbell, C., Somerville, M., Cardell, E., & Williams, L. T. (2021). Enhancing graduate employability through targeting ePortfolios to employer expectations: A systematic scoping review. *Journal of Teaching and Learning for Graduate Employability*, *12*(2), 82-98. https://doi.org/10.21153/jtlge2021vol12no2art1003

Morris, R., Perry, T., & Wardle, L. (2021). Formative assessment and feedback for learning in higher education: A systematic review. *Review of Education, 9(*3), e3292. https://doi.org/10.1002/rev3.3292

* Nicolaidou, I. (2013). E-portfolios supporting primary students' writing performance and peer feedback. *Computers & Education, 68*, 404-415. https://doi.org/10.1016/j.compedu.2013.06.004

Nitko, A. J., & Brookhart, S. M. (2014). *Educational assessment of students*. Pearson.

Oh, J. E., Chan, Y. K., & Kim, K. V. (2020). Social media and e-portfolios: impacting design students' motivation through project-based learning. *IAFOR Journal of Education*, *8*(3), 41-58. https://doi.org/10.22492/ije.8.3.03

Pachler, N., Daly, C., Mor, Y., & Mellar, H. (2010). Formative e-assessment: Practitioner cases. *Computers & Education*, *54*(3), 715-721. https://doi.org/10.1016/j.compedu.2009.09.032

Peacock, S., Scott, A., Murray, S., & Morss, K. (2012). Using feedback and ePortfolios to support professional competence in healthcare learners. *Research in Higher Education Journal*, 16. http://www.aabri.com/manuscripts/111070.pdf

Pierre, E., & Oughton, J. (2007). The Affective Domain: Undiscovered Country. *College Quarterly, 10*(4), 1-7.

Ramsey, B. & Duffy, A. (2016). *Formative assessment in the classroom: Findings from three districts.* Michael and Susan Dell Foundation and Education, 1. Retrieved from https://education-first.com/wpcontent/uploads/2016/05/MSDF-Formative-Assessment-Study-Final-Report.pdf

* Ruiz, J. G., Qadri, S. S., Karides, M., Castillo, C., Milanez, M., & Roos, B. A. (2009). Fellows' perceptions of a mandatory reflective electronic portfolio in a geriatric medicine fellowship program. *Educational Gerontology, 35*(7), 634-652. https://doi.org/10.1080/03601270902877360

* Sánchez Gómez, S., Ostos, E. M. C., Solano, J. M. M., & Salado, T. F. H. (2013). An electronic portfolio for quantitative assessment of surgical skills in undergraduate medical education. *BMC Medical Education, 13(*1), 1-9. https://doi.org/10.1186/1472-6920-13-65

Schildkamp, K., Van der Kleij, F. M., Heitink, M. C., Kippers, W. B., & Veldkamp, B. P. (2020). Formative assessment: A systematic review of critical teacher prerequisites for classroom practice. *International Journal of Educational Research, 103*, 101602. https://doi.org/10.1016/j.ijer.2020.101602

Segaran, M. K., & Hasim, Z. (2021). Self-regulated Learning through ePortfolio: A Meta-analysis. *Malaysian Journal of Learning and Instruction*, *18*(1), 131-156. Journal of Learning and Instruction, 18(1), 131-156. https://doi.org/10.32890/mjli2021.18.1.6

* Sharifi, M., Soleimani, H., & Jafarigohar, M. (2017). E-portfolio evaluation and vocabulary learning: Moving from pedagogy to andragogy. *British Journal of Educational Technology, 48(*6), 1441-1450. https://doi.org/10.1111/bjet.12479

* Shepherd, C. E., & Bolliger, D. U. (2011). The effects of electronic portfolio tools on online students' perceived support and cognitive load. *The Internet and Higher Education, 14(*3), 142-149. https://doi.org/10.1016/j.iheduc.2011.01.002

Slepcevic-Zach, P., & Stock, M. (2018). ePortfolio as a tool for reflection and self-reflection. *Reflective Practice, 19*(3), 291-307. https://doi.org/10.1080/14623943.2018.1437399

Stefani, L., Mason, R., & Pegler, C. (2007). *The educational potential of e-portfolios: Supporting personal development and reflective learning*. Routledge. https://doi.org/10.4324/9780203961292

Stemler, S. (2000). An overview of content analysis. *Practical assessment, research, and evaluation*, *7*(1), 17. https://doi.org/10.7275/z6fm-2e34

Stobart, G. (2008). *Testing times: The uses and abuses of assessment.* Routledge. https://doi.org/10.4324/9780203930502

* Struyven, K., Blieck, Y., & De Roeck, V. (2014). The electronic portfolio as a tool to develop and assess pre-service student teaching competences: Challenges for quality. *Studies in Educational Evaluation, 43,* 40-54. https://doi.org/10.1016/j.stueduc.2014.06.001

Sweet, D. (1993). Student portfolios: Classroom uses. *Office of Education Research: Consumer Guide, 8,* 1-5.

Tillema, H. H. (2001). Portfolios as developmental assessment tools. *International Journal of Training and Development*, *5*(2), 126-135. https://doi.org/10.1111/1468-2419.00127

Törmälä, V. (2021). ePortfolios in craft education at the primary level: Teachers' experiences on ICT integration. *Design and Technology Education,* *26*(2), 28-45. Retrieved from https://ojs.lboro.ac.uk/DATE/article/view/2918

Van der Kleij, F. M., Vermeulen, J. A., Schildkamp, K., & Eggen, T. J. (2015). Integrating data-based decision making, assessment for learning and diagnostic testing in formative assessment. *Assessment in Education: Principles, Policy & Practice, 22*(3), 324-343. https://doi.org/10.1080/0969594X.2014.999024

Vankúš, P. (2021). Influence of game-based learning in mathematics education on students' affective domain: A systematic review. *Mathematics, 9*(9), 986. https://doi.org/10.3390/math9090986

Vogel, M. (2018). Combining and integrating formative and summative assessment in mathematics teacher education. ZDM – *Mathematics Education, 50*(3), 519-531. https://doi.org/10.1007/s11858-018-0948-y

Wade, A., Abrami, P., & Sclater, J. (2005). An electronic portfolio to support learning. *Canadian Journal of Learning and Technology, 31*(3). https://doi.org/10.21432/T2H30P

* Wang, L. (2010). Integrating communities of practice in e-portfolio assessment: Effects and experiences of mutual assessment in an online course. *The Internet and Higher Education, 13*(4), 267-271. https://doi.org/10.1016/j.iheduc.2010.07.002

* Wang, P., & Jeffrey, R. (2017). Listening to learners: An investigation into college students' attitudes towards the adoption of e-portfolios in English assessment and learning. *British Journal of Educational Technology, 48*(6), 1451-1463. https://doi.org/10.1111/bjet.12513

Wiliam, D. (2011). What is assessment for learning?. *Studies in educational evaluation, 37*(1), 3-14. https://doi.org/10.1016/j.stueduc.2011.03.001

Wiliam, D., & Black, P. (1996). Meanings and consequences: a basis for distinguishing formative and summative functions of assessment. *British Educational Research Journal, 22*(5), 537–548. https://doi.org/10.1080/0141192960220502

Yancey, K. B. (2023). *ePortfolio as curriculum: Models and practices for developing students' ePortfolio literacy.* Taylor & Francis.

* Yang, M., Tai, M., & Lim, C. P. (2016). The role of e-portfolios in supporting productive learning. British *Journal of Educational Technology, 47*(6), 1276-1286. https://doi.org/10.1111/bjet.12316

Zhang, X., Olfman, L., Ractham, P., & Firpo, D. (2009, May). The implementation and evaluation of KEEP SLS: An ePortfolio system supporting social constructive learning. In *Proceedings of the special interest group on management information system's 47th annual conference on Computer personnel research* (pp. 13-18).

* Zheng, Y., & Barrot, J. S. (2022). Social media as an e-portfolio platform: Effects on L2 learners' speaking performance. *Language Learning & Technology, 26*(1). https://hdl.handle.net/10125/73487

ISSN: 1309 – 6575
Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi
Journal of Measurement and Evaluation in Education and Psychology
Research Article; 2023; 14(Special Issue); 320-331

Examining Students' Online Formative Test-Taking Behaviors Using Learning Analytics

Alper BAYAZIT* Denizer YILDIRIM** Gökhan AKÇAPINAR*** Hale ILGAZ****

Abstract

In online learning environments, assessment is an important dimension and also one of the most challenging parts of the online learning process. So, to provide effective learning, analyzing students' behaviors is important when designing online formative and summative assessment environments. In this study, students' profiles were analyzed within an online formative assessment environment and compared with a summative assessment environment based on attempt count, overall time spent, first-attempt score, and last-attempt score metrics. The within-subjects design, cluster analysis, and the Kruskal Wallis-H Test was carried out for analyzing behaviors. As a result, it was shown in the data that there were three main clusters. Cluster 1 showed a high number of interactions, and an increasing trend was observed in "grades" over "attempts". Additionally, Cluster 2 consisted of students who received the best grades in all other clusters, and finally, Cluster 3 consisted of students who interacted little and scored lower on formative assessments.

Keywords: test-taking behavior analysis, learning analytics, formative assessment, assessment analytics

Introduction

Assessment is one of the primary components of the online learning process. Two of the main approaches used in assessment design are formative and summative assessment. While both assessment approaches are focused on student development and progress, their approaches differ. Harlen and James (1997) defined formative assessment as an "assessment for learning" that focuses on student learning at the current stage and supports learning for the next step. While summative assessment, also defined as an "assessment of learning", is considered a more systematic and continuous recording of overall achievement. As a result, formative assessment focuses on the improvement of learning, whereas summative assessment focuses on providing information for accreditation and evaluation (Xiong et al., 2018). In an online course based on a formative assessment approach, students are aware of their strengths and weaknesses, can be more engaged and motivated, and monitor their progress (Crisp & Ward, 2008; Wolsey, 2008). Additionally, formative assessment can help decrease students' anxiety levels (Cassady & Gridley, 2005) and increase interaction between peers and instructors (Vonderwell et al., 2007). For this reason, the design of formative assessment is important in terms of providing information about how students perceive this process.

* Asst. Prof. Dr., Ankara University, Faculty of Medicine, Ankara-Turkiye, abayazit@ankara.edu.tr, ORCID ID: 0000-0003-4369-587X
** Dr., Ankara University, Faculty of Open and Distance Education, Ankara-Turkiye, dyildirim@ankara.edu.tr, ORCID ID: 0000-0002-4534-8153
*** Assoc. Prof. Dr., Hacettepe University, Faculty of Education, Ankara-Turkiye, gokhana@hacettepe.edu.tr, ORCID ID: 0000-0002-0742-1612
**** Assoc. Prof. Dr., Ankara University, Faculty of Open and Distance Education, Ankara-Turkiye, hilgaz@ankara.edu.tr, ORCID ID: 0000-0001-7011-5354

To cite this article:
Bayazıt, A., Yıldırım, D., Akçapınar, G. & Ilgaz, H. (2023). Examining students' online formative test-taking behaviors using learning analytics, *14*(Special Issue), 320-331. https://doi.org/10.21031/epod.1275597

Received: 2.04.2023
Accepted: 3.09.2023

Technology use during the formative assessment process is especially seen as a pillar of tracking students’ learning and performance within 21st-century learning environments (Shin et al., 2022). However, due to the nature of the formative assessment process, students sometimes tend to put less effort into these tests, which can ultimately cause poorer results compared to the summative assessment process (Wise, 2006; Yildirim-Erbasli & Bulut, 2022). The effort shown during test-taking is an important dimension of the formative assessment environment and allows the documentation of student engagement (Wise et al., 2013; Yildirim-Erbasli & Bulut, 2020). Test-taking frequency (Blondeel et al., 2023; Palmen et al., 2015; Shin et al., 2022) and response time and patterns (Man et al., 2018; Yildirim-Erbasli & Bulut, 2020; Wise & DeMars, 2006) are also indicators of students’ engagement and motivation within the formative assessment process.

It is essential to investigate students' behaviors during the self-assessment process to find patterns that have a negative effect on learning (Yang et al., 2022). Also, revealing test-taking behavior allows for the identification of students' trial-and-error patterns and cheating activities (Man et al., 2018) and provides timely feedback for maintaining mastery of learning (Hui, 2023). It also enables researchers and/or administrators to gain greater insight into the processes and behaviors that lead to a specific test outcome (Stadler et al., 2020). Investigation of formative assessment behaviors also helps to make intervention possible and can significantly impact student interests and achievements (Rakoczy et al., 2019). Typically, these scales have been employed by researchers to examine exam-taking behaviors; however, self-report measures are weak against many forms of bias, and at times these scales can be quite limited, as they only provide fundamental information regarding a student's motivation towards test taking. Additionally, it can be hard to know how accurately test takers complete the scale (Wise & Gao, 2017). Thus, log data analysis can provide important information regarding students' behavior in formative online assessments for both professionals and educators (Guo, 2021). This data can also be useful for detecting differences in students’ aberrant behavior in real time (Han & Kang, 2021) because it is a reliable predictor of the ability to be tested (Stadler et al., 2020). For example, students' interactions during the evaluation process are typically stored in log data. Some of these metrics include the number of student attempts, number of question views, number of hints viewed, number of submissions (Yang et al., 2022), response time data, total scores (Guo & Ercikan, 2021), test scores and submission times (Hui, 2023), individual mean item response times (Lee & Haberman, 2016), omitted items (Sarac & Loken, 2022), time-on-task and the number of interactions (Stadler et al., 2020), and flagged, reviewed, changed, or omitted items (Wise & Gao, 2017). These studies show that the trial scores, response time, and trial numbers are the most often used metrics. However, the pre-processing of data obtained from several assessment tools is seen as a limitation, and as a result, obtaining the necessary metrics from commonly used learning management systems can contribute to the literature and provide important insights into students' test-taking strategies and student profiles.

One test-taking performance study by Silm et al. (2013) focused on performance in low-stakes tests. According to their aim, they specifically searched for the number of items test-takers attempted to solve, the number of correct answers, overall time spent, and the speed of their accomplishments. The research was focused on 327 first-year students attending a higher education institution. It was found that when the difficulty levels of items are similar, the number of items solved and the mean time for each item can predict performance in low-stakes tests and short response times signal low-test scores. It was also discovered that the mean time for incorrect answers was shorter than the mean time for correct answers. Additionally, in another study using responses and response times for computer-based reading assessment, Yildirim-Erbasli and Bulut (2020) evaluated the effect of students' test-taking efforts on their reading growth. A quick screening tool was designed and applied to 7602 students over an academic year to monitor and assess their reading ability. They found that rapid guessing and slow responses can be helpful when calculating and interpreting students’ growth estimates. In a large-scale test, Programme for International Student Assessment (PISA) 2012, Lundgren and Eklöf (2020) focused on test-takers' within-item behaviors from a self-reported and behavioral effort perspective. Essentially, they analyzed time on task, time to first action, number of actions, unique routes, repeated wrong routes, and actions per minute variables with self-reported data by using math test scores. Thus, they determined that low levels of effort before completing a task may not be diagnostic of test-taking motivation, although low levels of effort before taking a test appear to be below the level of effort put in prior to giving up on a task. So, these variables/metrics provide important information when analyzing and interpreting the formative assessment results.

Clustering is employed for profiling purposes due to the clustering technique's interpretable insights regarding the relationship between test administration decisions and student performance profiles (Shin et al., 2022). Profiling test-takers or profiling students' test-taking behaviors are two of the primary approaches in clustering studies. For example, Yang et al. (2022) used cluster analysis to analyze how students interacted during formative assessments, which were given as a post-class self-evaluation. In the results, three distinct student profile clusters were determined. The students in Profile 1 were those who engaged the assessment system and hinted at it sparingly. While in Profile 2, students participated in exams, reviewed questions, and struggled to answer them. On the other hand, students in Profile 3 were those who successfully remembered information and used the exam system most. Therefore, according to these findings, students who completed online tests after class typically scored higher than those who did not, and those who engaged in non-standard behavior during the test did not increase their performance. Tempelaar et al. (2018) demonstrated that different at-risk groups might be identified by clustering several interaction data items including formative assessments. They highlighted that appropriate interventions are available by identifying these profiles. Additionally, Guo (2021) found four distinct student profiles in online exams taken at home or testing centers. Test-takers in only certain clusters tended to spend the majority of their time solving items, whereas, in other clusters, they were to have read the exam instructions for a significant portion of the time. In another study, Stenlund et al. (2018) clustered test-takers into groups and discovered three distinct student profiles: moderate, calm risk-taker, and test anxious risk-averse profile. Furthermore, in group difference studies, it was revealed that in terms of test performance, the calm risk-taker profile was the most successful, while the test anxious risk-averse profile was the least successful. Another approach in profiling studies is clustering individual test attempts rather than clustering an average over modules or students. Hui (2023) also stated that possible differences between scores and test types could be clustered and analyzed separately. Thus, computer-based techniques can be utilized to detect students who do not exhibit normal patterns by revealing behavior patterns through clustering. For example, Liao et al. (2021) claimed it is challenging to identify "item harvesters" who memorize or share test items. Thus, they offered a two-stage solution to identify this behavior, which, in the end, appeared to make tests less reliable. As a result, the initial phase should include cluster analysis to identify learners' test-taking behaviors, and then, abnormal behaviors must be marked for further investigation.

The clustering technique provides valuable information within a learning analytics framework to detect these general behaviors. It was shown in the presented literature that especially test-taking effort, engagement, and motivation play important roles in online formative assessments. When focusing on these dimensions, researchers worked in particular with the number of items that test-takers attempted to solve, the number of correct answers, overall time spent, and response time. Importantly, these user metrics had not been previously found to be analyzed by comparing summative and/or online formative test environments. Therefore, in the current study, students' profiles were analyzed within an online formative assessment environment and then compared with the summative assessment environment based on the metrics of attempt count, overall time spent, first attempt score, and the last attempt score.

Research Questions:

1. How many different groups are students divided into according to the metrics of taking the formative test?
2. Is there a significant difference between students in different groups in terms of their summative test scores?

Method

In the current study, a within-subjects design was used, and comparisons were made of the test-taking behaviors of participants in the weekly quizzes (formative assessment) and mid-term exams (summative assessment). Features related to the test-taking behaviors of students were extracted from the database through data mining methods. Also, cluster analysis and the Kruskal Wallis-H Test were used to test whether there was a statistically significant difference in test-taking behaviors of students for the online formative and summative tests.

Participants

This study was conducted at a state university with 66 vocational school students enrolled in Web Programming II courses as part of the Computer Programming and Internet and Network Technologies Distance Education Programs. The students' ages ranged from 18-45 years old, and the group included 51 males and 15 females. The duration of the course lasted approximately 14 weeks, and the theoretical transfer and implementation of the course were conducted online. Finally, the mid-term and final exams in this study were administered in a face-to-face setting.

Procedure

The research process was carried out between the sixth and eighth week of the course. Students had access to the formative assessment test following the lecture in week six until week eight when the mid-term exam was held. This process involved the formative assessment design that included the processes of "finding and handling information" and "assessment" in the Learning Design Taxonomy (Toetenel & Rientes, 2016). The formative test was structured using a question bank containing questions that had previously been used for the mid-term exam in the previous semester. Importantly, the question bank included subjects relevant to the mid-term examination and was systematically categorized into distinct units. Notably, the distribution of questions across units within the question bank was found to be uneven. Nonetheless, during the examination process, each student uniformly encounters an equivalent number of questions from each unit. To illustrate, each student has posed four questions sourced from both Unit 1 and Unit 2. These questions might manifest either as a duplicate or exhibit some variation amongst themselves. In each attempt of the formative test, the students were randomly assigned 20 questions from the question bank. The students were then asked to answer the questions they had been provided within 30 minutes. At the end of each attempt, the question itself, the students' answers, and feedback regarding the correct or not correct answers were shown. Furthermore, the students were not shown the correct answers to any of the questions. When the students did answer a question incorrectly, they were expected to peruse the course resources themselves to determine the correct answer as well as review the instructor's lecture recordings and discuss what they had learned.

The summative assessment was comprised of a total of 20 multiple-choice questions, and all students engaged in the examination under uniform conditions. Additionally, the examination was administered according to a specific online framework. Also, students were diligently supervised by an instructor through an online live virtual classroom. Furthermore, the students' interactions within the web browser were subject to scrutiny via the implementation of a Proctoring Moodle Plugin. As a result, following the conclusion of the examination, analysis of the test items was carried out and indices related to the item difficulty spanned a spectrum from 0.23 to 0.78. Therefore, the reliability coefficient of the items, as assessed through Cronbach's alpha measure, was determined to be 0.695.

Data Analysis

In answering the first research question, students' formative exam-taking behaviors (digital traces during the exam) were analyzed using the four metrics mentioned in Table 1. Cluster analysis was used to group the students based on their exam-taking behaviors, and with this analysis, it was determined there were common exam-taking behaviors among the students. Due to there being no prior insight into the number of clusters, the number was determined automatically using the Silhouette metric. Thus, the k-Means algorithm in Orange data mining software was used for the cluster analysis. Since the distributions of the data were not in a standard range, normalization was applied in the data pre-processing. Additionally, the Silhouette scores of clusters from 2 to 9 were calculated to determine the ideal number of clusters, and the number of clusters with the highest Silhouette score was considered as the ideal number of clusters.

Table 1.
Metrics used in the clustering and their explanations

Metric	Description
Attempt count	Total number of attempts by the student in the formative assessment activity
Overall time spent	Total time spent by the student in formative assessment activities (min)
First attempt grade	Student's grade for the first attempt in the formative assessment activity
Last attempt grade	Student's grade for the last attempt in the formative assessment activity

Next, to answer the second research question, a Kruskal Wallis-H test was used to compare summative test scores based on clusters. During this process, a non-parametric statistical analysis was also used due to the normality assumption not being met.

Results

In this study, students' profiles were analyzed within a formative assessment environment and then compared with an established summative assessment environment by gathering several learning management system metrics. Thus, these metrics were used for clustering students and determining their test-taking behaviors. As a result, to analyze this situation in further depth, two research questions were studied.

RQ1. How many different groups are students divided into according to the metrics of taking the formative test?

When the Silhouette scores of clusters 2-9 were examined, it was recognized that the students were ideally divided into three clusters. When considering the cluster centroids provided in Figure 1, along with the descriptive statistics provided in Table 2, it was determined that the students in Cluster 3 (n = 7) made fewer attempts (*Mdn* = 2) and spent less time (*Mdn* = 10 min) in the formative assessment activity than students in the other cluster. Additionally, when the first and last attempt grades were examined, it was noteworthy that both were found to be low, and there was a negligible score increase within students' last trial compared to their first trial.

Next, students in Cluster 1 (n = 20) were those who made the most attempts (*Mdn* = 11) and also spent the most time in the formative assessment activity (*Mdn* = 194 min). As a result, according to the median trial numbers presented in Table 2, students within this group made four times more attempts than those in Cluster 2 and six times more than those in Cluster 3. Similarly, considering the median time spent, students in Cluster 1 spent four times more time than those in Cluster 2 and 20 times more than those in Cluster 3. However, when their scores from the exams were examined, it was recognized that their first-attempt grades were low, and their last-attempt grades were seen to be high. In other words, it was observed that there was a three-fold increase between their first and last attempt grades.

On the other hand, students in Cluster 2 (n = 38) were found to fall in between Cluster 1 and Cluster 3 in terms of the number of attempts (*Mdn* = 3) and time they spent (*Mdn* = 50 min) when considering their cluster centroids. However, despite fewer attempts, their first-attempt grades were determined to be higher than both clusters (*Mdn* = 15). This value was two times higher than the median first-attempt grades of Cluster 1 and approximately three times higher than that of Cluster 3. Therefore, it was determined there was an increase of four points between the median grades of their first and last attempt.

Figure 1.
Cluster Centroids

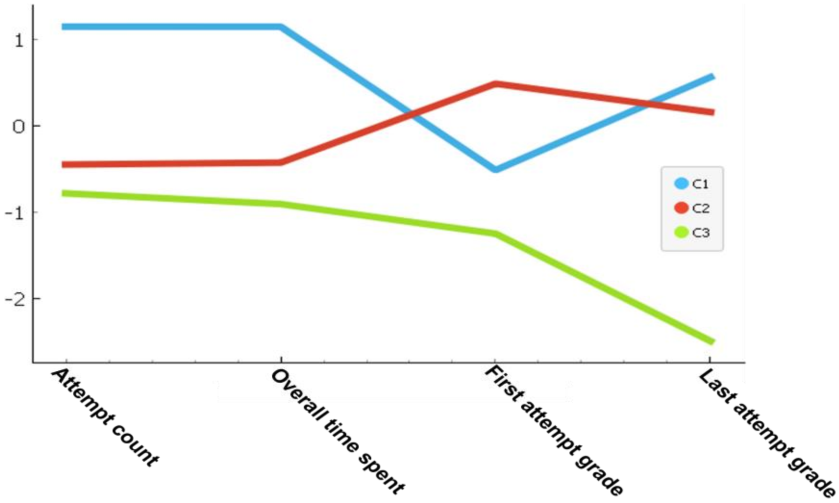


Table 2.
Descriptive statistics by clusters

Cluster	Statistics	Attempt count	Overall time spent	First attempt grade	Last attempt grade
Cluster 1	Mean	12.15	205.52	5.35	17.05
	N	20.00	20.00	20.00	20.00
	Std. Deviation	6.11	100.84	3.84	2.42
	Median	11.00	194.02	5.50	17.00
Cluster 2	Mean	3.28	55.18	10.69	15.00
	N	39.00	39.00	39.00	39.00
	Std. Deviation	1.88	36.19	4.68	2.47
	Median	3.00	50.30	11.00	15.00
Cluster 3	Mean	1.43	9.36	1.43	1.57
	N	7.00	7.00	7.00	7.00
	Std. Deviation	.53	9.40	2.51	2.70
	Median	2.00	9.60	4.00	5.00

RQ 2. Is there a significant difference between students in different groups in terms of summative test scores?

When comparing the summative test scores according to clusters, the non-parametric Bonferroni Correction with the Kruskal Wallis-H Test was applied since Levene's Test statistic was significant ($L = 5.890$, $df1 = 2$, $df2 = 59$, $p <. 01$).

Table 3.
Kruskal Wallis-H Test results

Cluster[a]	Mean	N	SD	Median	Mean Rank	Chi-Square[b]	df	p
Cluster 1	35.500	20	11.34	35.00	23.05	8.747	2	.013[c]
Cluster 2	49.359	39	18.36	40.00	36.67			
Cluster 3	33.333	3	5.77	30.00	20.67			
Total	44.113	62	17.28	40.00				

a. Kruskal Wallis-H Test
b. Grouping Variable: Cluster, Dependent Variable: Summative Grade
c. Significant: Bonferroni Correction $\frac{\rho}{n(n-1)/2}$ = *.05/3 = .01667* ⇒ .013 < *.01667*

According to the results from the Kruskal Wallis-H Test (Table 3), it was determined that there was a significant difference between clusters in terms of summative test scores (Chi-Square = 8.747; df = 2, < .0167). Thus, learners in the high-scoring group with fewer attempt counts in the formative test (Cluster 2, Mean: 49.359, Mean Rank: 36.67) achieved higher success within the summative test than those from the other two clusters. As a result, a significant difference was found to be present only between Cluster 1 and Cluster 2 in terms of summative assessment scores (U = 220, W = 430, Z = -2.734, p = .006, Adj. p = .016). Furthermore, no significant difference was found between Cluster 3 - Cluster 1 (U = 29, W = 35, Z = -0.93, p = .830, Adj. p = 1.00) and Cluster 3 - Cluster 2 (U = 27, W = 33, Z = -1.545, p = .137, Adj. p = .411).

Discussion

Analysis of learning progressions in exams, along with the dimensions of task design, trustworthiness, and fairness, is one way to take advantage of the potential of assessment analytics (Gašević et al., 2022). In the current study, the formative exam metrics collected from distance education students and students' test-taking behaviors in formative assessment were investigated through cluster analysis. As a result, it was revealed through the cluster analysis that three distinct student profiles were present via the assessment analytics.

Additionally, students in Cluster 1 showed a high number of interactions; while receiving low scores in their first attempts, they increased their scores in further attempts. When considering the first attempts, these students exhibited low achievement, but after making some effort, their scores improved. Furthermore, as a strategy, they may have reviewed the questions and then attempted to solve them by taking the exams again. Liao et al. (2021) discovered a similar behavior pattern in their studies and mentioned them as “item harvesters” who tended to remember, record, and then share items included in the test among their peers. Interestingly, these students caused security concerns within the high-stakes exams. On the other hand, the students could have discovered this pattern for themselves within the formative assessment process. Hui (2023) also determined similar patterns, stating some students exhibited developing patterns for discovering potential correct answers. In the current study, students in Cluster 1 made several attempts, received low scores on their first attempts, and then increased their scores in subsequent attempts. As a result, this may be evidence of a pattern related to formative assessment item harvesting via memorizing test items along with options for further attempts. Importantly, possible reasons for this may be a sense of curiosity, the goal of being more successful, and/or learning by trial-and-error. Hui (2023) stated in their findings that some students with high scores submitted the exams earlier with the aim of having additional time in case help was needed. Again, in the current study, students in the first group started their formative exams earlier than those in the other groups. This may have also indicated that their aim was to identify any misconceptions they may have had and correct those misconceptions within a sufficient amount of time required for the summative assessment. Students in Cluster 1 were also found to be the ones who spent the most amount of time on exam trials in terms of total time spent. Thus, as a result of the high number of trials, it was expected that the total time would increase as well. However, this could also be proof of their effort by considering their correct answers in subsequent attempts. As a result, students in this profile included those who were found to have the highest number of exam attempts. The fact that these students took the exams a considerable amount of time before the summative exam and regularly took the formative exams enabled them to increase their scores throughout this process. According to Hui (2023), trials that do not lead to progress in scoring can be described as trial-and-error, and trials that increase scores over time can be described as effortful improvement. Similarly, students with low scores show more random guessing behavior than those with high scores (Stenlund et al., 2017). In this respect, we can assume that students in Cluster 1 exhibited some form of trial-and-error or random guessing behavior as part of their first trials and then likely worked to show effortful improvement progressively.

Next, Cluster 2 consisted of students who received the highest scores among all the clusters. In the formative exam, their first-attempt scores were higher, yet their subsequent attempt scores were lower. Importantly, the number of trials was low, along with the number of interactions being lower than that of other students. Students who studied and desired to assess themselves prior to the final exam might have been part of this group. Furthermore, according to their behavior, they tended to wait until the last day, just before the final exam. Interestingly, this pattern resembled a similar pattern from Stadler et al. (2020), which demonstrated that higher-ability students spent more time in the problem-solving process but interacted less than others. The fact that these students earned high scores, more recently tested themselves, and showed little interaction indicated having a higher level of ability. Also, we can conclude that students in this group were successful due to more likely knowing what they wanted and not needing to make further attempts and/or spend further time due to their earning high scores on the first attempt. This pattern was also similar to the result of Hui (2023) in which students with high scores did not make any subsequent attempts and ultimately stopped making attempts. Additionally, Hui (2023) explained that these students did not benefit from additional time and, as a result, should not make further attempts during this period due to the exam being too easy for them and, therefore, not needing to carry out additional work. The fact that students in Cluster 2 earned high scores on the formative exam on their first attempt and that their scores decreased in the next application followed a similar pattern. On the other hand, Yang et al. (2022) stated that students who frequently participate in formative assessment following the lesson receive higher scores from the summative exam than those who do not. A similar pattern was also observed in the current study when students from Cluster 2 received high scores for both the formative and summative exams.

In Cluster 3, students who had limited interaction were found to also score lower on formative assessments. Additionally, several students in this subgroup did not complete the final exam. Hui (2023) claims that some students may struggle to understand course material, which causes them to lose interest. This could be what caused the students to disengage as well as preventing them from completing the final exam. Importantly, a risk of dropping out of this course was present for students from this cluster. In this regard, it is important to intervene and assist students who fail to participate and/or receive low scores in their formative assessment, and this group of students should be considered at-risk for drop-out students.

Conclusion, Limitations, and Future Research

In the current study, the formative assessment test-taking behavior of students was analyzed along with investigating the student profiles. Clustering was applied to the metrics collected from formative assignment interactions, and then these were compared with the metrics in terms of the differing student profiles. Importantly, meaningful differences were found between students’ formative assessment test-taking behaviors and summative test scores. Additionally, the research outcomes can enrich the literature by showcasing learners' interactions with the unique formative assessment design of the study and by discussing parallels with distinct test-taking profiles evident within only a limited body of work. Therefore, researchers and practitioners should be able to recognize the profile of Cluster 1 as a behavioral model that necessitates instituting precautions with high-risk assessments or as a trial-and-error behavioral model where interventions within the assessment design have the potential to enhance learning performance. Conversely, the profile of Cluster 2 can be appraised as students’ being self-directed learners capable of overseeing their own learning progress. Thus, to incentivize these learners to attain a higher level of performance, assessment designs featuring progressively more challenging questions for each attempt can be implemented. On the other hand, in the literature, the Cluster 3 profile can be assessed as learners at risk of dropping out and awaiting solutions. As a result, to mitigate drop-out occurrences and increase engagement within the learning process, interventions such as support for countering demotivation, provision of diverse tasks, and adaptation of assessment and content should be employed among students in this profile.

One of the limitations of the current study, in regard to generalizability, was the small sample size. Features of clusters should be compared with research results from large-scale samples, along with also confirming the comparison of summative assessment performance among clusters. Another limitation was related to the features used in clustering. In particular, by deepening the time metrics (response time for each question), clusters with divergent test-taking profiles can be obtained within the formative assessment. In the current study, students were encouraged through the formative test to increase their efforts toward learning. This may be due to the formative assessment being structured in a way that creates equivalent exams for students regarding each subject it covers. However, although it cannot be guaranteed that students take exams of equal difficulty in each attempt, validity and reliability concerns of the formative tests can be mitigated due to the random selection of questions.

Finally, in future studies, students within at-risk groups should be identified, and appropriate interventions should be applied to assist them. As a result, students can behave similarly to the successful students found in other clusters, and the contribution of these interventions can lead to further investigations. Additionally, more advanced metrics can be revealed for formative assessments held in the Moodle learning management system. Thus, due to these metrics, more detailed information can be obtained, especially in regard to the student test-taking strategies found in Clusters 1 and 2.

Declarations

Conflict of Interest: No potential conflict of interest was reported by the authors.

Ethical Approval: The data usage in this study was approved by the Ankara University Ethics Committee, Social Science Sub Committee (Document number:160, dated 17.05.2021).

References

Blondeel, E., Everaert, P., & Opdecam, E. (2023). Does practice make perfect? The effect of online formative assessments on students' self-efficacy and test anxiety. *The British Accounting Review* (101189). https://doi.org/10.1016/j.bar.2023.101189

Cassady, J. C., & Gridley, B. E. (2005). The Effects of Online Formative and Summative Assessment on Test Anxiety and Performance. *The Journal of Technology, Learning and Assessment, 4*(1). Retrieved from https://ejournals.bc.edu/index.php/jtla/article/view/1648

Crisp, V., & Ward, C. (2008). The development of a formative scenario-based computer assisted assessment tool in psychology for teachers: The PePCAA project. *Computers & Education, 50*(4), 1509-1526. https://doi.org/10.1016/j.compedu.2007.02.004

Gašević, D., Greiff, S., & Shaffer, D. W. (2022). Towards strengthening links between learning analytics and assessment: Challenges and potentials of a promising new bond. *Computers in Human Behavior, 134*(2022), 107304. https://doi.org/10.1016/j.chb.2022.107304

Guo, H. (2021). How Did Students Engage with a Remote Educational Assessment? A Case Study. *Educational Measurement: Issues and Practice, 41*(3), 58-68. https://doi.org/10.1111/emip.12476

Guo, H., & Ercikan, K. (2021). Comparing Test-Taking Behaviors of English Language Learners (ELLs) to Non-ELL Students: Use of Response Time in Measurement Comparability Research. *ETS Research Report Series, 2021*(1), 1-15. https://doi.org/10.1002/ets2.12340

Han, S., & Kang, H.-A. (2021). Sequential Monitoring of Aberrant Test-Taking Behaviors Based on Response Times. In M. Wiberg, D. Molenaar, J. González, U. Böckenholt, & J.-S. Kim, *Quantitative Psychology* Cham. https://doi.org/10.1007/978-3-030-74772-5_7

Harlen, W., & James, M. (1997). Assessment and Learning: differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy & Practice, 4*(3), 365-379. https://doi.org/10.1080/0969594970040304

Hui, B. (2023). *Are They Learning or Guessing? Investigating Trial-and-Error Behavior with Limited Test Attempts*. Paper presented at the LAK23: 13th International Learning Analytics and Knowledge Conference, Arlington, TX, USA. https://doi.org/10.1145/3576050.3576068

Lee, Y. H., & Haberman, S. J. (2016). Investigating Test-Taking Behaviors Using Timing and Process Data. *International Journal of Testing, 16*(3), 240-267. https://doi.org/10.1080/15305058.2015.1085385

Liao, M., Patton, J., Yan, R., & Jiao, H. (2021). Mining Process Data to Detect Aberrant Test Takers. *Measurement: Interdisciplinary Research and Perspectives, 19*(2), 93-105. https://doi.org/10.1080/15366367.2020.1827203

Lundgren, E., & Eklöf, H. (2020). Within-item response processes as indicators of test-taking effort and motivation. *Educational Research and Evaluation, 26*(5-6), 275-301. https://doi.org/10.1080/13803611.2021.1963940

Man, K., Harring, J. R., Ouyang, Y., & Thomas, S. L. (2018). Response Time Based Nonparametric Kullback-Leibler Divergence Measure for Detecting Aberrant Test-Taking Behavior. *International Journal of Testing, 18*(2), 155-177. https://doi.org/10.1080/15305058.2018.1429446

Palmen, L. N., Vorstenbosch, M. A. T. M., Tanck, E., & Kooloos, J. G. M. (2015). What is more effective: a daily or a weekly formative test? *Perspectives on Medical Education, 4*(2), 73-78. https://doi.org/10.1007/s40037-015-0178-8

Pools, E., & Monseur, C. (2021). Student test-taking effort in low-stakes assessments: evidence from the English version of the PISA 2015 science test. *Large-scale Assessments in Education, 9*(1), 10. https://doi.org/10.1186/s40536-021-00104-6

Rakoczy, K., Pinger, P., Hochweber, J., Klieme, E., Schütze, B., & Besser, M. (2019). Formative assessment in mathematics: Mediated by feedback's perceived usefulness and students' self-efficacy. *Learning and Instruction, 60*(2019), 154-165. https://doi.org/10.1016/j.learninstruc.2018.01.004

Rienties, B., & Toetenel, L. (2016). The impact of learning design on student behaviour, satisfaction and performance: A cross-institutional comparison across 151 modules. *Computers in Human Behavior, 60*(2016), 333–341. https://doi.org/10.1016/j.chb.2016.02.074

Sarac, M., & Loken, E. (2023). Examining patterns of omitted responses in a large-scale English language proficiency test. *International Journal of Testing*, *23*(1), 56-72. https://doi.org/10.1080/15305058.2022.2070756

Shin, J., Chen, F., Lu, C., & Bulut, O. (2022). Analyzing students' performance in computerized formative assessments to optimize teachers' test administration decisions using deep learning frameworks. *Journal of Computers in Education*, *9*(1), 71-91. https://doi.org/10.1007/s40692-021-00196-7

Silm, G., Must, O., & Täht, K. (2013). Test-taking effort as a predictor of performance in low-stakes tests. *Trames. Journal of the Humanities and Social Sciences*, *17(*67/62), 433-448. https://doi.org/10.3176/tr.2013.4.08

Silm, G., Pedaste, M., & Täht, K. (2020). The relationship between performance and test-taking effort when measured with self-report or time-based instruments: A meta-analytic review. *Educational Research Review*, *31*(100335), 1-22. https://doi.org/10.1016/j.edurev.2020.100335

Stadler, M., Hofer, S., & Greiff, S. (2020). First among equals: Log data indicates ability differences despite equal scores. *Computers in Human Behavior*, *111*(2020), 106442. https://doi.org/10.1016/j.chb.2020.106442

Stenlund, T., Eklof, H., & Lyren, P. E. (2017). Group differences in test-taking behaviour: an example from a high-stakes testing program. *Assessment in Education-Principles Policy & Practice, 24*(1), 4-20. https://doi.org/10.1080/0969594x.2016.1142935

Stenlund, T., Lyren, P. E., & Eklof, H. (2018). The successful test taker: exploring test-taking behavior profiles through cluster analysis. *European Journal of Psychology of Education*, *33*(2), 403-417. https://doi.org/10.1007/s10212-017-0332-2

Tempelaar, D., Rienties, B., Mittelmeier, J., & Nguyen, Q. (2018). Student profiling in a dispositional learning analytics application using formative assessment. *Computers in Human Behavior*, *78*(2018), 408-420. https://doi.org/10.1016/j.chb.2017.08.010

Toetenel, L., & Rienties, B. (2016). Learning Design – creative design to visualise learning activities. *Open Learning: The Journal of Open, Distance and e-Learning*, *31*(3), 233-244. https://doi.org/10.1080/02680513.2016.1213626

Vonderwell, S., Liang, X., & Alderman, K. (2007). Asynchronous Discussions and Assessment in Online Learning. *Journal of Research on Technology in Education, 39*(3), 309-328. https://doi.org/10.1080/15391523.2007.10782485

Wise, S. L. (2006). An Investigation of the Differential Effort Received by Items on a Low-Stakes Computer-Based Test. *Applied Measurement in Education*, *19*(2), 95-114. https://doi.org/10.1207/s15324818ame1902_2

Wise, S. L., & DeMars, C. E. (2006). An Application of Item Response Time: The Effort-Moderated IRT Model. *Journal of Educational Measurement*, *43*(1), 19-38. https://doi.org/https://doi.org/10.1111/j.1745-3984.2006.00002.x

Wise, S. L., Ma, L., Cronin, J., & Theaker, R. A. (2013). Student test-taking effort and the assessment of student growth in evaluating teacher effectiveness. *Annual conference of the American Educational Research Association*, San Francisco, CA,

Wise, S. L., & Gao, L. (2017). A General Approach to Measuring Test-Taking Effort on Computer-Based Tests. *Applied Measurement in Education*, *30*(4), 343-354. https://doi.org/10.1080/08957347.2017.1353992

Wolsey, T. (2008). Efficacy of Instructor Feedback on Written Work in an Online Program. *EdMedia + Innovate Learning Online 2022*, *7*(2), 311-329.

Xiong, Y., & Suen, H. K. (2018). Assessment approaches in massive open online courses: Possibilities, challenges, and future directions. *International Review of Education, 64*(2), 241-263. https://doi.org/10.1007/s11159-018-9710-5

Yang, A. C. M., Chen, I. Y. L., Flanagan, B., & Ogata, H. (2022). How students' self-assessment behavior affects their online learning performance. *Computers and Education: Artificial Intelligence, 3*(2022), 100058. https://doi.org/10.1016/j.caeai.2022.100058

Yildirim-Erbasli, S. N., & Bulut, O. (2020). The impact of students' test-taking effort on growth estimates in low-stakes educational assessments. *Educational Research and Evaluation, 26*(7-8), 368-386. https://doi.org/10.1080/13803611.2021.1977152

Yildirim-Erbasli, S. N., & Bulut, O. (2022). Designing Predictive Models for Early Prediction of Students' Test-taking Engagement in Computerized Formative Assessments. *Journal Of Applied Testing Technology, 22*(2), 1-14. Retrieved from https://jattjournal.net/index.php/atp/article/view/167548

ISSN: 1309 – 6575
Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi
Journal of Measurement and Evaluation in Education and Psychology
Research Article; 2023; 14(Special Issue); 332-344

Design and Development of an Interactive Video Player for Supporting Formative Assessment in Online Learning

Alper BAYAZIT* Gökhan AKÇAPINAR**

Abstract

In this study, the aim was to design a video player with embedded quizzes to enhance students' interactions with video learning materials and enable them to derive maximum benefit from these materials. The developed tool is integrated into the Moodle learning management system and presents questions from a question bank to students at predetermined time intervals set by the instructor. Additionally, it records interactions between the student and embedded quizzes, such as the number of attempts and the number of correct or incorrect answers. Furthermore, students' video interactions, encompassing actions like play, pause, and seek, are also stored in the database. The tool allows instructors to export 15 features related to students' videos and embedded quiz interactions. Consequently, students can assess their comprehension of the content they are viewing and receive immediate feedback, while instructors can access summary reports for all students. This enables them to find out how many students have watched the videos and their responses to the questions before or after the class. This paper explains the development process of the tool and presents findings from a pilot study utilizing the tool. The results of the pilot study revealed that students' video-watching behaviors in interactive videos differed from those in non-interactive videos. An interactive video player that allows embedding quiz questions to support formative assessment in online learning environments, can be advantageous for researchers, instructors, and learners.

Keywords: video analytics, in-video quizzes, interactive videos, video-based learning, online learning

Introduction

Due to the prevalence of mobile devices, the increasing number of Internet users, and the frequent utilization of online learning, video technology in education has advanced faster than ever (Sablić et al., 2020). According to Lacey and Wall (2020), utilizing video-based learning materials can improve student satisfaction (Üstün, 2023), student learning (Yoon et al., 2021), provide helpful feedback (Dohms et al., 2020), and increase student engagement (Mohammadhassan & Mitrovic, 2022). Additionally, using these materials can prepare students for real-world examinations (Lacey & Wall, 2020; Weeks & Horan, 2013), reduce test anxiety (Tripodi, 2018), enhance performance (Weeks & Horan, 2013), and lead to successful learning outcomes (Zaneldin et al., 2019). The importance of video materials is increasing for both students and teachers in flipped classrooms (Bakla & Mehdiyev, 2022; Rose et al., 2016; Xiu et al., 2018), hybrid, or distance learning environments (Barut Tugtekin & Dursun, 2022).

Crook et al. (2012) reported that students who only engage with video materials often need help receiving feedback. Furthermore, Montayre and Sparks (2018) have also found that they complain about the lack of interaction. Therefore, the students must understand if they have achieved the course objectives and need tools to test their knowledge. Furthermore, instructors need to know their students' level of interaction and progress to plan learning activities effectively. In-video quizzes offer an engaging and interactive way to learn new information (Cummins et al., 2016). Integrating formative assessment components into video materials can help learners improve their learning scores (Rice et al., 2019). In a way that is challenging in a traditional classroom environment, receiving immediate and individual feedback (Mirriahi et al., 2021) can make hybrid or online learning more effective.

* Asst. Prof. Dr., Ankara University, Faculty of Medicine, Ankara-Turkiye, e-mail: abayazit@ankara.edu.tr, ORCID ID: 0000-0003-4369-587X
** Assoc. Prof. Dr., Hacettepe University, Faculty of Education, Ankara- Turkiye, e-mail: gokhana@hacettepe.edu.tr, ORCID ID: 0000-0002-0742-1612

To cite this article:
Bayazıt, A., & Akçapınar, G. (2023). Design and Development of an Interactive Video Player for Supporting Formative Assessment in Online Learning, *14*(Special Issue), 332-344. https://doi.org/10.21031/epod.1286077
Received: 20.04.2023
Accepted: 10.09.2023

The instructors can use interactive in-quiz videos and integrate video interaction data to track students' progress in video-based learning. Researchers use video analytics to measure students' viewing behaviors, including the frequency of seeking, playing, and pausing (Yürüm et al., 2022). This data is analyzed to understand their viewing strategies (Akçapınar & Bayazıt, 2018; Yoon et al., 2021) and engagement with the course materials (Ifenthaler et al., 2023; Mohammadhassan & Mitrovic, 2022). Students may alter their approach to watching educational videos depending on their goals. As a result, their viewing strategies may vary (Yoon et al., 2021). Video analytics metrics, such as play, pause, seek, etc., can visually represent quantitative interaction data for a specific video timeline. Visual cues can help instructors improve their video material by identifying areas where students must review missed knowledge or replay a section for clarity (Kim et al., 2014). Furthermore, utilizing video analytics metrics can help develop a predictive model that identifies students at risk of underperforming (Mubarak et al., 2020).

In summary, educational videos are commonly used in education because they facilitate visual learning (Chen, 2020; Mohammadhassan & Mitrovic, 2022), improve comprehension (Coakley et al., 2020), and promote memory retention (Seo et al., 2021). Incorporating video-based learning into students' academic experiences can be beneficial, as they serve as multimedia learning tools (Park, 2022) and support asynchronous education (Choe et al., 2019). Students prefer video materials because they connect theoretical knowledge and practical application (Dohms et al., 2020; Evi-Colombo et al., 2022) while fostering essential skills like problem-solving and reflection (Liu et al., 2021), critical thinking, and reasoning (Chen, 2020; Gartmeier et al., 2019). By supporting various designs, they can be utilized to educate individuals in different fields. An example is the diverse learning objectives and content, including lecture lessons, hands-on programming tasks (Atapattu & Falkner, 2018), surgery training with a head camera and narration (Ahmet et al., 2018), and presentation recordings (Chorianopoulos, 2018). To support active learning, videos should be designed to be more effective and interactive. Adding questions to videos can help students evaluate their readiness, track their progress, and identify misconceptions. Furthermore, analyzing student interactions through video analytics can offer comprehensive insights into the efficacy of video designs and the learning methods utilized by students. Therefore, this study aims to design and develop an interactive video player for online courses. The rest of the paper explains the development process of the tool and presents findings from a pilot study utilizing the tool.

Literature Review

Formative Assessment and Learning Analytics

Formative assessment is an ongoing, dynamic, and informal evaluation process used in education to track and evaluate students' learning progress (Bell & Cowie, 2001). It provides feedback for students to understand their progress and improve, while also giving teachers insights into the teaching process (Nicol & Macfarlane-Dick, 2004). Therefore, summative assessments are called "assessments of learning" while formative assessments are called "assessments for learning" (Kulasegaram & Rangachari, 2018). Formative assessment can take several forms, including quizzes, discussions, surveys, peer assessments, or self-assessments.

Learning Analytics (LA) enhances learning by analyzing data during education and follows four stages (Clow, 2012). Learners improve their knowledge and skills by engaging in diverse applications across multiple environments. Various sources provide data, such as LMS, online platforms, assessments, digital resources, and student demographics. Machine learning algorithms and statistical methods analyze the data to identify patterns, trends, and relationships (Üstün et al., 2022). Personalized learning and identifying at-risk students are crucial to maintaining student success.

LA can enhance the potentialities of formative assessment, such as by creating materials with automatic formative assessment and creating collaborative activities for students (Barana et al., 2019). LA can enhance the eliciting of students' conceptions, and for investigating these strategies, formative assessment tasks can be used (Stanja et al., 2023). An early warning system can be designed for at-risk students, or personalized assignments based on student errors can be given (Rodríguez-Martínez et al., 2023).

Video Analytics Tools

Video analytics give insight into the teaching and learning process (Giannakos et al., 2015). Student engagement, interest, and behavior data can be measured through metrics such as the number of video view, play and seek (Bakla & Mehdiyev, 2022; Brinton et al., 2016; Seo et al., 2021). Using video timeline data, students get personalized content recommendations based on their interactions and preferences (Belarbi et al., 2019). Video analytics also offers automated assessment and feedback (Chatti et al., 2016). By analyzing video interactions, instructors can provide individualized feedback to students based on their challenges, assess the effectiveness of their videos, or understand the learning gaps.

Previous studies investigated the students' video interactions to understand the students' learning strategies, learner profiles, engagement with the materials, active learning processes, and performance predictions. For example, Liao and Wu (2023) analyzed the video log data of 47 graduate students. Four learner profiles were examined based on their learning improvements: the Advanced, the Diligent, the Indifferent, and the Persistent. Students who were diligent and persistent took frequent self-paced breaks to take notes. Advanced students had a high seeking backward frequency, and indifferent students had the highest ratio for seeking backwards, indicating distraction problems. As a result, video-based interactions are a reliable measure of learning motivation. Similarly, Yoon et al. (2021) analyzed student profiles during video-based online learning. Four behavior patterns were found: browsing, socializing, seeking info, and environment setup. These led to two clusters. Active learners, the first cluster, are frequently engaged in social interaction, information seeking, and environment configuration. The second cluster, passive learners, mainly browsed. The study found that active learners achieved higher learning outcomes than passive learners. Zamzuri (2022) analyzed YouTube videos based on student interaction metrics in another study. On average, students only watched one-third of the total video duration. However, they paid more attention during the demonstration sessions, primarily when the whiteboard was used and visualizations were included in the materials. Video analytics can assist in identifying the segments of videos that students tend to watch more frequently. Consequently, this information can be used to make inferences about video design. Yürüm et al. (2022) conducted a quasi-experimental study to examine how interactive videos affect students' engagement and satisfaction. After controlling for motivation, they discovered students were significantly more satisfied with interactive videos. Additionally, students were less likely to revisit essential points and more likely to skip unimportant ones in interactive videos. Their findings suggest that interactive video lectures can convert videos into interactive ones. This research indicates that including formative assessment features in videos can enhance their effectiveness. Combining self-reports and video analytics metrics like play, pause, seek, and question interaction could provide a more comprehensive outcome. In a quasi-experimental study, Mohammadhassan and Mitrovic (2022) examined data from a platform incorporating note-taking, peer-reviewing, and personalized prompts. They compared levels of engagement, such as the number of videos watched and comments, as well as learning gains. The results indicated that visual LA in videos encouraged constructive behavior and enhanced learning outcomes. Various factors, such as time and purpose, can influence video viewing strategies. For example, Seo et al. (2021) showed that students tend to perform more searches within the video during exam weeks and when rewatching videos. After analyzing video analytics tools utilized in previous research, we summarized their advantages and disadvantages for students and instructors in Table 1.

Table 1.

Video platforms used in educational video analytics studies

Tool	Metrics	LMS/MOOC integration	Video Markers	Assessment Features	Tracking Scope
Vimeo	Views, finishes, time watch	Embed	Yes	multiple-choice	Video
Panopto	Play, pause, seek,	LTI	No	multiple-choice	User
Youtube	Views, duration	Embed	No	No	Video
SocialSkip	Play, pause, seek,	Not supported	No	multiple-choice, true/false	User

Vimeo provides paid options with advanced analytics and the ability to include quiz questions. The system cannot directly match with LMS users, only through embedding. This restriction makes it difficult to conduct further investigations. Panopto provides valuable insights into video engagement for businesses and can be integrated with Moodle using the LTI plugin. While it has limitations in matching student IDs and numbers, it does allow for quizzes. However, more information is needed on adding markers to question areas. When analyzing student interactions, YouTube videos only provide video-focused reports and visualizations. Adding quiz questions through Google Classroom can help, but there needs to be a marker in the sections linked to the questions. Integrating with an LMS is possible through embedding, but collecting and analyzing student-based data has challenges. SocialSkip is a helpful plug-in that enables the collection of student interaction data. SocialSkip offers free access to various question types but must be fully integrated with Moodle.

Various educational video analytics tools support collecting metrics such as play, pause, and seek. These tools are typically commercial, with different prices depending on the number of users and videos. Some tools need to be able to connect to Moodle and require separate memberships for data collection. Adding a formative assessment item to a different system while keeping it in a pool in the existing question bank is impractical. Integrating LMS interaction data and video analytics data in the future may be complex due to incomplete user account integration.

Method

This is a tool development study, so from this perspective, details of the tool development process are reported in the method section. The purpose of the developed tool is to provide students with formative assessment during their video-watching sessions so that students can assess their comprehension of the content they are viewing and receive immediate feedback, while instructors can access summary reports for all students.

Components of the Tool

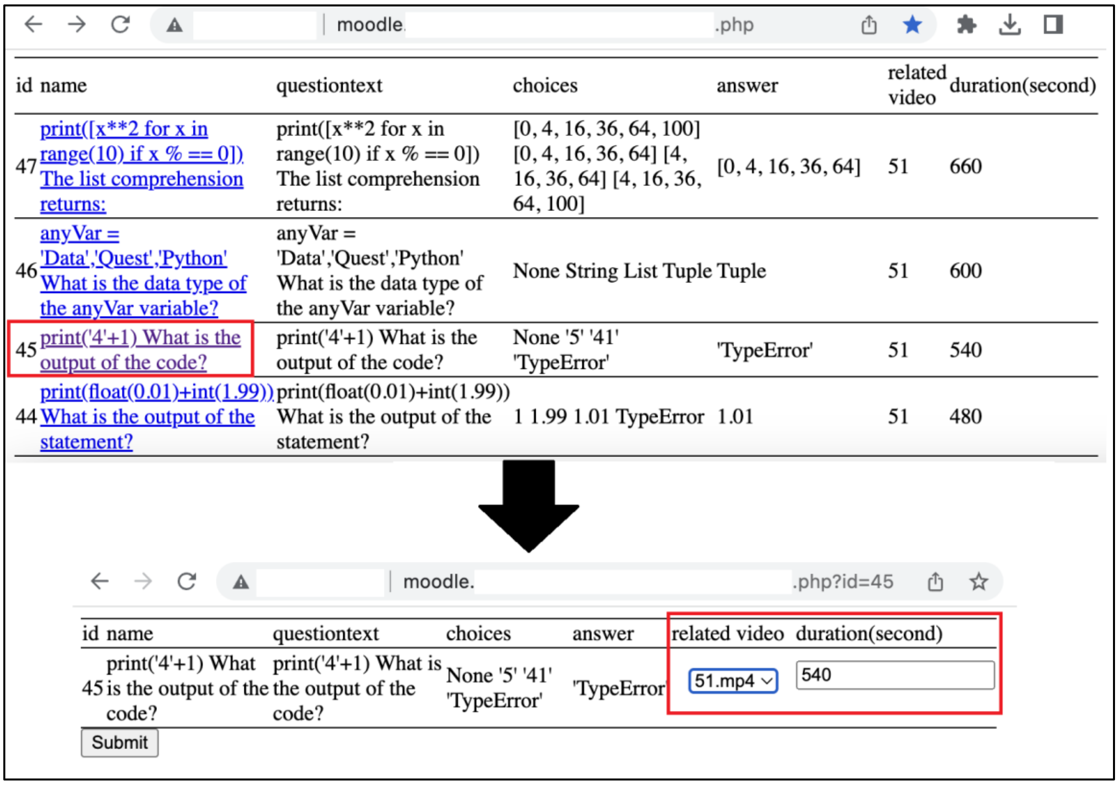
In this study, we designed and developed an interactive video player for supporting formative assessment in the Moodle environment. Moodle is one of the most widely used LMS for blended learning (Üstün et al., 2021; Ustun & Tracey, 2021), distance education, flipped classrooms (Üstün et al., 2022), or other online learning purposes in schools, universities, and other sectors.

The tool's video player component was created using JavaScript and HTML5. Using MySQL queries, interaction data is written to the database. The PHP programming language facilitates data transfer between the server database, assessment items, and video player. Multiple-choice questions in the Moodle question bank can be added using the tool at any point in the video timeline. Markers indicating specific questions can be added to the video timeline, specified in seconds. The student can view each question by directly clicking on a marker. The video pauses at the question marker on the timeline.

Questions associated with the Moodle question bank are displayed with options on the screen below the video (Figure 1). Thus, students can answer the formative assessment items and receive immediate feedback such as "correct answer, congratulations" or "wrong answer, try again". Students can answer the question repeatedly until they provide the correct response. The students' attempts to answer the questions are saved as a log, including interactions such as video play, pause, or seek. Instructors can view all student interactions and their performances in video-based assessments. Teachers and students must sign into the Moodle LMS to create interactive video materials. The multiple-choice questions in the Moodle question bank are listed on a page, including question texts, choices, correct answers, related video material names, and question display time. Teachers can edit the linked videos and the times when the questions are shown by clicking the question name.

Figure 1.

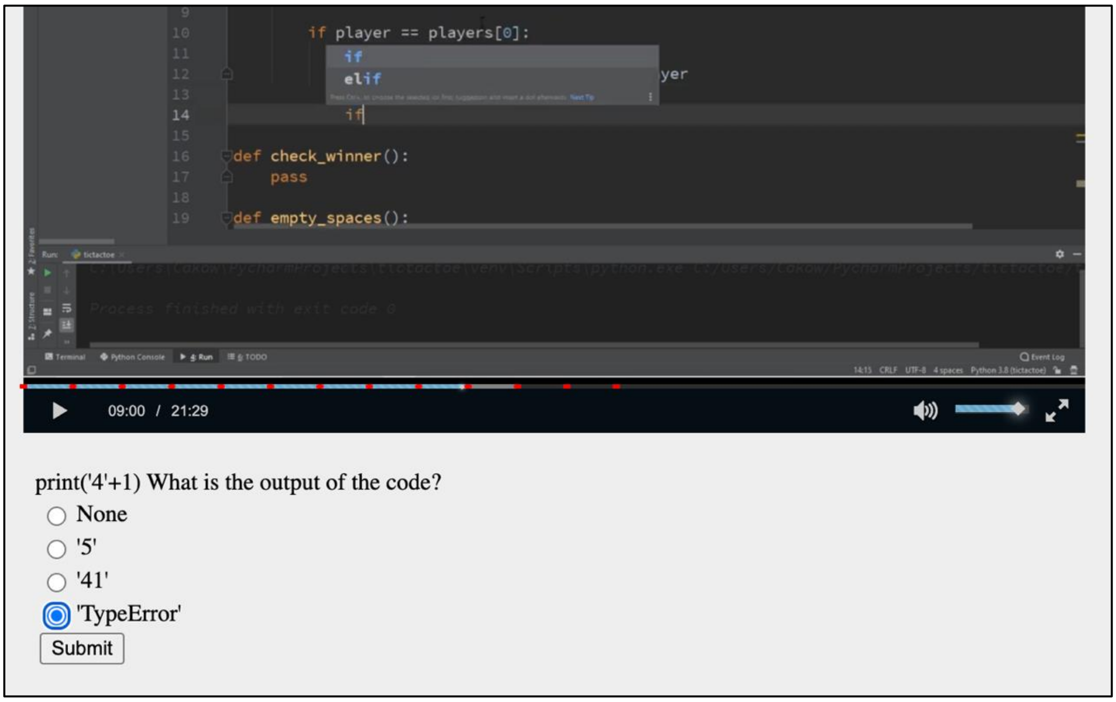
Interactive video player question bank interface



The interactive video player is embedded in the student interface. The tool generates a random ID number for the files submitted to the server to prevent students from accessing video materials through external links. Once students log in to the Moodle LMS, they can access the video link through the weekly course view. In the tables in the Moodle database, video file information, user information, and session information are recorded during the video viewing process. Teachers can also export students' item interactions and video-watching behaviors from the Moodle database as log files.

Figure 2.

Interactive video player student interface



The Interactive Video Player interface consists of three main parts; (1) the video player section where the video player and markers are located, (2) the section where the formative assessment items associated with the video markers are located, (3) the instant feedback area based on the selected options (Figure 2). The video player section allows users to watch videos, view the video duration and the current video time, and use video controls such as play, pause, adjust volume and screen size. In addition, it enables the students to seek forward or backwards while watching a video or to view the assessment items by directly jumping to the question time (by clicking on the marker). The question display area shows the student the items in the question pool at video time. In the feedback area, the answers to the questions are given with the knowledge of whether they are right or wrong. Based on the feedback, the student can either press the "play" button to keep watching the video or seek over the timeline to try to clear up any misunderstandings.

In summary, the tool that is integrated works with Moodle LMS and makes sure that the assessment items in the question bank show up at the correct times on the timeline in the student video-watching tasks. So students can receive instant feedback and increase their interaction. The teachers can investigate students' interactions with videos and their responses.

Interaction Data

The Interactive Video Player records students' formative assessment and interaction behaviors with videos in a new table created in the Moodle database by adding a single-line record for each behavior. These log records include the Moodle user ID, date and time stamp, interaction type, detailed information about the interaction, and video ID. Each interaction performed represents one row in the related table. In the interaction detail, information about question attempts and video interactions is provided (Figure 3).

Figure 3.
Interactive video player log data

id	userid	sessionid	time	event	detail	videoid
50117	404	b6a0fb2d00f91b826c4aa422d9ba77ef	1673087674	8	298 sec. video played.	51
50118	404	b6a0fb2d00f91b826c4aa422d9ba77ef	1673087675	11	300 sec. marker time, video paused	51
50119	404	b6a0fb2d00f91b826c4aa422d9ba77ef	1673087676	9	300 sec. video paused.	51
50120	394	32e74781bb5b4419be4d16477ba4e1ef	1673087730	8	72 sec. video played.	51
50121	394	32e74781bb5b4419be4d16477ba4e1ef	1673087781	9	121 sec. video paused.	51
50122	394	32e74781bb5b4419be4d16477ba4e1ef	1673087813	16	121 sec., 51 video ID, 38 question ID wrong answer	51
50124	394	32e74781bb5b4419be4d16477ba4e1ef	1673087829	15	121 sec., 51 video ID, 38 question ID correct answer	51
50125	394	32e74781bb5b4419be4d16477ba4e1ef	1673087831	8	121 sec. video played.	51
50129	394	32e74781bb5b4419be4d16477ba4e1ef	1673087891	9	180 sec. video paused.	51
50131	394	32e74781bb5b4419be4d16477ba4e1ef	1673087950	16	180 sec., 51 video ID, 39 question ID wrong answer	51
50133	394	32e74781bb5b4419be4d16477ba4e1ef	1673087956	16	180 sec., 51 video ID, 39 question ID wrong answer	51
50136	394	32e74781bb5b4419be4d16477ba4e1ef	1673087970	16	180 sec., 51 video ID, 39 question ID wrong answer	51
50137	394	32e74781bb5b4419be4d16477ba4e1ef	1673087973	15	180 sec., 51 video ID, 39 question ID correct answer	51
50138	394	32e74781bb5b4419be4d16477ba4e1ef	1673087998	8	180 sec. video played.	51
50139	394	32e74781bb5b4419be4d16477ba4e1ef	1673088001	9	182 sec. video paused.	51
50140	394	32e74781bb5b4419be4d16477ba4e1ef	1673088006	15	182 sec., 51 video ID, 39 question ID correct answer	51
50141	394	32e74781bb5b4419be4d16477ba4e1ef	1673088009	8	182 sec. video played.	51
50142	394	32e74781bb5b4419be4d16477ba4e1ef	1673088015	9	188 sec. video paused.	51
50144	404	b6a0fb2d00f91b826c4aa422d9ba77ef	1673088098	8	300 sec. video played.	51
50145	404	b6a0fb2d00f91b826c4aa422d9ba77ef	1673088100	11	301 sec. marker time, video paused	51
50146	404	b6a0fb2d00f91b826c4aa422d9ba77ef	1673088101	9	301 sec. video paused.	51

The raw dataset must be processed to identify the students' strategies during their interactions with the video tasks and assessment items. For this purpose, a data preprocessing tool was developed for the raw data exported. During the data pre-processing, first, the duplicate records were removed. Then, the tool transforms the interactions into quantitative features including the number of video views of the students, video watching behaviors shown in different sessions, video timeline interactions such as play-pause, seek, or marker clicking, number of wrong answers, number of correct answers, and number of question item views. As a result of the preprocessing, the following metrics were discovered regarding the students' interactive video player behaviors:

n_Session: Total number of sessions for watching videos.

n_QuestionMarkerAutoPause: Number of automated pauses by reaching the formative assessment item at a time on the video timeline.

n_QuestionMarkerClick: Number of question views by clicking the marker.

n_QuestionAnswerCorrect: Number of correct answers given to the formative assessment item on the video timeline.

n_QuestionAnswerWrong: Number of wrong answers given to the formative assessment item on the video timeline

n_TotalAction: Total number of interactions (play, pause, seek, etc.) performed on the video.

d_Time: Total time spent watching videos.

n_diffVideoCount: Number of unique video views.

n_VideoLoad: Total number of views of the video page.

n_VideoPlay: Total number of plays obtained by pressing the play button of the relevant video or clicking on a paused video.

n_VideoPause: Total number of pauses obtained by pressing the pause button of the relevant video or clicking on the playing video.

n_VideoSeek: Total number of seeks forward or backward on the video timeline.

n_ForwardSeek: Total number of seek forward behavior on the video timeline.

n_BackwardSeek: Total number of seek backward behavior on the video timeline.

Pilot Study

The pilot study was conducted with second-year medical school students in an elective introductory programming course. The students were given a homework assignment that required them to watch videos and submit their code files. They were given a week to complete each assignment. Students were asked to watch two videos via the developed video player. One with embedded formative assessment questions and the other without interactive features (standard video). The students were first asked to watch a standard video and complete the task of developing a calculator application using Python. The following week, they were asked to watch the interactive video and develop a simple game. Formative assessment items were added to the video timeline for the second task. The game development video has 12 multiple-choice formative assessment questions on the video timeline. The students (N=12) who participated in this pilot study provided 3502 lines of interaction data. Descriptive statistics of the interaction data related to interactive video are given as an example in Table 2.

Table 2.

Sample descriptive statistics of student interaction in the interactive video

Interaction Metrics	Mean	Median	SD
n_Session	1.58	1	.79
n_QuestionMarkerAutoPause	3.42	2.00	4.14
n_QuestionMarkerClick	4.42	2.00	5.90
n_QuestionAnswerCorrect	20.08	20.00	6.42
n_QuestionAnswerWrong	7.25	7.50	4.16
n_TotalAction	123.58	111.50	41.77
d_Time	53.00	50.00	16.17
n_diffVideoCount	1.00	1.00	.00
n_VideoLoad	2.58	2.50	1.17
n_VideoPlay	40.92	31.00	23.43
n_VideoPause	40.67	31.50	23.78
n_VideoSeek	3.25	2.00	4.64
n_ForwardSeek	1.83	.00	4.30
n_BackwardSeek	1.17	.50	1.40

Students were given a single video; therefore, a different video count metric is one (*diffVideoCount*). Students typically finish the task in one or two sessions (*n_Session*). Regarding formative assessment items, it may be claimed that students engage with them frequently (*n_QuestionAnswerCorrect*, *n_QuestionAnswerWrong*). While some students see the questions immediately by clicking the marker (*n_QuestionMarkerClick*), others wait until the video pauses automatically (*n_QuestionMarkerAutoPause*). In addition, students showed more playing and pausing (*n_VideoPlay*, *n_VideoPause*) behavior than seeking forward or backward (*n_ForwardSeek*, *n_BackwardSeek*). Students tend to watch coding video tasks linearly because these tasks often use the demonstration strategy.

The watching behaviors of the paired students (N=9) in the two videos were compared in terms of the main metrics (play, pause, seek). As a result of the Wilcoxon Signed-Rank test (see Table 3), there is a significant difference in the seeking backward variable from the main metrics, such as video play, pause, and seeking forward and backward. The number of seeking backwards in the interactive video task is greater than the number of seeking backwards in the non-interactive videos ($z = -2.88$, $p = .02$). As a result; the Interactive Video Player provides essential information about the monitoring processes for students' videos containing formative assessment. Analyzing the data makes it possible to identify students' video-watching behaviors, make comparisons, and determine their strategies.

Table 3.
Comparison of interaction metrics in interactive and non-interactive video players

Metrics	Non-Interactive Video Player		Interactive Video Player			
	Mean Rank	Sum of Ranks	Mean Rank	Sum of Ranks	z	p
n_VideoPlay	3.25	13.00	6.40	32.00	-1.13	.26
n_VideoPause	3.50	14.00	6.20	31.00	-1.00	.31
n_ForwardSeek	4.83	14.50	5.08	30.50	-.95	.34
n_BackwardSeek	1.50	1.50	4.93	34.50	-2.32	.02

Discussion

The objective of researchers is to enhance the interactivity of classroom video materials to engage students better and facilitate their learning. (Fatima et al., 2019; Seo et al., 2021; Yürüm et al., 2022). Including video assessment segments can enhance the learning experience and provide valuable feedback (Mirriahi et al., 2021). Supporting these processes with video analytics can help develop individualized learning environments, identify students' learning strategies (Akçapınar & Bayazıt, 2018), and experience a more effective learning process. Within the scope of this study, a tool has been designed and developed to track the students' question-answering and video-watching behaviors in the video-based learning tasks. Furthermore, this tool is integrated with Moodle LMS question pools and user information. Using the developed script, it is now possible to incorporate formative assessment items from the Moodle LMS question pool into the video timeline. Educators and researchers can also take advantage of this tool.

The tool that has been developed comprises various interaction metrics like play, pause, number of views, and seeks. These metrics are similar to those used to collect video interaction data on Vimeo, Panopto, YouTube, and SocialSkip platforms. It includes formative assessment features like other video analytics tools and superior, unique features that surpass them. First, it contains a video marker and can dynamically add it to the video table for each added question. Thus, different approaches can be revealed while examining video-watching strategies, such as linear watching, progress with direct formative assessment, or watching by seeking forward or backwards. Furthermore, by analyzing the collected metrics, we can evaluate how the strategies are impacted based on whether the questions' answers are accurate. Another note is that this tool was created using PHP programming as an open-source project and seamlessly integrated with Moodle tables. It is possible to incorporate questions from the Moodle question pool into a video uploaded to the Moodle server. This allows for monitoring student interactions and opens avenues for exploring various tasks and evaluation methods Moodle offers in future studies. A significant limitation of the tool is that logging is not possible on the iOS operating system due to its player.

Based on the pilot application results, it was found that student engagement with videos significantly increased when formative features were added to the materials. This indicates that incorporating assessment features can enhance student interaction with educational videos. The findings align with Zamzuri’s (2022) study, which noted that active learners engaged more with videos than passive learners. Therefore, we can infer that incorporating formative assessments and videos can promote active learning. Furthermore, Yürüm et al. (2022) noted that students often re-watch key sections in interactive videos. This may have led the students in the pilot study to consider these sections significant based on the feedback they received, resulting in increased interaction. According to Yoon et al. (2021), students with active learner profiles tend to engage with videos and display the behavior of seeking information. In our study, we observed that students tended to watch interactive videos with greater interest and sought less compared to standard non-interactive videos. This suggests that the interactive features of the videos served as a motivating factor for active learning.

In further studies, we will explore how students approach their studies by analyzing their formative assessments and video interactions. Our goal is to understand how these materials aid the learning process and investigate the impact of video-based feedback on their behavior and learning outcomes.

To achieve these objectives, we plan to integrate our tool with different question types (such as matching and filling in the blanks) in Moodle's question pool. This will provide a broader range of questions for students to engage with. Additionally, we will incorporate reporting features for video usage and formative assessments on Moodle. This will enable students to receive automatic feedback based on their learning objectives and progress compared to their peers.

Declarations

Author Contribution: Bayazıt, A.: conceptualization, investigation, methodology, writing - original draft, data curation, data analysis. Akçapınar, G.: conceptualization, methodology, data analysis, writing - review & editing.

Funding: This study was partially funded by the Scientific and Technological Research Council of Turkey (TÜBİTAK) ARDEB 1002, Grant No: 122K549.

Ethical Approval: Approval was obtained from the ethics committee of Ankara University Medical School (I07-429-22, 04 AGU 2022). The procedures used in this study adhere to the tenets of the Declaration of Helsinki.

Consent to Participate: All authors have given their consent to participate in submitting this manuscript to this journal.

Consent to Publish: Written consent was sought from each author to publish the manuscript.

Competing Interests: The authors have no relevant financial or non-financial interests to disclose.

References

Ahmet, A., Gamze, K., Rustem, M., & Sezen, K. A. (2018). Is Video-Based Education an Effective Method in Surgical Education? A Systematic Review. *J Surg Educ*, *75*(5), 1150-1158. https://doi.org/10.1016/j.jsurg.2018.01.014

Akçapınar, G., & Bayazıt, A. (2018). Investigating Video Viewing Behaviors of Students with Different Learning Approaches Using Video Analytics. *Turkish Online Journal of Distance Education*, *19*(4), 116-125. https://doi.org/10.17718/tojde.471907

Atapattu, T., & Falkner, K. (2018). Impact of Lecturer's Discourse for Student Video Interactions: Video Learning Analytics Case Study of MOOCs. *Journal of Learning Analytics*, *5*(3), 182-197. https://doi.org/10.18608/jla.2018.53.12

Bakla, A., & Mehdiyev, E. (2022). A qualitative study of teacher-created interactive videos versus YouTube videos in flipped learning. *E-Learning and Digital Media*, *19*(5), 495-514. https://doi.org/10.1177/20427530221107789

Barana, A., Conte, A., Fissore, C., Marchisio, M., & Rabellino, S. (2019). Learning analytics to improve formative assessment strategies. *JE-LKS. Journal of e-learning and knowledge society*, *15*(3), 75-88.

Barut Tugtekin, E., & Dursun, O. O. (2022). Effect of animated and interactive video variations on learners' motivation in distance Education. *Education and Information Technologies*, *27*(3), 3247-3276. https://doi.org/10.1007/s10639-021-10735-5

Belarbi, N., Chafiq, N., Talbi, M., Namir, A., & Benlahmar, H. (2019). A Recommender System for Videos Suggestion in a SPOC: A Proposed Personalized Learning Method. Big Data and Smart Digital Environment, Cham.

Bell, B., & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science education*, *85*(5), 536-553.

Brinton, C. G., Buccapatnam, S., Chiang, M., & Poor, H. V. (2016). Mining MOOC Clickstreams: Video-Watching Behavior vs. In-Video Quiz Performance. *IEEE Transactions on Signal Processing*, *64*(14), 3677-3692. https://doi.org/10.1109/Tsp.2016.2546228

Chatti, M. A., Marinov, M., Sabov, O., Laksono, R., Sofyan, Z., Fahmy Yousef, A. M., & Schroeder, U. (2016). Video annotation and analytics in CourseMapper. *Smart Learning Environments*, *3*(1), 10. https://doi.org/10.1186/s40561-016-0035-1

Chen, G. W. (2020). A visual learning analytics (VLA) approach to video-based teacher professional development: Impact on teachers' beliefs, self-efficacy, and classroom talk practice. *Computers & Education*, *144*. https://doi.org/10.1016/j.compedu.2019.103670

Choe, R. C., Scuric, Z., Eshkol, E., Cruser, S., Arndt, A., Cox, R., Toma, S. P., Shapiro, C., Levis-Fitzgerald, M., Barnes, G., & Crosbie, R. H. (2019). Student Satisfaction and Learning Outcomes in Asynchronous Online Lecture Videos. *CBE Life Sci Educ*, *18*(4), ar55. https://doi.org/10.1187/cbe.18-08-0171

Chorianopoulos, K. (2018). A Taxonomy of Asynchronous Instructional Video Styles. *International Review of Research in Open and Distributed Learning*, *19*(1), 294-311. https://doi.org/10.19173/irrodl.v19i1.2920

Clow, D. (2012). The learning analytics cycle: closing the loop effectively. Proceedings of the 2nd international conference on learning analytics and knowledge,

Coakley, A., Bailey, A., Tao, J., Liou, Y. L., Champlain, A., Ander, M., & Lake, E. (2020). Video education to improve clinical skills in the prevention of and response to vasovagal syncopal episodes. *Int J Womens Dermatol*, *6*(3), 186-190. https://doi.org/10.1016/j.ijwd.2020.02.002

Crook, A., Mauchline, A., Maw, S., Lawson, C., Drinkwater, R., Lundqvist, K., Orsmond, P., Gomez, S., & Park, J. (2012). The use of video technology for providing feedback to students: Can it enhance the feedback experience for staff and students? *Computers & Education*, *58*(1), 386-396. https://doi.org/10.1016/j.compedu.2011.08.025

Cummins, S., Beresford, A. R., Davies, I., & Rice, A. (2016, April 25-26, 2016). Supporting Scalable Data Sharing in Online Education. Proceedings of the Third (2016) ACM Conference on Learning@ Scale, Scotland, Edinburgh, UK.

Dohms, M. C., Collares, C. F., & Tiberio, I. C. (2020). Video-based feedback using real consultations for a formative assessment in communication skills. *BMC Med Educ*, *20*(1), 57. https://doi.org/10.1186/s12909-020-1955-6

Evi-Colombo, A., Cattaneo, A., & Bétrancourt, M. (2022). Procedural knowledge acquisition in a second-year nursing course. Effectiveness of a digital video-based collaborative learning-by-design activity using hypervideo. *Journal of Computer Assisted Learning*, *39*(2), 501-516. https://doi.org/10.1111/jcal.12758

Fatima, S. S., Ghias, K., Jabeen, K., & Sabzwari, S. (2019). Enhancing Cognitive Engagement of Pre-clinical Undergraduate Medical Students via Video Cases and Interactive Quizzes in Problem-based Learning. *Cureus*, *11*(1), e3832. https://doi.org/10.7759/cureus.3832

Gartmeier, M., Pfurtscheller, T., Hapfelmeier, A., Grunewald, M., Hausler, J., Seidel, T., & Berberat, P. O. (2019). Teacher questions and student responses in case-based learning: outcomes of a video study in medical education [Article]. *BMC Med Educ*, *19*(1), 455. https://doi.org/10.1186/s12909-019-1895-1

Giannakos, M. N., Chorianopoulos, K., & Chrisochoides, N. (2015). Making Sense of Video Analytics: Lessons Learned from Clickstream Interactions, Attitudes, and Learning Outcome in a Video-Assisted Course. *International Review of Research in Open and Distributed Learning*, *16*(1), 260-283. <Go to ISI>://WOS:000357189200014

Ifenthaler, D., Schumacher, C., & Kuzilek, J. (2023). Investigating students' use of self-assessments in higher education using learning analytics. *Journal of Computer Assisted Learning*, *39*(1), 255-268. https://doi.org/10.1111/jcal.12744

Kim, J., Guo, P. J., Seaton, D. T., Mitros, P., Gajos, K. Z., & Miller, R. C. (2014). *Understanding in-video dropouts and interaction peaks inonline lecture videos* Proceedings of the first ACM conference on Learning @ scale conference, Atlanta, Georgia, USA. https://doi.org/10.1145/2556325.2566237

Kulasegaram, K., & Rangachari, P. K. (2018). Beyond "formative": assessments to enrich student learning. *Advances in Physiology Education*, *42*(1), 5-14. https://doi.org/10.1152/advan.00122.2017

Lacey, K., & Wall, J. G. (2020). Video-based learning to enhance teaching of practical microbiology. *FEMS Microbiology Letters*, *368*(2). https://doi.org/10.1093/femsle/fnaa203

Liao, C. H., & Wu, J. Y. (2023). Learning analytics on video-viewing engagement in a flipped statistics course: Relating external video-viewing patterns to internal motivational dynamics and performance [Article]. *Computers & Education*, *197*, N.PAG-N.PAG. https://doi.org/10.1016/j.compedu.2023.104754

Liu, Z., Yin, H., Cui, W., Xu, B., & Zhang, M. (2021). How to reflect more effectively in online video learning: Balancing processes and outcomes. *British Journal of Educational Technology*, *53*(1), 114-129. https://doi.org/10.1111/bjet.13155

Mirriahi, N., Jovanovic, J., Lim, L. A., & Lodge, J. M. (2021). Two sides of the same coin: video annotations and in-video questions for active learning. *Etr&D-Educational Technology Research and Development*, *69*(5), 2571-2588. https://doi.org/10.1007/s11423-021-10041-4

Mohammadhassan, N., & Mitrovic, A. (2022). *Investigating the Effectiveness of Visual Learning Analytics in Active Video Watching* (Vol. 13355). Springer International Publishing. https://doi.org/10.1007/978-3-031-11644-5_11

Montayre, J., & Sparks, T. (2018). As I haven't seen a T-cell, video-streaming helps: Nursing students' preference towards online learning materials for biosciences. *Collegian*, *25*(5), 487-492. https://doi.org/https://doi.org/10.1016/j.colegn.2017.12.001

Mubarak, A. A., Cao, H., Zhang, W., & Zhang, W. (2020). Visual analytics of video-clickstream data and prediction of learners' performance using deep learning models in MOOCs' courses. *Computer Applications in Engineering Education*, *29*(4), 710-732. https://doi.org/10.1002/cae.22328

Nicol, D., & Macfarlane-Dick, D. (2004). Rethinking formative assessment in HE: a theoretical model and seven principles of good feedback practice. *C. Juwah, D. Macfarlane-Dick, B. Matthew, D. Nicol, D. & Smith, B.(2004) Enhancing student learning though effective formative feedback, York, The Higher Education Academy*.

Park, E. E. (2022). Expanding Reference through Cognitive Theory of Multimedia Learning Videos. *The Journal of Academic Librarianship*, *48*(3), 102522. https://doi.org/https://doi.org/10.1016/j.acalib.2022.102522

Rice, P., Beeson, P., & Blackmore-Wright, J. (2019). Evaluating the Impact of a Quiz Question within an Educational Video. *TechTrends*, *63*(5), 522-532. https://doi.org/10.1007/s11528-019-00374-6

Rodríguez-Martínez, J. A., González-Calero, J. A., del Olmo-Muñoz, J., Arnau, D., & Tirado-Olivares, S. (2023). Building personalised homework from a learning analytics based formative assessment: Effect on fifth-grade students' understanding of fractions. *British Journal of Educational Technology*, *54*(1), 76-97.

Rose, E., Claudius, I., Tabatabai, R., Kearl, L., Behar, S., & Jhun, P. (2016). The Flipped Classroom in Emergency Medicine Using Online Videos with Interpolated Questions. *J Emerg Med*, *51*(3), 284-291 e281. https://doi.org/10.1016/j.jemermed.2016.05.033

Sablić, M., Mirosavljević, A., & Škugor, A. (2020). Video-Based Learning (VBL)—Past, Present and Future: an Overview of the Research Published from 2008 to 2019. *Technology, Knowledge and Learning*, *26*(4), 1061-1077. https://doi.org/10.1007/s10758-020-09455-5

Seo, K., Dodson, S., Harandi, N. M., Roberson, N., Fels, S., & Roll, I. (2021). Active learning with online video: The impact of learning context on engagement. *Computers & Education*, *165*. https://doi.org/10.1016/j.compedu.2021.104132

Stanja, J., Gritz, W., Krugel, J., Hoppe, A., & Dannemann, S. (2023). Formative assessment strategies for students' conceptions—The potential of learning analytics. *British Journal of Educational Technology*, *54*(1), 58-75.

Tripodi, N. (2018). First-year osteopathic students' use and perceptions of complementary video-based learning. *International Journal of Osteopathic Medicine*, *30*, 35-43. https://doi.org/10.1016/j.ijosm.2018.09.004

Üstün, A. B., Karaoğlan-Yılmaz, F. G., & Yılmaz, R. (2021). Investigating the role of accepting learning management system on students' engagement and sense of community in blended learning. *Education and information technologies*, *26*(4), 4751-4769. https://doi.org/10.1007/s10639-021-10500-8

Üstün, A. B., & Tracey, M. W. (2021). An innovative way of designing blended learning through design-based research in higher education. *Turkish Online Journal of Distance Education*, *22*(2), 126-146.

Üstün, A. B., Zhang, K., Karaoğlan-Yılmaz, F. G., & Yılmaz, R. (2022). Learning analytics based feedback and recommendations in flipped classrooms: an experimental study in higher education. *Journal of Research on Technology in Education*, 1-17. https://doi.org/10.1080/15391523.2022.2040401

Üstün, A. B. (2023). Investigating Impacts of Using Mobile Video Lectures on Student Satisfaction and Academic Achievement in Blended Learning. *Malaysian Online Journal of Educational Technology*, *11*(3), 199-210.

Weeks, B. K., & Horan, S. A. (2013). A video-based learning activity is effective for preparing physiotherapy students for practical examinations. *Physiotherapy*, *99*(4), 292-297. https://doi.org/10.1016/j.physio.2013.02.002

Xiu, Y., Moore, M. E., Thompson, P., & French, D. P. (2018). Student Perceptions of Lecture-Capture Video to Facilitate Learning in a Flipped Classroom. *TechTrends*, *63*(4), 369-375. https://doi.org/10.1007/s11528-018-0293-6

Yoon, M., Lee, J., & Jo, I. H. (2021). Video learning analytics: Investigating behavioral patterns and learner clusters in video-based online learning. *Internet and Higher Education*, *50*. https://doi.org/10.1016/j.iheduc.2021.100806

Yürüm, O. R., Taşkaya-Temizel, T., & Yıldırım, S. (2022). The use of video clickstream data to predict university students' test performance: A comprehensive educational data mining approach. *Education and Information Technologies*, 1-32. https://doi.org/10.1007/s10639-022-11403-y

Yürüm, O. R., Yıldırım, S., & Taşkaya-Temizel, T. (2022). An intervention framework for developing interactive video lectures based on video clickstream behavior: a quasi-experimental evaluation. *Interactive Learning Environments*, 1-16. https://doi.org/10.1080/10494820.2022.2042312

Zamzuri, Z. H. (2022). Assessing Students' Focus and Interest Using Youtube Analytics. *Asean Journal of Teaching and Learning in Higher Education*, *14*(1), 47-58. https://doi.org/10.17576/ajtlhe.1401.2022.05

Zaneldin, E., Ahmed, W., & El-Ariss, B. (2019). Video-based e-learning for an undergraduate engineering course. *E-Learning and Digital Media*, *16*(6), 475-496. https://doi.org/10.1177/2042753019870938

ISSN: 1309 – 6575
Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi
Journal of Measurement and Evaluation in Education and Psychology
Research Article; 2023; 14(Special Issue); 345-358

Investigation of the Effect of Online (Web-Based) Formative Assessment Applications on Students' Academic Achievement*

Şeref AKPINAR** Bayram ÇETİN***

Abstract

The purpose of this research is to examine the effects of providing resources for learning disabilities by the system and providing feedback regarding learning disabilities by the teacher within the scope of an online (web-based) formative evaluation application for 10th grade secondary school students in mathematics course quadratic equations on students' success. In the research, a quasi-experimental design was used. Pre-test - post-test achievement tests and follow-up tests were used. The research was conducted in the 2022-2023 academic year with a total of 302 students selected from 4 schools and 12 branches in Göksun and Andırın districts using the stratified, random cluster sampling method. Data were examined by one-way analysis of variance (ANOVA) and analysis of covariance (ANCOVA). According to the research results, it was determined that there was no statistically significant difference between the pre-test averages of the groups, but a statistically significant difference emerged in the post-test. Providing resources for learning disabilities by the system applied to the Experiment-2 group and Cognitive Diagnostic Modeling (CDM) for learning disabilities by the teacher. Providing detailed feedback according to the system, providing resources for learning disabilities by the system applied to the Experiment-1 group, and normal teaching applied to the Control group; It was determined that the system applied to the Experiment-1 group provided resources for learning disabilities and was more effective than the normal teaching applied to the control group. In addition, according to the results of the experimental process, Experiment-2 showed a higher level of improvement than Experiment-1 and Experiment-1 between the pre-post test averages of the Control group.

Keywords: Online Formative Assessment, Cognitive Diagnostic Model, Academic Achievement

Introduction

Although evaluation has no standard use, it can be done in many different ways, in many different contexts, and for many different purposes. However, in general terms, assessment can be defined as the process of collecting information about student achievements (Phye, 1997). According to Black and Wiliam (2006), the first and most important purpose of evaluation in education is to support learning (Cited by Yan & Cheng, 2015). Formative assessment, on the other hand, contains features that will meet the needs of teachers, as it provides valuable information to both students and teachers (Cauley & McMillan, 2010).

The main purpose of formative assessment is to provide feedback that can be used to increase the student's content knowledge, skills, and understanding. Strategies for obtaining meaningful feedback; The router should be response-specific, targeted, continuous, and delivered immediately (Shute, 2008).

* This study is a part of doctoral thesis conducted under the supervision of Prof. Dr. Bayram ÇETİN and prepared by Şeref AKPINAR.

**Director, Ministry of National Education, Ankara-Turkiye, sempatikbaskan46@gmail.com, ORCID ID: 0009-0001-0482-0108
** Prof. Dr., Gaziantep University, Faculty of Education, Gaziantep-Turkiye, bcetin27@gmail.com, ORCID ID: 0000-0001-5321-8028

To cite this article:
Akpınar, Ş., & Çetin, B. (2023). Investigation of the effect of online (web-based) formative assessment applications on students' academic achievement, *14*(Special Issue), 345-358. https://doi.org/10.21031/epod.1320182

Received: 26.06.2023
Accepted: 10.09.2023

Preparing a formative assessment application is a laborious and time-consuming task. In addition, extra time is needed outside the curriculum to give immediate feedback to the student after the assessment. Whereas, feedback for web-based formatting is usually embedded in the system. Evaluation for web-based formatting can be defined as evaluation practices in which the processes of evaluation and feedback to the learner are carried out through information and communication technologies. It is seen that there are many studies in which positive results are obtained for the learning process of the learner with formative assessment applications integrated into web-based learning environments (Brewer, 2004; Buchanan, 2000; Gardner, Sheridan, & White, 2002; Henly, 2003; Justham & Timmons, 2005; Khan, Davies, and Gupta, 2001; Peat and Franklin, 2002; Velan, Kumar, Dziegielewski, & Wakefield, 2002). Web-based assessment makes it possible to evaluate independently of time and place. In addition, all interactions of the learner can be recorded and then this information can be directed to the learner (Bayrak & Yurdugül, 2015). Evaluation tools for web-based formatting can be designed and used in order to eliminate time and space problems, contribute to the learning processes of learners, and provide immediate feedback to the learner (Cukusic et al., 2014).

Tekin (2010) determined that formative assessment has positive effects on success, attitude and remembering what has been learned in mathematics lessons. Formative assessment, which provides students with the opportunity to evaluate themselves and monitor their individual development, also contributes to the development of students' metacognitive awareness (Jones, 2007).

Formative assessment helps to achieve individual and specific goals in the mathematics learning and teaching process (Ginsburg, 2009). The use of formative assessment in mathematics teaching contributes positively to the success of students (Tempelaar et al., 2012; Tekin, 2010). In addition, formative assessment positively affects students' attitudes towards mathematics and the permanence of the knowledge they have acquired (Tekin, 2010).

Pierce and Ball (2009) recommend the use of technology in structuring mathematics-related assessment processes. By using technology in formative assessment in mathematics lessons, instant and individual feedback can be given to students (Stacey and Wiliam, 2013). Since formative assessment is a time-consuming practice for teachers and requires extracurricular work (Tekin & Özdemir, 2014), the use of technology in this process is important.

With this research, a different way is proposed to determine student knowledge and skills. It may be possible to identify the learning deficiencies in students and obtain profiles regarding the subjects in which students are weak and strong. In the literature review, it was seen that formative evaluation studies have been carried out historically in the past (İnaltun 2018; Tor and Bektaş 2023). It has been determined that online measurement and evaluation studies have been carried out later, and in recent years, online formative evaluation studies have been carried out. Arslan and Yetkin (2020) stated the studies conducted in a content analysis study on the use of online evaluation systems in education.

If we look at the studies carried out in recent years; In his study, Alır (2015) determined that he examined secondary school students' acceptance structures of the web-based formative evaluation system and their interactions with the feedback in the system. In 2016, Cabı examined student perceptions on e-assessment in distance education. It was observed that Demir (2017) discussed the effects of feedback given through computer-assisted formative assessment on the transfer of learning. Again, Başokçu et al. (2018), in their project study, realized the effectiveness of the cognitive memory-based monitoring model in order to increase Turkey's mathematics success in international large-scale exams. It was determined that another researcher, Hotaman (2020), conducted a study on the importance of formative assessment in terms of the success of online education.

In addition, in this study, online (web-based) formative assessment application and monitoring tests were prepared according to Cognitive Diagnostic Modeling, and after the monitoring tests, teachers provided detailed feedback according to Cognitive Diagnostic Modeling. In Cognitive Diagnostic Modeling, instead of the total scores and questions in the test, each individual taking the test is measured regarding their possession of each feature/qualities and sub-features/qualities that are tried to be measured in the questions, and the opportunity to give feedback based on this measurement is provided. For this reason, tests developed with Cognitive Diagnostic Models also serve to determine the educational needs of each student and provide feedback (Cheng and Chang, 2007).

By looking at the studies conducted, it has been determined that the research on students' weak and strong knowledge levels and skill levels is quite limited and very few studies have been conducted in Turkey. In addition, it has been determined that such a study has never been conducted to determine learning deficiencies in the field of mathematics in order to monitor the development of success by providing detailed feedback according to Cognitive Diagnostic Modeling through online (web-based) formative assessment application. For this reason, it is thought that the results of this study will make important contributions to the literature. In the light of the information to be obtained by this method, an alternative assessment application will be presented for teachers. With the formative evaluation application, students will not be evaluated only according to level or result, but an evaluation that is not intended for grading will be made during the process. Teachers' information and technology literacy will increase, and there will be an opportunity to monitor the development of students' success and identify students' learning deficiencies by providing feedback through online (web-based) formative assessment application.

In this study, Student/Teacher Support System (ÖDS) and Education Information Network (EBA) platforms created by the Ministry of National Education (MEB) were used. It is thought that the implementation of our study in a national infrastructure created by the Ministry of National Education (MEB) will provide easy applicability throughout the country, will not impose a lot of burden on practitioners, will be economical in terms of implementation, will be easily accessible, and will be a useful infrastructure application due to all these situations. In addition, implementing this study in a national infrastructure created by the Ministry of National Education (MEB) will provide a more reliable application infrastructure.

It is thought that our study can contribute to policymakers in making decisions about the teaching process and curriculum and provide important information to researchers in their studies.

The purpose of this research is to examine the effects of providing resources for learning disabilities by the system and providing feedback regarding learning disabilities by the teacher within the scope of an online (web-based) formative evaluation application for 10th grade secondary school students in mathematics course quadratic equations on students' success.

In this research, the answers to the following sub-problems will be sought:

1. How are the proficiency levels of the experimental and control group students in the field of Second Degree Equations learning before the experiment?

2. How are the proficiency levels of the experimental and control group students in the field of Second Degree Equations learning after the experiment?

3. Is there a difference between the pre-test and post-test results of the experimental and control group students in the field of Second Order Equations learning?

Methods

The research was conducted using a semi-experimental design. Semi-experimental designs are usually the best type of design that can be used in field studies where a person wants to make causal inferences. In the process of applying the semi-experimental design, the process of determining the groups is important. It is necessary to try to equalize the groups in which the research will be conducted as much as possible in terms of the variables subject to the research (Christensen, Johnson, & Turner, 2015).

Information about the research design and data collection process is presented in Table 1.

Table 1

Research Design and Data Collection Process

Group	Pre-Test	Experimental Procedure	Final Test
Experiment-2	İDDT	Within the scope of the formative assessment application, the system provides resources for learning disabilities with a total score, detailed feedback by the teacher on learning disabilities according to CDM	İDDT
Experiment-1	İDDT	Within the scope of the formative assessment application, the system provides resources for learning disabilities with the total score.	İDDT
Control	İDDT	No action will be taken	İDDT

İDDT: Second Degree Equations

Table 1.as can be seen in the research, the experimental process of the research will be conducted through three groups in the form of Experiment-1, Experiment-2 and control group. The experimental groups were divided into two different groups within the research. Before the experimental procedure, the “Second Degree Equations Test (IDDT)” was applied as a preliminary test. The experiment-1 group consists of 104 students; Experiment-2 consists of 102 students; and the control group consists of 96 students. As a result of the monitoring tests applied to the Experiment-1 students, the total score obtained by the system within the scope of the formative assessment application was indicated and resources for learning disabilities were presented. On the other hand, Experiment-2 students were given detailed feedback from the teacher for learning disabilities in accordance with the CDM, as well as the system's overall score as part of the formative assessment application as a result of the monitoring tests applied. According to the CDM, detailed feedback was made according to the characteristics and sub-characteristics that students wanted to acquire in the Mathematics course on Second-Degree Equations.

Only pre- and post-tests were applied to the control group students. Experiment-1 and Experiment-2 were created by the teacher to determine the effect of giving feedback according to CDM; the control group was formed to determine the effect of both monitoring tests and feedback. The “Second Degree Equations Test (IDDT)” was applied to all groups as the final test at the end of the experimental process.

Working Group

In the research, using the stratified, random cluster sampling method, a total of 302 students selected from 4 schools in Göksun and Andırın districts and 3 branches from each school and 12 branches were determined as the sample of the research. According to the design, the determined sample was divided into Experiment-1, Experiment-2 and Control groups in each school. Experiment-1 group consisted of 102 students from 4 branches of 4 schools; Experiment-2, from a total of 104 students in 4 branches of 4 schools; The control group consisted of 96 students in 4 branches of 4 schools.

Application Process

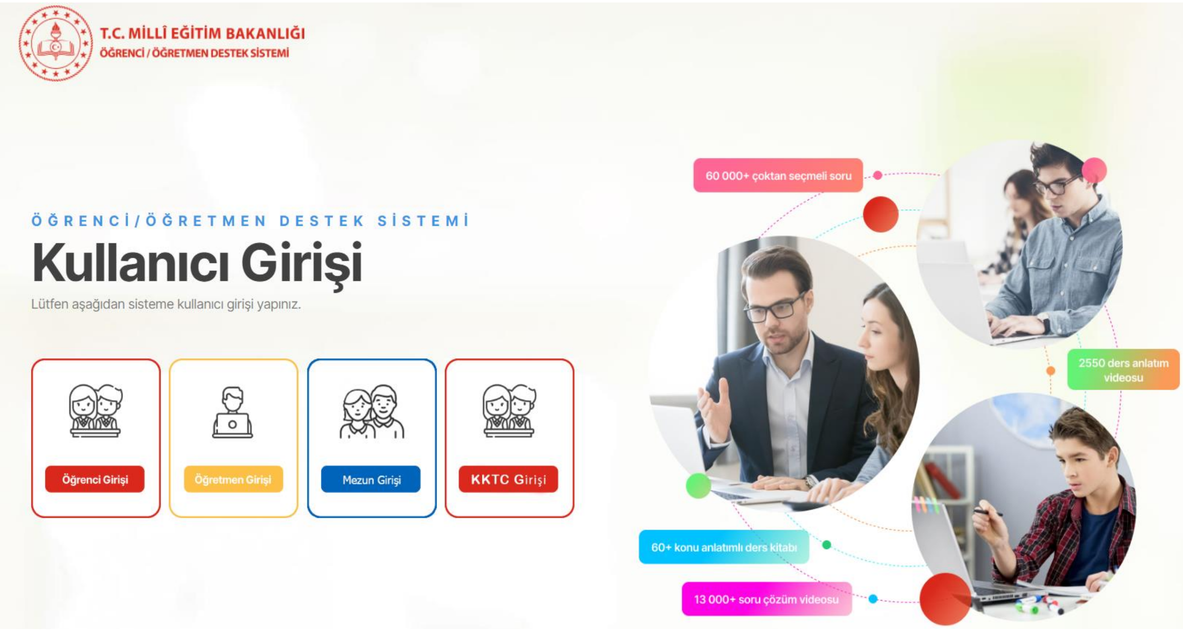
The research was carried out in a Science High School, two Anatolian High Schools and an Anatolian Imam Hatip High School located in Göksun and Andırın districts of Kahramanmaraş province. A total of four Mathematics teachers, one from each of the schools mentioned in the research, worked. During the experimental process of the research, three follow-up tests were applied to the 10th Grade

Experiment-1 and Experiment-2 groups on the subject of Second Degree Equations in the Mathematics course for an average of 6 weeks. A pilot application of the test was carried out before each follow-up test application. In order to prepare the monitoring tests, they were created from the question pool in ÖDS (Student-Teacher Support System) affiliated with the Ministry of National Education, based on the achievements in the secondary school mathematics course curriculum, and opinions were received from mathematics teachers working at the school and experts in the fields of measurement and evaluation for the content validity of the questions. An achievement test was created by randomly selecting questions from the pool to measure each achievement.

The prepared follow-up tests were sent to the relevant groups via ÖDS (Student-Teacher Support System) at 2-week intervals, according to the subject achievements. IDAT (Quadratic Equations test) was used as pretest and posttest at the beginning and end of the application. Figure 1 shows the screen where students can enter the ÖDS system.

Figure 1

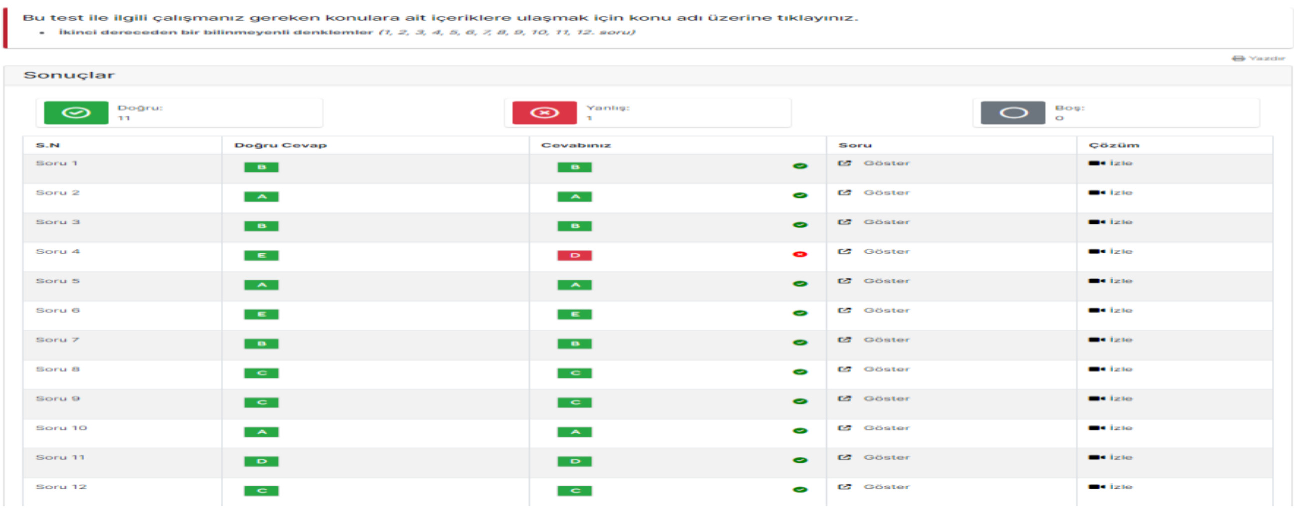
STUDENT/TEACHER SUPPORT SYSTEM (ÖDS) Login Screen



During the implementation process, after a lesson was taught approximately every two weeks and the relevant achievement was given, the monitoring test was uploaded to the ÖDS system and the necessary sharing and information was provided for the students to access the questions. Students logged into the system using their computers, tablets or smartphones. Then, the students solved the follow-up tests whenever and wherever they wanted. An example of a screen showing the total scores students received after their answers and the resources for their learning disabilities is shown in Figure 2.

Figure 2

ÖDS Assignment Result Screen



At the end of the application, the follow-up test results for the Experiment-2 group were examined one by one through the ÖDS system. For each student's learning disabilities, detailed feedback and assignments were made by the teacher according to CTM (Cognitive Diagnostic Modeling) over the EBA system, which is also a different application. According to CDM, detailed feedback was given according to the characteristics and sub-features that students wanted to gain in the Mathematics lesson Second Degree Equations. These feedbacks were given to the same students by the same teachers after three follow-up tests with two-week intervals. Experiment-1 and Experiment-2 group students were advised to watch the lesson videos suggested by the system before each viewing test, but their viewing status could not be followed. It was thought that making a force on this subject would create forcing results for both the practice teachers and the students of the experimental groups.

An example of a screen showing the message sent by the teacher via the EBA system is presented in Figure 3.

Figure 3

EBA Message Screen

Data Collection Tools

As a data collection tool, the "Second Degree Equations Test (IDDT)" was applied as a pre-test and post-test before and after the experimental process.

Second Degree Equations Test (IDDT)

The Second Degree Equations Test "IDDT" was used to measure the pre- and post-application knowledge levels of the students in the experimental and control groups about second degree equations. In order to develop the test, a question pool was created based on the achievements in the secondary mathematics curriculum for the validity of the questions, opinions were obtained from mathematics teachers working in schools and experts in the field of educational programs and teaching. The questions from the pool were randomly generated to measure each win. The pilot application of the test was carried out before the application of the Second Degree Equations Test. Item analysis was performed after the pilot application. As a result of the analyses carried out after the pilot application and in accordance with the expert opinion, an 18-question Second-Order Equations Test was created. The KR-20 value of the test is 0.89. According to this result, IDDT is reliable.

Analysis of the Data

Before the statistical analyses, it will be tested whether the quantitative data meet the prerequisites of the analyses. In this context, loss value analyses, normality test and homogeneity of variances were checked.

In the study, whether the differences between the pre-test score averages of the groups were significant was examined by one-way analysis of variance (ANOVA).

Covariance analysis (ANCOVA) is one of the most frequently used analyses in experimental designs using pre-post test. In this analysis, in which the pre-test scores are taken as covariates, the differences between the groups are calculated by checking the pre-test scores Decently. According to the fact that the pretest scores did not differ between the groups in the research design, they were examined by covariance analysis of the experimental process.

Results

Pre-Test Variance Analyses for the Proficiency Levels of Research Groups in the Field of Second Degree Equations Learning Before the Experiment

The results of the analyses conducted to decipher whether there is a significant difference between the group averages of the pre-test scores of the students in the research groups are given in Table 2 and Table 3:

Table 2

Descriptive Statistics of the Pre-Test Total Scores

Group	N	$\bar{x}$	s	Skewness
Control	96	4.00	1.66	.33
Experiment1(System)	102	4.06	1.79	.22
Experiment2(Feedback)	104	3.92	1.58	.13
Total	302	4.01	1.67	.24

When Table 2 is examined, it is seen that the total pre-test scores of the research groups are between the deciency coefficients of -1.1 for all groups. These results mean that not all groups showed excessive deviations from the normal distribution (Leech, Barrett, & Morgan, 2007). However, the Levene statistic shows that the variances of the groups are homogeneous (.525, p>0.05). This indicates that the data set meets the assumptions of ANOVA analysis.
Table 2 shows the differences between the pre-test score averages of the research groups when examined .06, .08 and .it is seen that it is 14. One-way analysis of variance (ANOVA) was applied to determine whether these differences were significant. The results of the analysis are given in Table 3.

Table 3
ANOVA Analysis of the Pre-Test Scores of the Groups

	Sum of Squares	sd	Mean Square	F	P
Between Groups	.96	2	.48	.17	.85
Within Groups	845.03	299	2.83		
Total	845.98	301			

When Table 3 is examined, no significant difference was found between the pre-test scores of the research groups (F(2-301) = .17, p>.05). These findings show that the pre-test scores of the research groups are equal.

Post-Test Analysis of Variance for the Proficiency Levels of the Research Groups in the Field of Learning Second Order Equations after the Experiment

The results of the analyses carried out to reveal whether there is a significant difference between the group averages regarding the post-test scores of the students in the research groups are given in Tables 4 and 5:

Table 4
Descriptive Statistics of Post-Test Total Scores

Group	N	$\bar{x}$	s	Skewness
Control	96	8.44	3.40	.72
Experiment1(System)	102	10.16	3.33	.57
Experiment2(Feedback)	104	12.11	4.04	.92
Total	302	10.28	3.90	.16

When Table 4 is examined, it is seen that the total pre-test scores of the research groups are between the deciency coefficients of -1.1 for all groups. These results mean that not all groups showed excessive deviations from the normal distribution (Leech, Barrett, & Morgan, 2007). However, the Levene statistic (.09, p>0.05) shows that the variances of the groups are homogeneous. This indicates that the data set meets the assumptions of ANOVA analysis.

When Table 4 is examined, it is seen that the differences between the decal-test score averages of the research groups are 1,719, 1,949 and 3,668. One-way analysis of variance (ANOVA) was applied to determine whether these differences were significant. The results of the analysis are given in Table 5.

Table 5

ANOVA Analysis of the Post-Test Scores of the Group

	Sum of Squares	sd	Mean Square	F	P	η2	Groups Difference
Between Groups	674.13	2	338.06	25.85	.00	.15	Exp-2-Exp-1;
Within Groups	3898.95	299	13.04				Exp-2-Control;
Total	4573.08	301					Exp-1-Control

When Table 5 is examined, there is a statistically significant difference between the post-test scores of the research groups ($F(2\text{-}301)= 25.85$, $p>0.05$). The significance of the difference in the group effect ($\eta 2 = .12$), it can be stated that it is of medium effect size. The differences between the groups were examined by the Bonferroni tes According to the results of the Bonferroni test, the mean of the final test of Experiment-2 group (M=12.11, s=4.04) is statistically higher than the mean of Experiment-1 (M=10.16, s=3.33) and the Control group of final tests (M=8.44, s=3.40). In the same way, it was found that there was a statistically significant difference between the post-test averages of the Experimental-1 group and the control group in favor of the Experimental-1 group.

3. Covariance Analysis of the Experimental Process

Another analysis that is frequently used in experimental patterns using pre-post test is covariance analysis (ANCOVA). In this analysis, in which the pre-test scores are taken as covariates, the differences between the groups are calculated by checking the pre-test scores decently. Although it was seen that the preliminary test scores did not differ between the groups in the research pattern, it was found that the experimental process should be examined again by covariance analysis taking into account the sample size. ANCOVA results for final test scores are given in Table 6.

Table 6

ANCOVA Analysis of the Final Test Scores

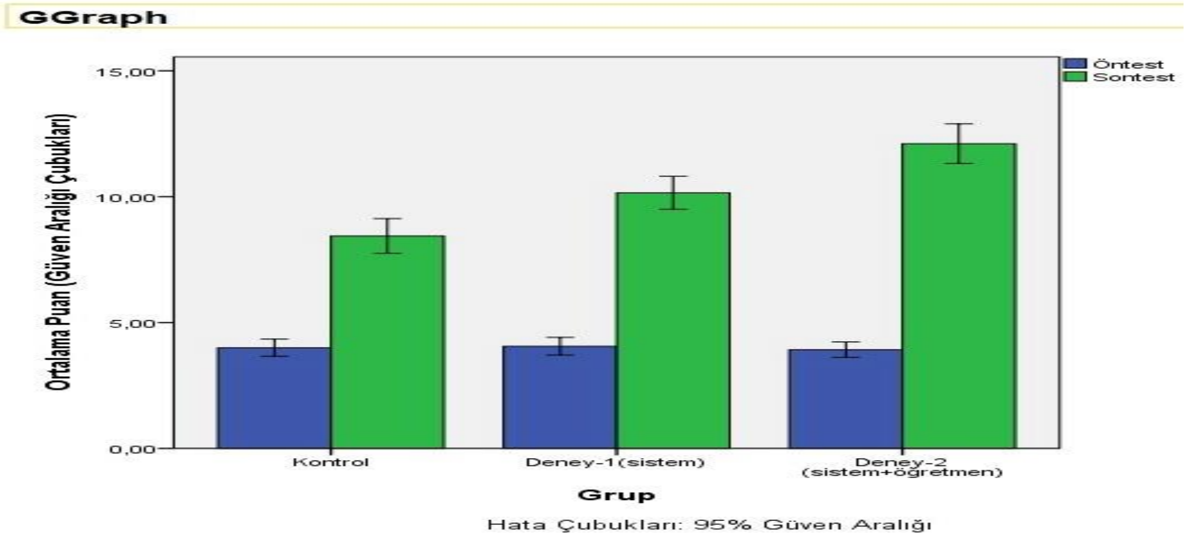
	Sum of Squares	sd	Mean Square	F	P	η2	Groups Difference
Pre-Test	2080.42	1	2080.42	340.92	.00	.534	Exp-2- Exp -1;
Group	724.29	2	362.14	59.34	.00	.285	Exp -2-Control; Exp -1-
Error	1818.53	298	6.10				Control
Total	36497.00	302					

According to the analysis results in Table 6, when the differences between the pre-test scores of the groups were controlled, it was found that the differences between the post-test scores of the groups were significant. ($F(2,302) = 59.34$, $p=.00$). The Deciency between which groups the difference was significant was examined by the Bonferroni test. According to the analysis results, it was found that there was a significant difference in favor of Experiment-2 Deciency between Experiment-2

(M=12.22) and Experiment-1 (M=10.05) and Control (M=8.43) groups. At the same time, the average of the Experimental-1 group is also statistically significantly higher than the control group.
The change in marginal averages is given in Figure 4.

Figure 4

Marginal Group Averages



When Figure 4 is examined, the change in the averages of the groups between the pre and post test is. The difference between the pre- and post-test in each group was found to be statistically significant. However, considering the common effect, it is observed that the average final test score of the Experiment-2 group increased statistically at a higher level than both groups, and the Experimental-1 group increased statistically higher than the Control group.

Discussion and Conclusion

The purpose of this research is to examine the effects of providing resources for learning disabilities by the system and providing feedback regarding learning disabilities by the teacher within the scope of the online (web-based) formative assessment application of secondary school 10th grade students in mathematics course quadratic equations on the students' success.
According to the results obtained, it was determined that there was no statistically significant difference between the pre-test averages of the groups, but a statistically significant difference emerged in the post-test. The post-test average of the Experiment-2 group was higher than the post-test average of the Experiment-1 and Control groups; The post-test mean of the Experiment-1 group also showed a statistically significant difference from the post-test mean of the control group. In addition, according to the results of the experimental process, Experiment-2 showed a higher level of improvement than Experiment-1 and Experiment-1 showed a higher level of improvement between the pre-post test averages than the Control group.
According to the results of the research, the system applied to the Experiment-2 group provided resources for learning deficiencies and the teacher provided feedback regarding learning deficiencies,

the system applied to the Experiment-1 group provided resources for learning deficiencies, and the normal education applied to the Control group; It has been determined that the system applied to the Experiment-1 group and the provision of resources for learning disabilities are more effective than the normal education applied to the Control group.

When the literature related to the study is examined, results that support our research results are seen.

Baleni (2015) stated in his study that effective online formative assessment can foster student and assessment-centered focus through formative feedback and enrich student engagement with valuable learning experiences. Ongoing reliable assessment tasks and interactive formative feedback have been identified as important features to deal with intimidations towards rationality and reliability in the online formative assessment environment.

Başokçu et al. (2018) in their project study, no difference was observed between the pre-test averages between the Experiment-1 group, which was given detailed feedback, and the Experiment-2 group, which was given feedback only on the total score, and between Experiment-2 and the Control group, which was not given any feedback, while the final significant differences were detected in the test averages.

In the study of Hotaman (2020), it was evaluated that formative assessment, which can provide students with the necessary rapid feedback during online courses, will become important for success. Hotaman's study is parallel with our research result, showing that the Experiment-2 group, receiving feedback from both the system and the teacher, and the Experiment-1 group, receiving feedback from the system alone, were more effective than the Control group, which received no feedback.

In their study, Karadağ and Özgür (2021) stated that more feedback should be given to learners regarding the learning and evaluation process. Additionally, Demir (2017) stated in his study that the average scores of students who received detailed feedback increased more than other groups. Both Karadağ and Özgür (2021) and Demir (2017) found that the Experiment-2 group, receiving feedback from both the system and the teacher, performed better than the Experiment-1 group, which received feedback from the system alone.

Pekcan and Toraman (2022) stated in their study that measurement and evaluation in distance education is useful in detecting learning deficiencies and increasing the quality of learning. The result stated by Pekcan and Toraman (2022) supports the conclusion that we obtained as a result of our research that the average of the groups in which online formative evaluation was applied is more effective than the average of the group in which online formative evaluation was not applied and student success is higher.

Hannah, James and Williams (2014); Shirley and Irving (2015); Reeves, Gunter and Lacey (2017); Faber, Luyten and Visscher (2017); In their study, Pemberton (2018) concluded that technology-supported formative assessment practices in mathematics education have a positive effect on student success. All these results are parallel to the result we obtained form our research that online formative assessment application increases students mathematics achievement.

The following recommendations are made according to the experiences gained during the study and the results of the study.

In this study, the Student/Teacher Support System (ÖDS) system created by the Ministry of National Education (MEB) was used. Since the teacher does not have a link to send messages to students in the ÖDS system, the Education Information Network (EBA) created by the Ministry of National Education (MEB) was used. According to the results of our study, the provision of resources for learning disabilities by the system applied to the Experiment-2 group and the feedback provided by the teacher for the learning disabilities differed statistically significantly from the system applied to the Experiment-1 group and the normal teaching applied to the control group showed a higher level of improvement among the pre-post test averages. Considering this result, it will be a more useful system if a message link is added to the ÖDS (Student/Teacher Support System) by the Ministry of National Education (MEB).

Again, in the studies carried out during the study process, a sufficient amount of question and solution videos were found for the subject studied in the ÖDS system. However, in the examination of other courses and topics, it has been determined that some courses and topics do not have enough questions and solution videos. It is recommended to increase the questions and solution videos of all the lessons and subjects in the ÖDS system in a sufficient amount in order to contribute more to the scientific studies to be made and to the students. According to the results of our study, the Experiment-2 group, which was given detailed feedback and assignment according to CDM (Cognitive Diagnostic Modeling) through the EBA system by the teacher for the learning disabilities of each student, differed significantly from the other groups and showed a high level of improvement from the other groups. . In addition, it has been determined that the online (web-based) formative assessment application also increases the academic success of the students. Considering these results, it is recommended that our teachers use online (web-based) formative assessment applications and, in particular, provide detailed feedback and assignments according to CDM (Cognitive Diagnostic Modeling).

As a result of our study, it was determined that the online (web-based) formative assessment application increased the academic success of the students. It can be interpreted that the source of the change in this result is the process-based, non-grading, online (web-based) formative assessment application that has no time and place limitations, as well as detailed feedback on more specific sub-features under the outcome, not just based on the outcome. Considering the result of our study, it is thought that increasing the studies on online applications and using CDM in giving feedback will serve to make clearer sentences on this subject.

Declarations

Author Contribution: Şeref Akpınar: Conceptualization, methodology, analysis, writing & editing, visualization. Bayram Çetin: Conceptualization, methodology, writing-review & editing, supervision.

Conflict of Interest: No potential conflict of interest was declared by the authors.

Ethical Approval: The study was ethically approved by Gazi University Rectorate Ethics Commission (decision number: 2023-110, 09/02/2023).

References

Alır, A. (2015). Investigating secondary school students' acceptance of web based formative assessment system and students' interaction with different feedback types (Thesis No. 381419) [Master's thesis, Hacettepe University]. Council of Higher Education Thesis Center. https://tez.yok.gov.tr/UlusalTezMerkezi/

Arslan, K. & Yetgin, G. (2020). Use of online assessment results in education: a content analysis. Turkish Studies - Education, 15(2), 651-671. https://dx.doi.org/10.29228/TurkishStudies.39646

Baleni, G. Z. (2015). Online formative assessment in higher education: Its pros and cons. The Electronic Journal of e-Learning, 13(4) 2015, 228-236. https://files.eric.ed.gov/fulltext/EJ1062122.pdf

Başokçu, T. O., Bardakçı, V., Çakıroğlu, E., Öğretmen, T., Yurdakul, B., & Akyüz, G. (2018). A model proposal to increase Turkey's mathematics achievement in international large-scale exams: The effectiveness of the cognitive diagnosis-based monitoring model (Project: SOBAG 3501. Project No: 115K531). https://open.metu.edu.tr/bitstream/handle/11511/50310/TVRnMU56SXk.pdf

Bayrak, F., & Yurdugül, H. (2015). E-Assessment and E-Feedback. Educational Technology Readings, (s. 449-468).

Black, P., & Wiliam, D. (2006) Developing a theory of formative assessment. In J. Gardner (Ed.), Assessment and learning (pp. 81-100). London: Sage.

Black, P., & Wiliam, D. (2006). Assessment for learning in the classroom. In J. Gardner (Ed.), Assessment and Learning (pp. 9-25). London: Sage Publications.

Brewer C.A. (2004). Near real-time assessment of student learning and understanding in Biology courses. Bioscience 54, 1034-1039.

Buchanan, T. (2000). The efficacy of a World Wide Web mediated formative assessment. Journal of Computer Assisted Learning, 16, 193–200.

Cabı, E. (2016). The perception of students on e-assessment in distance education. Journal of Higher Education of Science. 6(1). 94- 101. https://dergipark.org.tr/tr/download/article-file/1711633

Cauley, K. M., & McMillan, J. H. (2010). Formative assessment techniques to support student motivation and achievement. The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 83(1), 1-6. DOI: 10.1080/00098650903267784 To link to this article: https://doi.org/10.1080/00098650903267784

Cheng, Y. ve Chang, H. (2007). The modified maximum global discrimination index method for cognitive diagnostic computerized adaptive testing. CAT and Cognitive Structure Paper Session, June, 7.

Christensen, L.B., Johnson, R.B., and Turner, L.A., (2015). Research methods, design, and analysis. Boston: Pearson Education Limited.

Cukusic, M., Garaca, Z., & Jadric, M. (2014). Online self-assessment and students' success in higher education instituions. Computers & Education, 72, 100-109.

Demir, D. (2017). The effects of immediate feedback interventions with different contexts through computer assisted formative assessment on transfer of learning. (Thesis No. 381419) [Master's thesis, Bahçeşehir University]. Council of Higher Education Thesis Center. https://tez.yok.gov.tr/UlusalTezMerkezi/

Faber, Janke & Luyten, Hans & Visscher, Adrie. (2017). The effects of a digital formative assessment tool on mathematics achievement and student motivation: Results of a randomized experiment. Computers & Education. 106. 10.1016/j.compedu.2016.12.001.

Gardner, L., Sheridan D. & White D. (2002) A web-based learning and assessment system to support flexible education. Journal of Computer Assisted Learning, 18, 125-136.

Ginsburg, H. P. (2009). The challenge of formative assessment in mathematics education: Children's minds, teachers' minds. Human Development, 52(2), 109-128. DOI: 10.1159/000202729

Hannah, John & James, Alex & Williams, Phillipa. (2014). Does computer-aided formative assessment improve learning outcomes. International Journal of Mathematical Education in Science and Technology. 45(2), 269-281. 10.1080/0020739X.2013.822583.

Henly, D. C. (2003) Use of Web-based formative assessment to support student learning in a metabolism/nutrition unit. European Journal of Dental Education, 7, 116–122.

Hotaman, D, (2020). The importance of formative evaluation for success of online education. International Journal of Social Studies, 13(73), 729-738. https://www.sosyalarastirmalar.com/articles/onlne-etmn-baarisi-aisindan-bmlendrc-deerlendrmenn-nem-the-importance-of-formative-evaluation-for-success-of-online-educ.pdf

İnaltun, H., & Ateş, S. (2018). Formative assessment in science education: A literature review. Gazi University Journal of Gazi Educational Faculty, 38(2), 567-612. https://doi.org/10.17152/gefad.353975

Jones, D. (2007). Speaking, listening, planning and assessing: the teacher's role in developing metacognitive awareness. Early Child Development and Care, 177:6-7,569-579.

Justham, D. & Timmons, S. (2005). An evaluation of using a web-based statistics test to teach statistics to postregistration nursing students. Nurse Education Today, 25, 156–163.

Karadağ, N, Özgür, A. Z. (2021). Examining the opinions of learners on assessment and evaluation practices in Anadolu University Open Education System. Anadolu University Journal of Social Sciences, 21 (2), 505-530. https://dergipark.org.tr/en/download/article-file/1849391

Khan, K. S., Davies, D. A. & Gupta, J. K. (2001) Formative selfassessment using multiple true-false questions on the Internet: feedback according to confidence about correct knowledge. Medical Teacher, 23, 158–163.

Leech, N., Barrett, K., & Morgan, G. A. (2007). SPSS for Intermediate Statistics: Use and Interpretation, Third Edition. Taylor & Francis.

Peat, M., & Franklin, S. (2002). Supporting student learning: the use of computer-based formative assessment modules. British Journal of Educational Technology, 33(5), 515-523.

Pekcan, N. & Toraman, Ç. (2022). Investigation of online assessment and evaluation applications in the pandemic according to teacher-student views. Journal of Erzincan University Faculty of Education, 24 (1), 120-129. DOI: 10.17556/erziefd.862654

Pemberton, Katie Jo. (2018). Assessment for feedback and achievement growth for middle school math students. [The doctoral study, Walden University]. Walden Dissertations and Doctoral Studies Collection. https://scholarworks.waldenu.edu/dissertations/5928

Phye, G. D. (1997). Handbook of classroom assessment: Learning, achievement, and adjustment. Academic Press. San Diego: Academic Press.

Pierce, R., & Ball, L. (2009). Perceptions that may affect teachers' intention to use technology in secondary mathematics classes. Educational Studies in Mathematics, 71(3), 299-317.

Reeves, Jennifer & Gunter, Glenda & Lacey, C., (2017). Mobile learning in pre-kindergarten: Using student feedback to inform practice. 20. 37-44.

Shirley, M. L., & Irving, K. E. (2015). Connected classroom technology facilitates multiple components of formative assessment practice. Journal of Science EducationaAnd Technology, 24(1), 56-68. 10.1007/s10956-014-9520.

Shute, V. J. (2008). Focus on formative feedback. Review of Educational Research, 78(1), 153-189. https://doi.org/10.3102/0034654307313795.

Tekin, E. G. (2010). The effect of formative assessment in mathematics education (Unpublished Master's Thesis). Marmara University Institute of Educational Sciences.

Tekin, E. G., & Özdemir, A. Ş. (2014). The effect of formative assessment on mathematics achievement and recall. International Conference on Education in Mathematics, Science & Technology. Necmettin Erbakan University, Konya.

Torun, B., & Pektaş, M. (2023). Content analysis of postgraduate theses on formative assessment and education in Turkey. Journal of Turkısh Academic Publications (TAY Journal), 7(2), 407-442.

Velan, G. M., Kumar, R. K., Dziegielewski, M. & Wakefield, D. (2002) Web-based self-assessments in pathology with Questionmark Perception. Pathology, 34, 282–284.

Yan, Zi & Cheng, Eric. (2015). Primary teachers' attitudes, intentions and practices regarding formative assessment. Teaching and Teacher Education. 45. 128–136. 10.1016/j.tate.2014.10.002.

ISSN: 1309 – 6575
Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi
Journal of Measurement and Evaluation in Education and Psychology
Research Article; 2023; 14(Special Issue); 359-381

Training 21st Century English Language Teachers in Turkish Context: Development of a Technology-Enhanced Measurement Curriculum

Burcu ŞENTÜRK* Beyza AKSU DÜNYA** Mehmet Can DEMİR***

Abstract

A case study that included 26 English Language teacher candidates was designed to develop an evidence-based measurement curriculum in Turkey, examining teacher candidates' experiences on the newly developed course and taking remedial actions to update the syllabus if needed. Data was collected using multiple sources: a pre-course survey, a weekly discussion board on Edmodo and a post-course survey. Survey data obtained from rating-scale items was analyzed using descriptive statistics and data visualization packages embedded in R. Open-ended survey data and discussion board data were content-analyzed using MaxQDA software. The results revealed that students had limited awareness regarding assessment for learning concepts and digital tools that could be used for assessment for learning purposes at the beginning of the course. Course content, in-class activities and projects helped them develop hands-on skills in developing sound language assessments as well as raised their awareness with respect to the importance of computer-based language assessment.

Keywords: Digital literacy, Language assessment, Computer-based assessment, Language assessment literacy, English language teaching

Introduction

Digital literacy is an important 21st century skill for teacher candidates. The idea of literacy practice in the 21st century has been dramatically impacted by the technological revolution and globalization, which highlights the necessity for educators to implement effective teaching strategies that include or blend traditional and emergent literacies (Yang et al., 2022). They can discuss the learning material in greater depth and provide more sophisticated knowledge in the context of daily life (Comeaux, 2002). Technology improves teaching and assessment capacities and opens doors for growth and variety in how learners are evaluated, including textual communication skills, cooperation, teamwork, and reflective thinking (Eyal, 2010; Liang & Creasy, 2004). Additionally, by assigning assessment activities and allowing students to progress at a speed that suits their needs, digital environments can address the diversity of learners (Alderson, 2000). The use of digital technologies in assessment becomes even more important for domains where performance-based skills (i.e., speaking) are at the forefront, necessitating English Language Teaching (ELT) teachers to utilize technology to design and conduct effective performance-based assessments. However, the Turkish ELT programs' curriculum lacks supporting teacher candidates to gain the needed skills for planning and implementing technology-enhanced assessment.

Since English is the most widely used language for commerce and international communication, many non-English speaking nations have long been affected by English on a global scale, and Turkey is no

* Asst. Prof. Dr., Bartın University, Faculty of Education, Bartın- Turkiye, bsenturk@bartin.edu.tr, ORCID ID: 0000-0001-8951-3256
** Assoc. Prof. Dr., Bartın University, Faculty of EducationBartın- Turkiye, baksu2@uic.edu, ORCID ID: 0000-0003-4994-1429
*** Research Assistant, Bartın University, Faculty of Education, Bartın-Turkiye, mehmetdemir@bartin.edu.tr, ORCID ID: 0000-0001-7849-7078

To cite this article:
Şentürk, B., Aksu Dünya, B., & Demir, M. C. (2023). Training 21st century english language teachers in turkish context: Development of a technology-enhanced measurement curriculum, *Journal of Measurement and Evaluation in Education and Psychology, 14*(Special Issue), 359-381. https://doi.org/10.21031/epod.1261763

Received: 7.03.2023
Accepted: 29.09.2023

exception (Balıkcıoğlu Akkuş & Altay, 2023). Consequently, it has become necessary to adopt modern educational policies and procedures in order to stay up with the most recent advancements in foreign language education. Learning English is viewed as a necessity for university students to achieve academic success and to have better career chances in the future, along with the rise in partly-English and fully-English departments in higher education (Doğançay-Aktuna & Kızıltepe, 2005). As a result, there is a lot of work being done to help learners acquire the language skills they need and to enhance their communicative and linguistic proficiency, but this process is not without its challenges (Balıkcıoğlu Akkuş & Altay, 2023).

It is clear that technology is employed in every course in educational settings nowadays. Technology, which is employed in many courses for various reasons and methodologies, not only enhances the lectures for students but also has a good impact on their achievement. The student whose course achievement rises also has a favorable view of their learning, gets more engaged in class, and gains self-assurance towards learning. The ability to acquire knowledge in only a few seconds is one of technology's greatest contributions to learning environments. In addition to books, like in traditional teaching, platforms with various software and features based on the internet are also used to gain the knowledge and skills linked to numerous themes in different classes; nevertheless, it is also crucial to note that this method of learning is not the only one (Arslan, 2023).

According to Vähäsantanen (2015), curriculum change is viewed as a dynamic and multifaceted reality in teachers' professional lives. To effectively support the objectives of curriculum reform, EFL teachers must continuously improve their professionalism (Jiang and Zhang, 2021). However, the cultivation of teachers' new roles as well as ongoing learning are both essential to the professional growth of EFL teachers (Jiang, 2022; Lei & Medwell, 2022; Lei & Xu, 2022; Tao & Gao, 2017; Vähäsantanen & Eteläpelto, 2009; Yang, 2015).

Uzun (2016) explained that teacher training programs (TTP) in Turkey have undergone extensive investigation and modification since institutions and faculties of education were established. This has resulted from the innovative approaches to and models of teaching that have been suggested by the trends and findings of the time. However, despite a willingness to adopt international patterns, the Turkish educational system has had significant issues bridging theory and practice.

The challenges to technology integration that instructors perceive have been the subject of numerous research (Atman Uslu, 2022). According to Mercader & Gairin (2020), barriers such as a lack of education, ignorance of digital technology teaching methods, a lack of planning, an excessive workload, a lack of time, the generational divide, technophobia, a lack of evaluation and incentives were given priority by university teachers. According to a study done with Indonesian instructors, there are significant challenges, including a lack of expertise and experience in ICT education, a lack of time and resources, and a bad Internet connection (Muslem et al., 2018).

This study intends to investigate the development of a technology-enhanced measuring curriculum for training English language instructors for the twenty-first century in Turkey. As technology presents several chances to improve teaching efficacy, engage students, and promote authentic language acquisition, its integration into language instruction has grown in importance (Greenier et al., 2021). However, effective technology integration calls for teachers to have a firm grasp of both language pedagogy and technological resources (Jiang, 2022).

The suggested curriculum places a strong emphasis on the measuring side of teaching languages, acknowledging the crucial role that evaluation and assessment play in determining how best to teach and keeping track of student progress (Lei, 2022). It seeks to provide English language instructors in Turkey with the knowledge and abilities required to successfully plan, carry out, and assess technologically enhanced exams.

This curriculum intends to enable instructors to use technology in their teaching practices while also assuring the validity, reliability, and fairness of assessment procedures by addressing the unique demands and challenges faced by English language teachers in the Turkish setting (Arslan, 2023). The curriculum is made to give instructors awareness of a range of technology resources and tools that might improve measurement, enabling them to build fun and interactive tests that accurately reflect students' language (Atman Uslu, 2022).

In conclusion, developing a technology-enhanced measuring curriculum for English language instructors in Turkey is an essential first step in preparing teachers for the opportunities and difficulties of the 21st century. Teachers can use technology to develop relevant evaluations that support good language learning outcomes by including it in the measurement process. This curriculum aims to empower English language instructors by giving them the knowledge and abilities they need to successfully navigate the digital world and deliver high-quality instruction to their students in a society that is continually changing (Arslan, 2023).

Curriculum Innovation in ELT

The term "curriculum evaluation" is a set of actions used to gather data on how policies, programs, curricula, courses, educational software, and other instructional resources operate and have an impact on students (Gredler, 1996). The creation, implementation, and maintenance of curricula depend on curriculum evaluation. The goal of curriculum assessment is to ascertain the advantages and disadvantages of the curriculum before implementation as well as the efficiency of its delivery following implementation. The health of education and its programs depends on evaluation (Ornstein & Hunkins, 2014). Moreover, Posavac and Carey (2003) explain six purposes of program evaluation which are to:

a. assess unmet needs,
b. document implementation,
c. measure results,
d. compare alternative programs,
e. provide information to maintain and
f. develop quality and detect negative side effects.

Peacock (2009) claims that establishing a causal link between expectations and results is no longer the only emphasis of program assessment. Instead, it is increasingly used to make program decisions based on a range of systematic data collecting and analysis methods that are related to effectiveness, efficiency, value, and appropriateness.

The Council of Higher Education (CoHE) decided on, reorganized, and introduced ELTTP in Turkey in the 2006–2007 academic year. The package program is used consistently throughout Turkey's English Language Teaching (ELT) Departments, despite possible minor variations in course titles, content, and methods among various universities. The technical courses in the proposed curriculum were designed to advance knowledge in the specialized field of English language (EL) instruction. Thus, the technical courses concentrated on additional EL education-related topics (such as linguistics, pragmatics, syntax, discourse analysis, etc.) as well as on related fields that could aid in the implementation of EL teaching and/or learning (such as computers, English literature, scientific research methods, testing and evaluation, materials development, etc.). The development of 21st-century skills that may be strongly linked to successful ICT use can naturally be anticipated to benefit from and be assisted by these types of courses. There are 18 total technical courses throughout the whole curriculum (49 credits-75 ECTS). Sixteen instructional courses make up the entire program's total (59 credits-81 ECTS).

Since the CoHE revised the curricula of the education faculties in 2018 to be responsive to the shifting demands and needs of the social, educational, and political domains as well as local, national, and international requirements, the current ELT program has been in use. The new program comprises a great number of obligatory and a few elective courses. The program's components include general knowledge, field knowledge (linguistic competence), teacher education (pedagogic competence), and teaching practice. The new curriculum (Seferoğlu, 2006) appears to place greater emphasis on teaching technique and practice components than the previous one, which was in operation from 1998 until 2006. The courses have a total of 175 class hours, of which 143 hours are spent on theory-based coursework and 32 hours on practice-based coursework such as teaching techniques, computer skills, unique teaching methods, and so on.

Essential Characteristics of English Language Teachers

According to Nunan (2003), an ideal teacher should have the following four types of essential qualities: a general level of education, subject competence (a certain level of English proficiency is required), professional competence (such as planning and management skills), and positive attitudes and beliefs. In addition, Selvi (2010) mentioned cultural competency among the teaching competencies and subject skills should be attained to a greater degree before joining the faculty so that trainees may focus more on learning how to teach English and spend less time studying English.

Shulman (1987) recommends creating three more knowledge domains in addition to these three, namely knowledge of learners and their characteristics, knowledge of educational environments, and understanding of educational purposes (as cited in Grenfell & Jones, 2003). Technology should be incorporated into the curriculum, according to Nunan (2003). In this approach, teacher candidates might be informed about contemporary technology and their instructional applications. Nevertheless, the curriculum (Karakas, 2012) covers these areas of expertise excessively.

The distribution of pedagogical and linguistic courses may change because only three elective courses are available at the institutions' discretion, but language assessment courses are unaffected because elective courses are primarily made up of field-related subjects. As a result, the general distribution of the courses in the standardized program would remain roughly the same. If more technologically informed educational processes are incorporated in the courses that will teach both innovative knowledge of the subject and how these may be embedded in creative classrooms, teachers can increase both their pedagogical abilities and also their technical skills.

Some Studies on Educational/Language Program Evaluation

Early studies focused on debating certain ideological, social, and political issues, helping aspiring language teachers find the best and most suitable training programs, or creating new models that adopted preexisting program evaluation models and modified them for the evaluation of educational and language programs (e.g., Collins, 1992; Grosse & Benseler, 1991; Lynch, 1990). Interest in this field has grown as the potential of evaluation and assessment has been identified and proven in educational environments. There has been an enormous amount of research into the evaluation of educational programs, both internationally (e.g., Angell et al., 2008; Dunworth, 2008; Fox & Diaz-Greenberg, 2006; Harris, 2009; Lee et al., 2008; Llosa & Slayton, 2009; Lozano et al., 2002; Lozano et al., 2008 Luke & Britten, 2007; Peacock, 2009; Rivera & Matsuzawa, 2007; Rolstad et al., 2005; Romeo & Dyer, 2004; Sullivan, 2006) and in Turkey (e.g., Biyik, 2007; Coskun, 2009; Er, 2006; Erdem, 2009; Güven & Demirhan Iscan, 2006; Oguz, 2009; Uslu, 2006; Yildiz, 2003; Yilmaz, 2005; Zehir Topkaya & Kücük, 2010). None of the aforementioned studies, however, looked at the ELTTP from the perspective of the current study, that is, whether the program meets and satisfies the needs of the teacher trainees and to what extent each course in the ELTTP has been helpful in terms of the sufficiency and efficiency of the lecturers/instructors, the contents of the given courses, and their practicality. In order to fill this vacuum in the literature, the current study explores these elements from the perspective of the students.

Curriculum Innovation and ICT Integration in Language Assessment

Turkey made various changes to its educational system to improve educational monitoring. ICT integration into instruction was one of the main goals of the 2005-launched curriculum reform initiatives, which attracted the attention of numerous stakeholders while significantly altering the whole national curriculum. One language testing and assessment (LTA) course is being offered as a required course at the eighth semester of the fourth year of undergraduate education, and a measurement and evaluation course (MEC) related to general testing and assessment in education is being taught in Turkish in the sixth semester, according to the curriculum created and revised in 2009 by the CoHE in Turkey.

Currently, it appears that there is no parallelism between the semesters in which the ELTEC and MEC are taught when looking at the course catalogs of EL teacher education programs in Turkey. Intriguingly, some universities (such as Bahçeşehir and Maltepe Universities) only include a measurement and evaluation course taught in Turkish, where terms and principles of testing and assessment in education are taught in general, not specifically relating them to language teaching and learning, while other universities (such as Boğaziçi and Istanbul Bilgi Universities) offer specific language testing and assessment courses. The language assessment literacy development of pre-service EFL teachers and the teacher educators who are in charge of preparing these future EFL teachers to conduct accurate and appropriate assessment practices when they start their careers are likely to face difficulties as a result of these inconsistencies among the EL teacher education programs in Turkey. In order to fully understand ELTEC in Turkish EL teacher education programs, it is important to include the perspectives of ELTEC instructors and pre-service EFL teachers. This will help to create a thorough image of the organization.

Aiming to fill this gap, we initiated an elective course titled "Computer-Assisted Educational Measurement" that can be modeled by other ELT programs. This study has two purposes: a) Presenting curriculum development stages for the program and b) conducting an impact analysis to revise the curriculum and take remedial actions if needed. To address these objectives, the following research questions were examined:

***Pre-course questions**:*

1. To what extent were the student teachers aware of the various digital tools in language assessment at the beginning of the course?

2. What are the student teachers' perceptions regarding the potential drawbacks of using computers and digital tools in order to design language assessments?

3. What are the initial survey results suggested for the elective course curriculum? How is the final course curriculum that was designed based on stakeholder data?

***Post-course questions**:*

4. Which digital assessment tools do the student teachers plan to use in their future teaching?

5. What does the student teachers' survey data suggest in terms of student perception change through the semester?

6. What remedial actions should be taken on the course curriculum and activities for a similar course?

Methods

Research Design

This study employed a single case study approach to examine teacher candidates' experiences on a newly developed course. Case studies help researchers deeply understand similar cases such as a class, course, school, or a community (Cohen et al., 2007). The study was conducted over the course of the

2021-2022 academic year after obtaining ethical approval from the institutional research board. The course titled "Computer-Assisted Educational Measurement" was offered for the first time as an elective course in the ELT program at a public university that can be modeled by similar ELT programs in the future. The course's main objective was to train ELT teacher candidates in using technology while assessing second language skills effectively in different language domains.

Participants

The study included 26 sophomore (equivalent to the 3rd year in the program) ELT teacher candidates who were identified through convenience sampling at a public university. They were admitted to the program based on a nationwide standardized large-scale assessment. Despite not having previously enrolled in any digital literacy or technology courses within the program, the participants successfully completed the fundamentals of assessment course, which is a standard, mandatory course offered in College of Education programs in Turkey. The convenience sampling approach was chosen due to its suitability for the research context. This particular approach was selected because one of the researchers had the opportunity to offer the course within the ELT program with the aim of designing an innovative and need-based course. Thus, the researchers were able to access a readily available pool of participants who were directly involved in the program, enhancing the relevance and practicality of the study's findings. 11 (42.3 %) of the students were female and 15 (57.7 %) of the students were male. Before the semester started officially, all sophomore ELT teacher candidates in the department were invited to participate in a pre-course survey. After analyzing survey data, the researchers finalized and announced the course syllabus to the faculty's course offering catalog. The course syllabus is presented in Appendix A. All the sophomore ELT students (n=26) registered for the elective course. In addition, one student from the Psychological Counseling Department registered for the course due to his personal interest in the subject matter.

Course Procedure

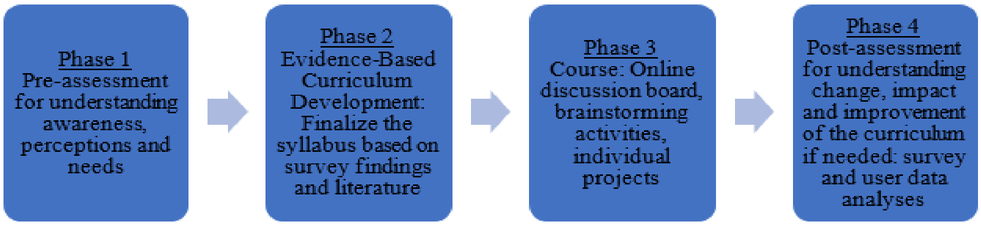
The class met weekly at a regular time through Zoom (2022) since it allowed file-sharing, group activity in small rooms, interaction with peers and with the instructor as well as effective video-conferencing. The course instructor, whose expertise is assessment and measurement, offered this specific class for the first time due to the existing gap in ELT teacher education programs. The class was held over 13 weeks by following a traditional flipped learning approach. Each week, the instructor posted course materials that included short tutorials, videos, articles, and a discussion board question relevant to the associated week's topic. The discussion board activity consisted of an integral part of the formative assessment in the class since it promoted students to conduct their own research before coming to class, reflect on the topic and interact with their peers. The rubric that was used to evaluate students' performance on the discussion board is presented in Appendix B. The synchronous course activities included Kahoot and Edmodo quizzes, pools over Edmodo, and brainstorming activities on Zoom, followed by the weekly lecture.

Data Collection and Analysis

The data collection and analysis procedures can be dissected into four phases as presented in Figure 1. First, students' initial perceptions, needs and awareness were examined through a pre-course survey and an initial discussion board activity. Their course participation, engagement with the subject matter and changing expectations were tracked through weekly discussion board activities, two large-scale projects, and self-evaluation activities. In the last phase of the study, course effectiveness, student change and prospective updates on the curriculum were evaluated through post-course surveys and learning management system (LMS) user data. The details of each study phase are presented below.

Figure 1

Phases of the Study



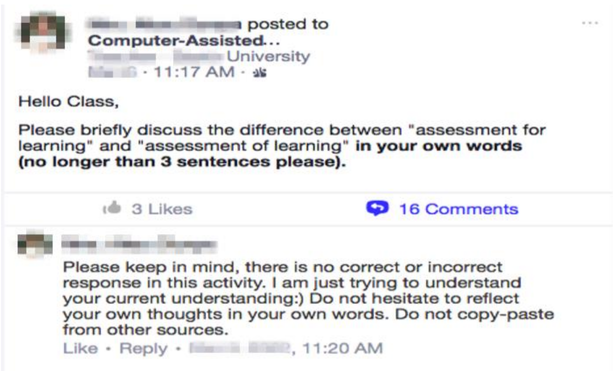
Diagnostic Assessment for Evidence-Based Curriculum Development: Data Collection and Analyses

A survey was designed and implemented before the semester started to understand students' awareness, perceptions and needs in terms of computer-assisted language assessment. The survey consisted of a total of 26 rating-scale items asking about students' previous computer-based assessment (CBA) experiences and expectations from the course. The design of the survey was informed by existing literature on computer-assisted language assessment in order to determine the key elements of the concept. Once the survey items were developed, they underwent expert reviews conducted by two specialists, including a measurement expert and a language expert. These efforts contributed to the content validity evidence.

The data was collected by the teaching assistant through the course LMS before the instructor introduced herself in order to eliminate any social-acceptability bias effect. Data analyses included the calculation of item-by-item descriptive statistics. In addition, a discussion board activity was held before the first week in order to understand what some key concepts such as CBA and assessment for learning (AfL) refer to the students, as seen in Figure 2. The qualitative data obtained through the discussion board was content-analyzed in order to enrich the survey results. The trustworthiness of qualitative data is a crucial aspect and can be ensured through various methods, including triangulation (Stahl & King, 2020). In this study, triangulation of qualitative data sources was employed to establish consistent and identifiable patterns related to the research questions. By observing similar outcomes across multiple data sources, such as discussion board data, self-assessment data and open-ended survey data, it was possible to enhance the trustworthiness of the findings. The triangulated results from the diagnostic assessment and the open-ended survey data helped the finalization of the course syllabus.

Figure 2

First Discussion Board Query for Finalizing Syllabus



Formative Assessment: Data Collection and Analyses

Students' progress and changing perceptions as measures of course effectiveness were tracked through a weekly discussion portal, an extensive midterm project on Padlet and self-assessment. Discussion board data was extracted to MaxQDA (VERBI Software, 2019) and analyzed using content analysis. A midterm project was completed on Padlet by each student and their performance was evaluated based on an analytical rubric with 3 performance criteria and 4 number of performance levels (see Appendix B). The development of the rubric was guided by the course objectives and further validated through an external expert review from the assessment field. Students also completed a self-assessment form (see Appendix C) to elaborate on their own progress critically. Self-assessment data was collected periodically and analyzed using MaxQDA as supplemental evidence for RQs 4, 5, and 6.

Summative Assessment: Data Collection and Analyses

Data was collected via a post-course survey and an extensive semester project for the summative assessment part of the study. Students responded to the survey, which was composed of 27 rating scale items and three open-ended items. In a similar manner to the construction of the pre-course survey, expert reviews from the same two experts were employed to ensure content validity. The final project required each student to prepare a 12-week syllabus for an imaginary technology-enhanced language course that could be taught in K-12. The students were asked to design their syllabus so that each week would include at least one computer-based formative assessment activity (i.e., discussion board, forum, presentation, self or peer assessment using technology) embedded into the class. Analyses of survey data using descriptive statistics (i.e., frequencies, percentages) and data visualization packages (“ggpubr”, Kassambara, 2020; “ggplot2”, Wickham, 2016) in R (R Core Team, 2022) consisted of the evidence for student change, improvement, and course effectiveness.

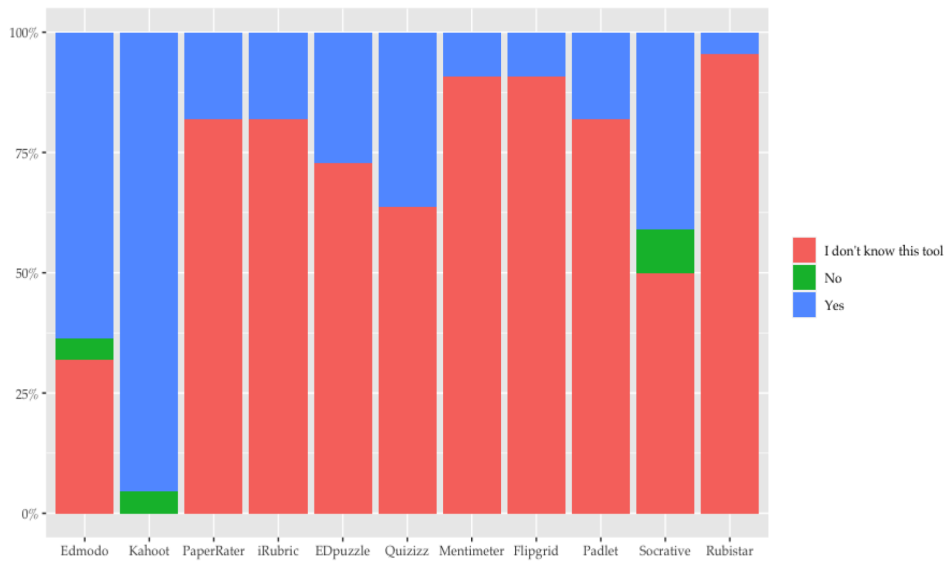
Results

Results with Respect Pre-Course Questions

The first research question was examined based on survey data (n=22) and the results were summarized in Figure 3 below. As seen in Figure 3, most students were aware of Kahoot (95%) and Edmodo (63%) as tools for digital language assessment. However, tools such as Rubistar (95%), Flipgrid (91%) and Mentimeter (91%) were known by almost none of the students despite their practicality in formative language assessment and rubric design.

Figure 3

Students’ Digital Tool Awareness at the Beginning of the Course



Students' perceptions regarding computer-based language assessment were further elaborated through a discussion board activity where the instructor asked the students to discuss the potential drawbacks of CBA in their context (RQ2). As presented in Table 1, the teacher candidates' discussion data revealed four initial codes: computer literacy, software-hardware issues, cheating, and lack of interaction.

Table 1

Perceived Drawbacks of CBA

Themes	Codes	Keywords
Differential opportunity	Computer literacy	Judgmental peers, experience, familiarity
	Software and hardware issues	Power shortage, rural areas, connectivity, camera, and microphone
	Random error	Easier cheating
Traditional habits	Lack of interaction	Human-mediated assessment, human-by-human interaction

As summarized in Table 1, the identified concerns can be categorized into two primary themes: disparate access and traditional practices and habits. Participants expressed their concerns regarding uneven access to technology, varying levels of digital literacy, and potential sources of error that pose risks to online assessments, such as cheating. Additionally, they traditionally viewed human-mediated assessments as more beneficial than computer-mediated assessments.

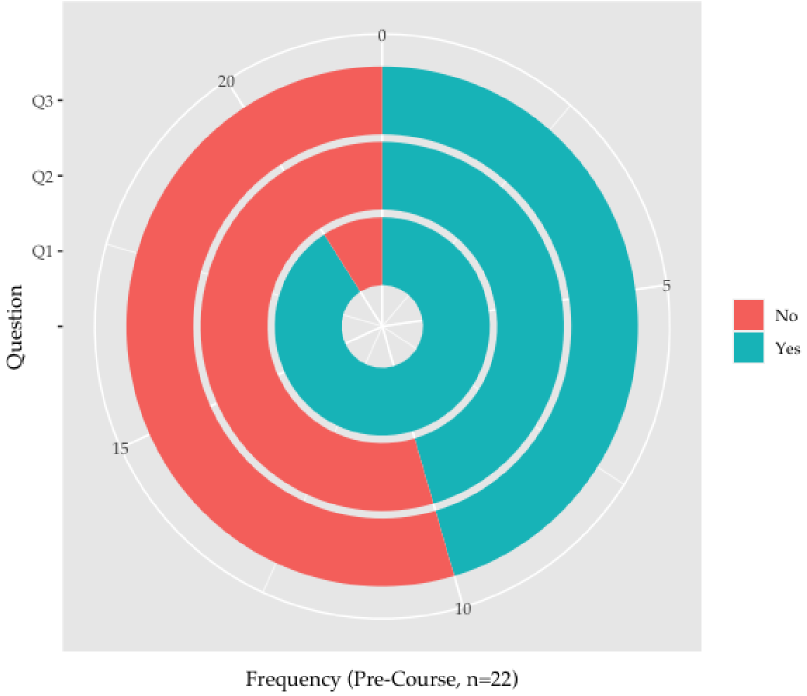
Specifically, teacher candidates presented their worries in computer-assisted language assessment due to variations in their access to and readiness to utilize technology. Furthermore, they perceived computer-assisted language assessment as challenging to implement across various language domains, primarily due to the inherent nature of the construct being measured. One student stated that:

S5: "The language learning and assessment process, which is especially communication-oriented, will still be a disadvantage with computers since there is no interaction as in the classroom."

Students' responses to the potential drawbacks alerted the instructor to introduce various computer-based tools for assessing different language skills (especially in more interactive skills such as speaking and listening) and help students experience them throughout the course. The third research question was examined first by elaborating students' previous experiences and formal preparation related to computer-assisted assessment. As seen in Figure 4, almost 91% of the course-takers had taken a computer-based language exam earlier. Yet, more than half of them (54.5%) had not taken any assessment-related course before.

Figure 4

*Students' Previous Assessment-Related Experiences**



*Q1: Have you ever taken any computer-assisted/based exam (e-YDS, TOEFL, online test etc.) ?

Q2: Have you ever taken any assessment or measurement-related course, seminar or webinar?

Q3: Have you ever used any computer-assisted/based assessment tool for your teaching?

Despite lack of enough formal preparation in terms of assessment and measurement, the discussion board data revealed that the students were able to define CBA appropriately:

S1: "The delivery and assessment processes completed through computers"

Few students viewed CBA as a cure for pandemic conditions:

S2: "Computer-based assessment should be used for online education, especially during a pandemic."

while a large number of students stated that computer-based language assessment is a necessity for 21st century teachers:

S3: "In order to grow as skilled prospective teachers, learning skills for conducting computer-based assessment is crucial for our generation of teachers, so it represents a necessity to me."

Another discussion board activity that was designed to deeply investigate students' understanding of the purpose of formative assessment revealed that students had a novice understanding of these two distinct concepts. Some students defined AfL as a process that could only be conducted during class time while only few students were able to distinguish their purposes appropriately:

S4: "Assessment for learning is done throughout the teaching period to identify necessary improvements and adjustments and to observe students' progress. Assessment of learning is done at the end of the teaching period to evaluate students' understanding of the taught subject"

Given the evidence based on both survey and discussion board data, the researchers concluded that the students were aware of the concept of CBA although many of them perceived that CBA could be used for summative assessment solely. In addition, students had noticeable misconceptions on AfL and formative assessment concepts. This finding urged the researchers that the syllabus must have emphasis on the use of computer-based/assisted strategies for formative assessment and AfL in language classrooms.

As a result of the interpretation of multiple data sources, four stages were planned to form the syllabus. In the first stage, students were going to be prepared to use CBA in their future teaching by presenting the rationale to adopt computers for assessment. Then, students were going to be trained to develop technological as well as assessment literacy through activities, group projects, self-assessment, and systematic feedback. Each week in the syllabus included hands-on activities for this purpose. Then, two extensive projects (one midterm and one final) were added to the syllabus to transform the students into confident users of computers for assessing different language skills by considering methods to ensure validity and reliability. Therefore, a student who completed this course would be a competent user of technology who could integrate assessment and digital skills for designing effective language assessment. Table 2 provides an overview of the primary tools and technologies employed in the course content, aligning them with the designed rubric. These tools significantly enhanced the assessment capabilities, particularly in terms of formative assessment/AfL, as highlighted in Table 2.

Table 2

Major Tools Utilized for AfL in Class

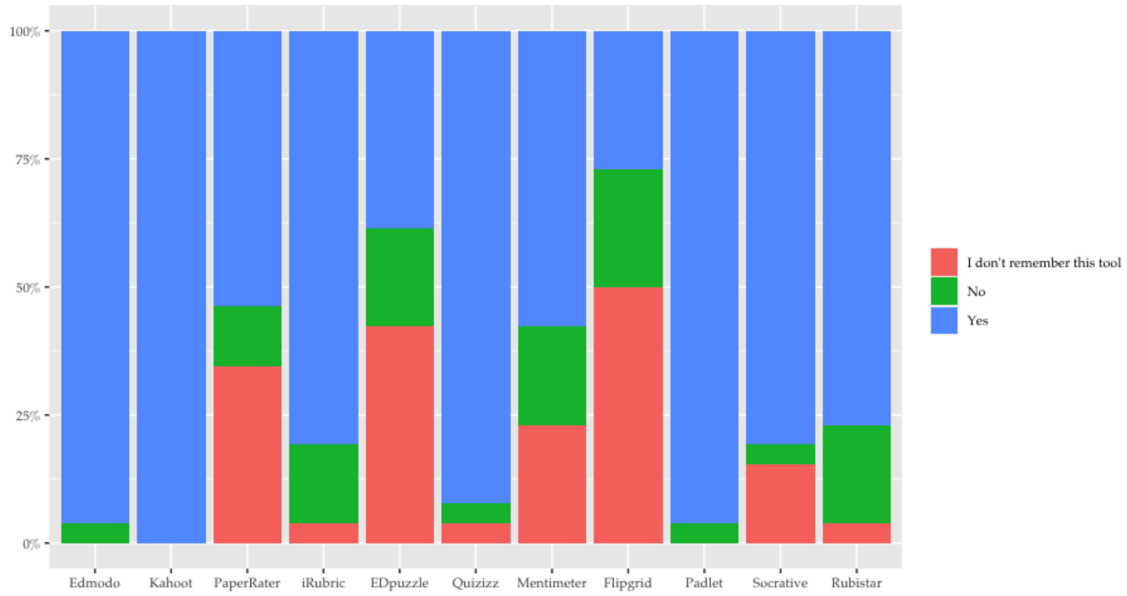
Tool	Which aspect of formative assessment was the tool utilized for?
Edmodo	Discussion board activity, self and peer assessment
Kahoot	In-class quizzes, end-of-class recap, gathering feedback and course evaluations
PaperRater	Peer and self-assessment of writing
iRubric	Creating and editing rubrics for performance assessments
EDpuzzle	Conducting quizzes embedded into instructional videos
Rubistar	Creating customizable rubrics for performance assessment
Padlet	Creating personal and group digital boards, , creating e-portfolios, brainstorming, discussion board activity, peer assessment
Mentimeter	Live quizzes, creating word clouds and concept maps,

Results with Respect to Post-Course Questions

The students were exposed to various Web 2.0 tools, websites, and computer programs to design sound assessments in different language skills during the semester. As seen in Figure 5, among the tools practiced throughout the semester, students found Padlet (96.2%) the most effective, followed by Quizizz (92.3%) and iRubric (80.8%). It is noticeable that before taking the course, most of the students were not aware of the tools that were found the most effective (see Figure 5) at the end of the course (n=26).

Figure 5

Students' Perceptions on the Effectiveness of the Tools



The RQ5 (What does the students' survey data suggest in terms of student perception change through the semester?) was examined through pre and post-survey results and participation statistics for each formative assessment activity. The same set of 13 survey questions was presented to the students before and after the semester. The common questions in both survey forms can be seen below in Table 3.

Table 3

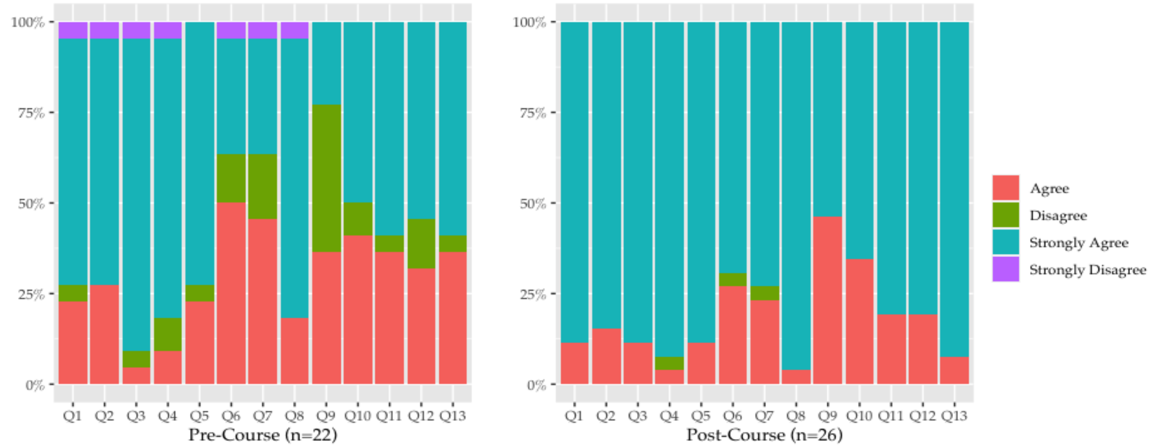
Survey Items

#	Item
1	Using the computer-assisted/based assessment will improve my work.
2	Using the computer-assisted/based assessment will enhance my effectiveness.
3	I could complete an exam or homework task using the computer.
4	I could complete an exam or homework using the computer if someone showed me how to do it first.
5	I can navigate easily through the Web to find any information I need.
6	Computer-assisted/based assessment tools that I know have clear instructions.
7	Computer-assisted/based assessment questions that I have seen are easy to answer.
8	Computer-assisted/based assessment tools that I know will be useful for my teaching.
9	My personal preparation for the computer-assisted/based is sufficient.
10	My performance expectations for the computer-assisted/based assessment are high.
11	Using computer-assisted/based assessment tools (online exams, online assignments, etc.) makes my learning enjoyable.
12	Using computer-assisted/based assessment stimulates my curiosity.
13	Using computer-assisted/based assessment will lead to my exploration.

Figure 6 depicts the change in students' perceptions throughout the semester. The largest change occurred on questions thirteen (Using computer-assisted/based assessment will lead to my exploration) and eight (Computer-assisted/based assessment tools that I know will be useful for my teaching). The general tendency of the change occurred towards positive regarding computer-based assessment, as seen in Figure 6.

Figure 6

Students' Perceptions Before and After the Semester



An important objective of designing this study and collecting data from different sources was to update the syllabus if needed. RQ6 (What remedial actions should be taken on the course curriculum and activities for a similar course?) served for this purpose. Analyses of discussion forums and open-ended survey data revealed that students found the course extremely effective, useful and practical for learning time-saver applications. They suggested that integrating more applications and tools for language assessment into the syllabus would improve the course content:

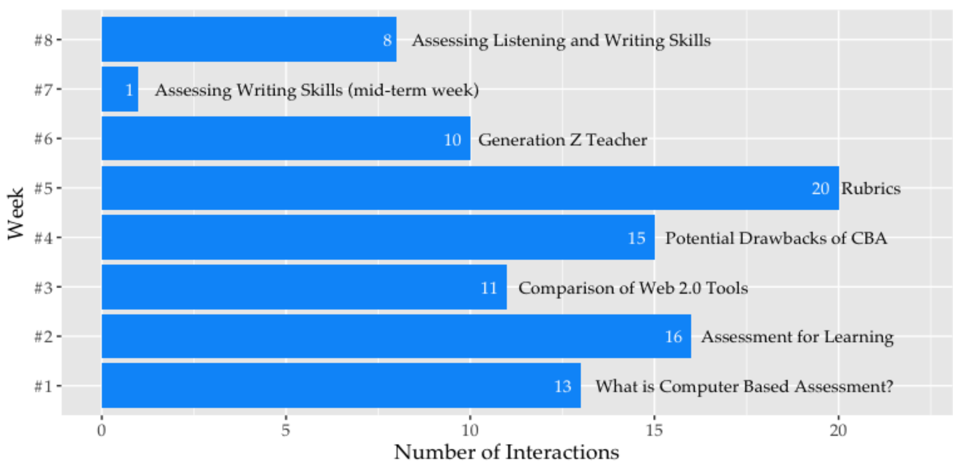
S7: "I will have the opportunity to create reliable activities for the students where both they are engaged and have fun, and I have the ease of evaluation, which is quite practical and time-saving."

S8: "The course would be improved by Integrating more tools, applications or websites to use. "

The discussion board data, as summarized in Figure 7, revealed that students' most interacted with the topic of rubrics as followed by AfL concepts. Therefore, these two topics should be covered more in the syllabus. Given the students' misconceptions on AfL at the beginning of the course and their interest in the topic on virtual discussion, it should be concluded that there would be a separate week on the syllabus allocated for the theory and applications of the AfL concept. That particular week might include strategies for AfL in language classrooms as well as similarities and differences between formative assessment, summative assessment and AfL.

Figure 7

Students' Interaction with Discussion Topics



Discussion

Assessment and Measurement courses are mandatory in Teacher Education programs in Turkey. A standard, theory-dominated curriculum was used for every undergraduate program although each program, particularly the applied ones such as ELT, Physical Education Teaching, Visual Arts Teaching are supposed to have differing needs and strategies in terms of assessment. As Atman Uslu (2022) stated, despite the latest changes in the curriculum of ELT in 2018, there is still a gap fulfilling the needs of the 21st century language teacher candidates' digital assessment literacies. The present courses are very limited in number and still lack a technology-integrated curriculum and practice for assessment. To fill this gap, the purpose of this study is to design an evidence-based second language assessment curriculum, apply it and evaluate its effectiveness through various data types.

The goal of the course was to help student instructors comprehend the advantages of implementing ICT in their future English classes. In order to promote technology-assisted assessment courses in EFL instruction, the course's student teachers were supposed to develop a variety of technology capabilities and pedagogical knowledge. The fact that they themselves had a positive experience with technology in the course of their language learning may be one of the most significant reasons for prospective teachers to incorporate technology as a part of their future teaching practice. In order to help Turkish EFL student teachers successfully apply their unique language testing experiences to their future careers as EFL teachers, this course attempted to first give students valuable personal language assessment experiences using ICT.

This study aims to offer both a local view on language assessment literacy (LAL) implementation in the Turkish EFL educational context and a worldwide perspective on LAL. In this study, Turkish EFL student teachers who had taken technology-integrated courses showed a desire to use novel teaching methods going forward. The students, teachers in particular, could recognize that their experience in their language classes motivates them to integrate technology in their actual teaching environments. This finding supports Arslan (2023) in that using digital tools in their teaching and assessment practices motivated students. They could also understand how educational technology has affected how languages are learned and taught (Schmid & Hegelheimer, 2014). The possibility and driving forces for future change for LAL would be provided by their views on the effects of integrating ICT and digital technologies during the prospective teacher preparation period (Atman Uslu, 2022; Jiang, 2021; Schmid & Hegelheimer, 2014).

As Arslan (2023) states, more possibilities should be provided for aspiring teachers to use ICT in the classroom. To successfully integrate technology, future teacher educators in EFL subjects need to update university curricula. This finding is in line with Lei (2022) since he claims that technology integration is crucial in teacher education curricula. Future teachers should learn how and why to employ technology to better their own language learning and future instruction in prospective teacher education courses (Arslan, 2023; Atman Uslu, 2022; Compton, 2009; Jiang, 2022). Institutions of higher learning that offer bachelor's degrees in education ought to include at least a training course in the foundations of educational technology because incorporating technology into classroom instruction can aid EFL students (Arslan, 2023; Atman Uslu, 2022; Masood, 2010).

Despite the ICT progress in Turkey, student instructors in the EFL educational context do not receive enough educational preparation about technology applications in their future teaching. This is in line with Uzun (2016) in that in today's technological world, it has utmost importance for the teachers to be digital literate. Most of the faculty in the department of EFL teacher education are not digital natives; hence, neither are they capable of using ICT nor are they interested in doing so (Arslan, 2023; Uzun, 2016). Additionally, during their years of academic study, student teachers themselves rarely encountered ICT integration (Uzun, 2016). For them to effectively employ technology in their future teaching, it was crucial to inspire and train them.

This course attempted to assist aspiring educators in preparing for their upcoming students who are digital natives by giving them valuable and instructive experience. This study may have ramifications for how this technology-integrated EFL student-teacher preparation course can identify and close the gap between the present and future state of education and testing in terms of the use of technology in Turkish EFL classrooms. Turkish EFL student instructors may understand how technology has improved their English language learning and testing through the technology-integrated teacher training course, and these experiences and beliefs would then be incorporated into their future classrooms (Jeong, 2017).

The current syllabus was found effective in increasing students' technology-enhanced assessment skills and changing their perceptions, as evidenced by student data. However, it should be noted that the syllabus must be updated periodically since applications, tools, and technologies (i.e., natural language processing) are continuously developing/changing. For example, Edmodo, which has been a very popular and convenient tool for almost 15 years, has been shut down permanently as of September 2022 (Edmodo, 2022).

One key issue that affected students' perceptions of CBA was differential access to technology. The most recent literature on assessment in the pandemic also supports this finding (Kim & Padilla, 2020). Students who had inadequate access to technology, internet and computers struggled most in assessment during the pandemic. However, it is a fact that the problem is not related to the pandemic only since the issue was the same before the pandemic as well (Ahn & McEachin, 2017; Center for Research on Educational Outcomes [CREDO], 2015). The teacher education programs do not offer the necessary courses that includes digital assessment literacies. This need necessitated teacher education programs to offer this kind of course to fulfill their needs. Additionally, as students proposed, hardware and software accessibility should be reinforced in and out of the campus because technology-enhanced assessment must consist of two components equally for each student: hardware and adequate training to use them.

Another finding related to the students' concern about adopting CBA is that it is prone to cheating since it lacks real-time human interaction. However, the most recent studies stated that there is no difference in the perceived cheating behavior of the students in online and face-to-face education (Yazici et al., 2022). This empirical evidence challenges the perception that CBA inherently enables more cheating compared to traditional assessment methods. Thus, these findings contribute to alleviating concerns surrounding cheating in CBA by highlighting that it may not be a substantiated issue.

It is important to acknowledge and address certain limitations associated with the findings of the study. Firstly, it is crucial to recognize that this research is based on a case study, which inherently relies on specific contextual factors. Therefore, caution should be exercised when attempting to generalize the findings beyond the scope of this particular study. While the insights gained are valuable within the context in which they were obtained, further research is needed to validate and corroborate these findings across different educational settings.

Furthermore, it is crucial to highlight that the curriculum utilized in this study was newly developed and implemented within the specific institution. As is customary with any implementation, there may be aspects that require further refinement. The experiences and feedback gathered during the initial application of the curriculum will serve as invaluable input for future iterations of the course. Continuous improvements will be made in response to student feedback, advancements in technology, and their suitability for enhancing EFL assessment practices. This iterative process ensures that the curriculum remains responsive and adaptable, keeping pace with emerging trends in the field.

Conclusion

This study has shed light on the implementation of technology-integrated language assessment literacy (LAL) in the English as a Foreign Language (EFL) educational context while also providing a global perspective on LAL. The findings of this research highlight the significance of integrating technology into language teacher education in assessment, as evidenced by the enthusiasm and desire demonstrated by EFL student teachers who had undergone the technology-integrated assessment course.

The practical implications of this study are significant. Firstly, it is imperative for higher education institutions and policymakers in Turkey to recognize the importance of equipping EFL student teachers with adequate assessment literacy and digital skills. By incorporating technology-enhanced assessment courses into teacher education programs, educators can empower future teachers to effectively assess their students' language proficiency and make informed instructional decisions based on accurate and reliable assessment data. Furthermore, providing student teachers with technology-integrated courses offers them an opportunity to explore innovative teaching methods and utilize digital tools to enhance language learning experiences for their future students, particularly in the era of artificial intelligence.

Moreover, the findings of this study have implications for the global educational landscape as well. Language assessment literacy is not confined to the Turkish EFL context alone; it is a crucial aspect of language education worldwide. Therefore, the insights gained from this research can inform educators and policymakers in other countries to enhance the assessment literacy of their language teachers. Integrating technology into assessment courses can be a valuable strategy to foster the development of assessment literacy, as it not only enhances teachers' assessment competencies but also encourages them to embrace novel assessment approaches that promote active engagement and foster meaningful learning experiences for students.

Declarations

Conflict of Interest: The authors report there are no potential conflicts of interest to declare.

Ethical Approval: This study was approved by the Bartın University Rectorate Board of Ethics for the Social Sciences and Humanities (Date: 09.02.2023, Protocol number: 2023-SBB-0048).

References

Ahn, J., & McEachin, A. (2017). Student enrollment patterns and achievement in Ohio's online charter schools. *Educational Researcher*, *46*(1), 44-57. https://doi.org/10.3102/0013189X17692999

Alderson, J. C. (2000). Technology in testing: The present and the future. *System*, *28*(4), 593–603. https://doi.org/10.1016/S0346-251X(00)00040-3

Angell, J., DuBravac, S., & Gonglewski, M. (2008). Thinking globally, acting locally: Selecting textbooks for college-level language programs. *Foreign Language Annals*, *41*(3), 562–573. https://doi.org/10.1111/j.1944-9720.2008.tb03312.x

Arslan, A. (2023). Determining the Effect of University Students' Technology Usage Levels on Digital Reading Self-Efficacy. *Bartın University Journal of Faculty of Education, 12*(1), 73-84. https://doi.org/10.14686/buefad.993603

Atman Uslu, N. (2022). Exploring Pre-service Teachers' Perception Regarding Factors in Technology Integration with Q Methodology. *Bartın University Journal of Faculty of Education, 11*(3), 543-558. https://doi.org/10.14686/buefad.1001562

Balıkcıoğlu Akkuş, G. & Altay, İ. F. (2023). Instructor's Written Feedback in Emergency Remote Teaching: EFL Learners' Perspectives. *Bartın University Journal of Faculty of Education, 12* (1), 47-56. https://doi.org/10.14686/buefad.977978

Biyik, C. O. (2007). A preliminary evaluation of the distance English language teacher training program (DELTTP) in Anadolu University, Turkey. *Turkish Online Journal of Distance Education-TOJDE*, *8*(1), 143–162.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research Methods in Education*. Routledge.

Collins, J. (1992). Our ideologies and theirs. *Pragmatics*, *2*(3), 405–415.

Comeaux, P. (2002). *Communication and collaboration in the online classroom*. Anker.

Compton, L. K. (2009). Preparing language teachers to teach language online: A look at skills, roles, and responsibilities. *Computer Assisted Language Learning*, *22*(1), 73-99. https://doi.org/10.1080/09588220802613831

Coskun, H. (2009). Türkiye ve Almanya'da yabancı dil öğretmeni yetiştirme programlarının karşılaştırılması [The comparison of teacher training programmes in Turkey and Germany]. *C.Ü. Sosyal Bilimler Dergisi*, *33*(1), 61–73.

Center for Research on Educational Outcomes (CREDO). (2015). *Online charter school study*. Stanford University. https://credo.stanford.edu/sites/g/files/sbiybj6481/f/online_charter_study_final.pdf

Doğançay-Aktuna, S., & Kızıltepe, Z. (2005). English in Turkey. *World Englishes*, *24*(2), 253-265. https://doi.org/10.1111/j.1467-971X.2005.00408.x

Dunworth, K. (2008). Ideas and Realities: Investigating good practice in the management of transnational English language programmes for the higher education sector. *Quality in Higher Education*, *14*(2), 95–107. https://doi.org/10.1080/13538320802278099

Edmodo. (2022). *About Edmodo*. Retrieved from https://go.edmodo.com/about/

Er, K. O. (2006). Evaluation of English curricula in 4th and 5th grade primary schools. *Ankara University, Journal of Faculty of Educational Sciences*, *39*(2), 1–25.

Erdem, A. (2009). Comparing the language curriculums of Turkey and Ireland. *Procedia Social and Behavioral Sciences*, *1*(1), 529–535. https://doi.org/10.1016/j.sbspro.2009.01.095

Eyal, L. (2010). *The reciprocity between learning-content management system (LCMS) and the assessment of learners in E-learning courses* (Unpublished doctoral dissertation). Bar-Ilan University, Ramat-Gan, Israel.

Fox, R. K., & Diaz-Greenberg, R. (2006). Culture, multiculturalism, and foreign/world language standards in U.S. teacher preparation programs: Toward a discourse of dissonance. *European Journal of Teacher Education*, *29*(3), 401–422. https://doi.org/10.1080/02619760600795270

Greenier, V., Derakhshan, A., and Fathi, J. (2021). Emotion regulation and psychological well-being in teacher work engagement: a case of British and Iranian English language teachers. *System*, *97*:102446. https://doi.org/10.1016/j.system.2020.102446

Grenfell, M., Kelly, M., & Jones,D. (2003). *The European language teacher: recent trends and future developments in teacher education*. Peter Lang.

Gredler, M. E. (1996). Educational games and simulation: A technology in search of a research paradigm. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 521-540). Simon & Schuster Macmillan.

Grosse, C. U., & Benseler, D. P. (1991). *Directory of foreign language teacher preparation programs in the United States: A preliminary report*. In Acting on Priorities: A Commitment to Excellence. Report of Southern Conference on Language Teaching, FL, 20p.

Güven, I., & Demirhan Iscan, C. (2006). The reflections of new elementary education curriculum on media. *Ankara University, Journal of Faculty of Educational Sciences*, *39*(2), 95–123. https://doi.org/10.1501/Egifak_0000000142

Harris, J. (2009). Late-stage refocusing of Irish-language programme evaluation: Maximizing the potential for productive debate and remediation. *Language Teaching Research*, *13*(1), 55-76. https://doi.org/10.1177/1362168808095523

Jeong, K.-O. (2017) Preparing EFL student teachers with new technologies in the Korean context. *Computer Assisted Language Learning*, *30*(6), 488-509. https://doi.org/10.1080/09588221.2017.1321554

Jiang, A. L. (2022). Identity work as ethical self-formation: the case of two Chinese EFL teachers in the context of curriculum reform. *Frontiers in Psychology*, *12*, 6355. https://doi.org/10.3389/fpsyg.2021.774759

Jiang, A. L., and Zhang, L. J. (2021). Teacher learning as identity change: the case of EFL teachers in the context of curriculum reform. *TESOL Quarterly*, *55*(1), 271–284. https://doi.org/10.1002/tesq.3017

Karakas, A. (2012). Evaluation of the English Language Teacher Education Program in Turkey. *ELT Weekly*, *4*(15). Retrieved from http://www.eltweekly.com/elt-newsletter/2012/04/vol-4-issue-15-research-article-evaluation-of-the-english-language-teachereducation-program-in-turkey-by-ali-karakas/

Kassambara, A. (2020). *Ggpubr: 'ggplot2' based publication ready plots*. https://CRAN. R-project.org/package=ggpubr

Kim, C. J. H., & Padilla, A. M. (2020). Technology for educational purposes among low-income Latino children living in a mobile park in Silicon Valley: A case study before and during COVID-19. *Hispanic Journal of Behavioral Sciences*, *42*(4), 497-514. https://doi.org/10.1177/0739986320959764

Lee, Y.-F., Altschuld, J. W., & Hung, H.-L. (2008). Practices and challenges in educational program evaluation in the Asia-Pacific region: Results of a Delphi study. *Evaluation and Program Planning*, *31*(4), 368–375. https://doi.org/10.1016/j.evalprogplan.2008.08.003

Lei, M. (2022). *Teachers' Perceptions of Curriculum Innovation in English Language Teaching in China*. Wuhan: Huazhong University of Science and Technology Press.

Lei, M., & Medwell, J. (2022). The Changing Role of Chinese English-as-Foreign-Language Teachers in the Context of Curriculum Reform: Teachers' Understanding of Their New Role. *Frontiers in Psychology*, *13:*904071. https://doi.org/10.3389/fpsyg.2022.904071

Lei, P., & Xu, J. (2022). Development of college English teachers' psychological empowerment scale: a validation study. *Frontiers in Psychology, 13*:846081. https://doi.org/10.3389/fpsyg.2022.846081

Liang, X., & Creasy, K. (2004). Classroom assessment in web-based instructional environment: Instructors' experience. *Practical Assessment, Research, and Evaluation, 9*(1), 7. https://doi.org/10.7275/84mr-wp41

Llosa, L., & Slayton, J. (2009). Using program evaluation to inform and improve the education of young English language learners in US schools. *Language Teaching Research*, *13*(1), 35–54. https://doi.org/10.1177/1362168808095522

Lozano, A. S., Sung, H., Padilla, A. M., & Silva, D. M. (2002). Evaluation of professional development for language teachers in California. *Foreign Language Annals*, *35*(2), 161–170. https://doi.org/10.1111/j.1944-9720.2002.tb03152.x

Lozano, A. S., Sung, H., Padilla, A. M., & Silva, D. M. (2008). A statewide professional development program for California foreign language teachers. *Foreign Language Annals*, *37*(2), 301–309. https://doi.org/10.1111/j.1944-9720.2004.tb02202.x

Luke, C. L., & Britten, J. S. (2007). The expanding role of technology in foreign language teacher education programs. *CALICO Journal*, *24*(2), 253–267.

Lynch, B. K. (1990). A context-adaptive model for program evaluation. *TESOL Quarterly*, *24*(1), 23–42. https://doi.org/10.2307/3586850

Masood, M. (2010). An Initial Comparison of Educational Technology Courses for Training Teachers at Malaysian Universities: A Comparative Study. *Turkish Online Journal of Educational Technology-TOJET*, *9*(1), 23-27.

Mercader, C., & Gairín, J. (2020). University teachers' perception of barriers to the use of digital technologies: the importance of the academic discipline. *International Journal of Educational Technology in Higher Education*, *17*(1), 4. https://doi.org/10.1186/s41239-020-0182-x

Muslem, A., Yusuf, Y. Q., & Juliana, R. (2018). Perceptions and barriers to ICT use among English teachers in Indonesia. *Teaching English with Technology, 18*(1), 3-23.

Nunan, D. (2003). The impact of English as a global language on educational policies and practices in the Asia-Pacific Region. *TESOL Quarterly, 37*(4), 589-613. https://doi.org/10.2307/3588214

Oguz, A. (2009). An evaluation of suitability of practices in teacher training programs for the constructivism depending on the teacher trainees' views. *Ankara University, Journal of Faculty of Educational Sciences*, *42*(1), 129–155. https://doi.org/10.1501/Egifak_0000001141

Ornstein, A. C., & Hunkins, F. P. (2014). *Curriculum: Foundations, principles and theory*. Sage.

Peacock, M. (2009). The evaluation of foreign-language-teacher education programmes. *Language Teaching Research*, *13*(3), 259–278. https://doi.org/10.1177/1362168809104698

Posavac, E. J., & Carey, R. G. (2003). *Program evaluation methods and case studies* (6th ed.). Prentice Hall.

R Core Team (2022). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria. https://www.R-project.org/

Rivera, G. M., & Matsuzawa, C. (2007). Multiple-language program assessment: Learners' perspectives on first- and second-year college foreign language programs and their implications for program improvement. *Foreign Language Annals*, *40*(4), 569–583. https://doi.org/10.1111/j.1944-9720.2007.tb02881.x

Rolstad, K., Mahoney, K., & Glass, G. V. (2005). The big picture: A meta-analysis of program effectiveness research on English language learners. *Educational Policy*, *19*(4), 572–594. https://doi.org/10.1177/0895904805278067

Romeo, K., & Dyer, M. (2004). Adapting to changing needs: A teacher-led Japanese FLES program. *Learning Languages*, *10*(1), 10–15.

Schmid, E. C., & Hegelheimer, V. (2014). Collaborative research projects in the technology-enhanced language classroom: Pre-service and in-service teachers exchange knowledge about technology. *ReCALL*, *26*(3), 315-332. https://doi.org/10.1017/S0958344014000135

Seferoğlu, G. (2006) Teacher candidates' reflections on some components of a pre-service English teacher education programme in Turkey. *Journal of Education for Teaching*, *32*(4), 369-378. https://doi.org/10.1080/02607470600981953

Selvi, K. (2010). Motivating factors in online courses. *Procedia - Social and Behavioral Sciences, 2*(2), 819-824. https://doi.org/10.1016/j.sbspro.2010.03.110

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, *57*(1), 1-23. https://doi.org/10.17763/haer.57.1.j463w79r56455411

Stahl, N. A., & King, J. R. (2020). Expanding approaches for research: Understanding and using trustworthiness in qualitative research. *Journal of Developmental Education*, *44*(1), 26-28.

Sullivan, J. H. (2006). The importance of program evaluation in collegiate foreign language programs. *The Modern Language Journal*, *90*(4), 590–593.

Tao, J., and Gao, X. (2017). Teacher agency and identity commitment in curricular reform. *Teaching and Teacher Education, 63,* 346–355. https://doi.org/10.1016/j.tate.2017.01.010

Uslu, Z. (2006). Almanca Öğretmenliği bölümlerinde secmeli kırk kredilik İngilizce dersinin etkileri [The effect of the elective forty-credits English course in the German language teacher training departments]. *Uludağ Üniversitesi Eğitim Fakültesi Dergisi*, 19, 415–430.

Uzun, L. (2016). The educational and technical courses in the ELT program in Turkey: Do they contribute to ICT skills? *Cogent Education*, *3*(1), 1-12. https://doi.org/10.1080/2331186X.2016.1141454

Vähäsantanen, K. (2015). Professional agency in the stream of change: Understanding educational change and teachers' professional identities. *Teaching and Teacher Education, 47*, 1-12. https://doi.org/10.1016/j.tate.2014.11.006

Vähäsantanen, K., and Eteläpelto, A. (2009). Vocational teachers in the face of a major educational reform: Individual ways of negotiating professional identities. *Journal of Education and Work, 22*(1), 15-33. https://doi.org/10.1080/13639080802709620

Yang, H. (2015). *Teacher mediated agency in educational reform in China*. Springer International Publishing. https://doi.org/10.1007/978-3-319-15925-6

VERBI Software. (2019). *MAXQDA 2020* [computer software]. Berlin, Germany: VERBI Software. Available from https://maxqda.com

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Yang, Y. T. C., Chen, Y. C., & Hung, H. T. (2022). Digital storytelling as an interdisciplinary project to improve students' English speaking and creative thinking. *Computer Assisted Language Learning*, *35*(4), 840-862. https://doi.org/10.1080/09588221.2020.1750431

Yazici, S., Yildiz Durak, H., Aksu Dünya, B., & Şentürk, B. (2022). Online versus face-to-face cheating: The prevalence of cheating behaviours during the pandemic compared to the pre-pandemic among Turkish University students. *Journal of Computer Assisted Learning,* 39(1), 231-254. https://doi.org/10.1111/jcal.12743

Yildiz, C. (2003). Yabancı dil öğretmeni yetiştirme programında yeni yapılanma (Dersler-İçerikler: Eleştirel bir yaklaşım) [A new construction in the foreign language teacher training programme (Lessons-Contents: A Critical Approach)]. *Dil Dergisi*, *121*, 7-21.

Yilmaz, C. (2005). How far are the teacher trainees' communicative needs are taken into account in ELT departments? *Ondokuzmayıs Üniversitesi Egitim Fakültesi Dergisi*, 19, 92–103.

Zehir Topkaya, E., & Kücük, Ö. (2010). An evaluation of the 4th and 5th grade English language teaching program. *Elementary Education Online*, *9*(1), 52–65.

Zoom Video Communications, Inc. (2020). *ZOOM cloud meetings* (Version 5.11.10). Available from https://zoom.com

Appendix A

Computer-Assisted Educational Measurement

Course Code: GKS-021

Instructor: XXX

Course Objectives: Developing students’ skills in digital assessment tools and computer-assisted language testing approaches that can be utilized in technology-enhanced ELT classrooms.

Week	Topic	Pre-Course and In-class Activities
1	Introduction to the course. Warming up to course materials and syllabus Introducing Edmodo (Sign-up, usage)	Sign up for the course on Edmodo
2	Introduction to language assessment: assessment types by purpose, formative and summative assessment methods	Participate in Edmodo Discussion
3	An overview of strategies for using technology while measuring language skills: Online quizzes, discussion boards, mind mapping, peer and self-assessment	Participate in Edmodo Discussion In-class: Develop an online quiz on Kahoot
4	Benefits and drawbacks of computer-based assessment Large-scale computer-based language testing programs: TOEFL iBT, IELTS, DuoLingo, PISA Foreign Language Assessment, Global Test of English Communication for Students	Participate in Edmodo Discussion
5	Performance assessments and rubrics: Rubric types, developing and revising rubrics using technology: ● Computerized Rubric Building Tools: i.e., RUBISTAR, iRubric	Participate in Edmodo Discussion
6	Assessing reading skills in computer-assisted environments: ● Web 2.0 Tools: i.e., Socrative, Mentimeter, Quizziz, Padlet, Edpuzzle	Participate in Edmodo Discussion
7	Assessing writing skills in computer-assisted environments: ● Web 2.0 Tools (i.e., Padlet, Blogger)	Participate in Edmodo Discussion

	• Artificial Intelligence and Writing Assessment: Paperrater	
8	Assessing listening skills in computer-assisted environments: • Web 2.0 Tools • Podcasts (i.e., Listen Notes. Synth, Spotify) • Text-to-Speech Read Aloud tools that can be plugged into Browser • TeachVid: Using YouTube for Listening Assessment	Participate in Edmodo Discussion
9	Assessing speaking skills in computer-assisted environments: • Web 2.0 Tools: FlipGrid, Padlet, Google Classroom • Pronunciation tools (i.e., Accent Roise) • Using Web and Phone apps for informal speaking assessment (i.e., WhatsApp)	Participate in Edmodo Discussion
10	Peer and self-assessment: Theory and application • Online tools: i.e., Blogger, Twitter • Web 2.0 Tools: i.e., Padlet, Edmodo	Participate in Edmodo Discussion
11	Rater-mediated assessments and rater effects in language testing	Participate in Edmodo Discussion
12	Introduction to computer adaptive testing	Participate in Edmodo Discussion
13	Recap and presentation of the final project	

Course Evaluation

- Midterm and final projects: You will be expected to complete two projects. The midterm project will include designing a virtual classroom on a web tool (Padlet) and will compose 30% of your final grade. The final project will include creating a 12-week technology-enhanced course syllabus. This will compose 60% of your final grade. The instructions will be posted later.

- Weekly discussion forum: Every week, the instructor will initiate to post a comment such as an issue of the week or just a pep talk in order to support students both for their academic challenges and for affective challenges. Then students will be encouraged to respond to the posting, and in turn, they could also initiate posting a new issue. Participation performance in this activity will be counted towards 5% of your final grade.

- Peer assessment practice: Each student will randomly be assigned to evaluate the midterm and final project of another student. Participation performance in this activity will be counted towards 5% of your final grade.

Appendix B

Discussion Board Rubric*

Criteria	**Excellent**	**Good**	**Fair**	**Poor**
Timeliness and quantity of discussion responses	Postings well distributed throughout the week	Just one posting a week well before the due date	Just one posting close to the due date	Just one posting after the due date
Responsiveness to discussion topics and demonstration of knowledge and understanding from assigned readings.	Readings were understood and incorporated into discussion as it relates to the topic.	Readings were understood and incorporated into discussion as it relates to the topic.	Little use made of readings.	Little or no use made of readings. Postings have questionable relationships to discussion questions and/or readings; they are non-substantive
The ability of postings to move discussion forward.	Two or more responses add significantly to the discussions (e.g., identifying important relationships, offering a fresh perspective or critique of a point; offers supporting evidence).	At least one posting adds significantly to the discussion.	At least two postings supplement or add moderately to the discussion	Postings do little to move discussion forward
Points	**4**	**3**	**2**	**1**

*Adapted from a rubric from Purdue Repository for Online Teaching and Learning

Appendix C

Computer-Assisted Measurement Midterm Project (PLEASE RESPOND TO THIS LINK!) ...

Self-Assessment Activity

Please be sincere and accurate in your responses. The self-assessment form will be scored based on completion ONLY. So please do not hesitate to share your true experiences and self-evaluate yourself.

2. Please describe what did you learn through this project?

Enter your answer

3. Please describe how might this project contribute to your future teaching practice?

Enter your answer

4. .Please describe what did you struggle with the most while completing this project?

Enter your answer

Back

Submit

ISSN: 1309 – 6575
Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi
Journal of Measurement and Evaluation in Education and Psychology
Research Article; 2023; 14(Special Issue); 382-395

Learning Analytics in Formative Assessment: A Systematic Literature Review

Ke ZHANG * Ramazan YILMAZ ** Ahmet Berk USTUN ***
Fatma Gizem KARAOĞLAN YILMAZ****

Abstract

This systematic review examines the use of learning analytics (LA) in formative assessment (FA). LA is a powerful tool that can support FA by providing real-time feedback to students and teachers. The review analyzes studies published on Web of Science and Scopus databases between 2011 and 2022 that provide an overview of the current state of published research on the use of LA for FA in diverse learning environments and through different delivery modes. This review also explores the significant potential of LA in FA practices in digital learning. A total of 63 studies met all selection criteria and were fully reviewed by conducting multiple analyses including selected bibliometrics, a categorical meta-trends analysis and inductive content analysis. The results indicate that the number of LA in FA studies has experienced a significant surge over the past decade. The results also show the current state of research on LA in FA, through a range of disciplines, journals, research methods, learning environments and delivery modes. This review can help inform the implementation of LA in educational contexts to support effective FA practices. However, the review also highlights the need for further research.

Keywords: Learning analytics, formative assessment, assessment analytics, bibliometrics

Introduction

Formative Assessment

In the learning process, it is vital for the teacher to ascertain what the student already knows and teach accordingly (Ausubel, 1968). In this sense, assessment is an essential factor in the learning process. Students' performance and progress can be measured by assessment. Also, it shows what needs to be improved in the learning and teaching process. According to Lubinescu et al. (2001), assessment is a key factor for accreditation and evidence in the learning process. It occurs over the course of time by collecting evidence of learning in a systematic and planned way to determine whether a student achieved learning (Harlen et al., 2002). Two types of assessments encompassing assessment for formative and summative purposes have been emphasized in the literature. There is a distinction between these types of assessments. While summative assessment summarizes learning in order to make a decision related to recording, marking or certifying performance and achievements (Harlen & James, 1997), the formative assessment identifies aspects of learning by monitoring student learning during the learning process to provide feedback, modify learning and teaching activities and strengthen subsequent learning.

Formative assessment is a continuous process of evaluating student learning to identify areas of student weakness and make adjustments to instruction for improving student outcomes (Black & Wiliam, 1998). It involves ongoing monitoring and gathering evidence of students' progress during the learning process

* Professor, Wayne State University, Faculty of Education, Detroit-USA, ke.zhang@wayne.edu , ORCID ID: 0000-0002-4690-7586
** Assoc. Prof., Bartın University, Faculty of Science, Bartın-Turkiye, yilmaz04@gmail.com, ORCID ID: 0000-0002-2041-1750
*** Asst. Prof., Bartın University, Faculty of Science, Bartın- Turkiye, ustun.ab@gmail.com , ORCID ID: 0000-0002-1640-4291
**** Assoc. Prof., Bartın University, Faculty of Science, Bartın- Turkiye, gkaraoglanyilmaz@gmail.com, ORCID ID: 0000-0003-4963-8083

To cite this article:
Zhang, K., Yılmaz, R., Ustun, A. B., & Karaoğlan-Yılmaz, F. G. (2023). Learning analytics in formative assessment: a systematic literature review. *Journal of Measurement and Evaluation in Education and Psychology, 14*(Special Issue), 382-395. https://doi.org/10.21031/epod.1272054

Received: 28.03.2023
Accepted: 8.10.2023

(Yan et al., 2021). Based on collecting evidence of students' progress, feedback is provided to students as a key factor of formative assessment (Black & Wiliam, 1998; Stobart, 2008). Evidence based feedback is a useful strategy to foster students' learning outcomes in different circumstances. Furthermore, one way to extend formative assessment is to incorporate more technology into the process. For example, making online quizzes or assessments can provide immediate feedback to students and can help teachers identify areas of weakness more quickly (Karaoglan-Yilmaz et al., 2020; Ustun & Tracey; 2021). Additionally, using analytical tools in a learning management system (LMS) or any other smart system can allow teachers to track student progress over time and make data-driven decisions about instruction. This process can include the use of data from formative assessments, as well as data from other sources, such as data for demographic, student performance and student engagement (Karaoglan Yilmaz et al., 2022). By analyzing this data, educators can identify patterns and trends that can inform instruction and help to improve student outcomes.

Learning Analytics

The demand of extracting meaningful insights from high-volume data requires automated analytical analyses in order to strengthen and shape the learning environments and experience (Ustun et al., 2022). High-volume data should be turned into meaningful information about the learning and teaching processes through analytical analyses using statistical algorithms and mathematical techniques. Analytical analyses can be performed by Learning Analytics (LA) which provides information about students and the learning environment in order to "access, elicit, and analyse them for modelling, prediction, and optimization of learning processes" (Mah, 2016, p. 288). LA is an emerging field that potentially revolutionizes how we understand and improve learning. It can be defined as "the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs" (The Society for Learning Analytics (SoLAR, http://solaresearch.org/)). In other words, data that students generate can be collected, analyzed and reported to understand and optimize the teaching and learning process and the learning environment. Ultimately, LA uncovers students' learning patterns and behaviors to predict student learning outcomes (Xing et al. 2015) and also discovering their learning patterns and behaviors provides opportunities for teachers to tailor education by offering more personalized experiences or adaptive learning materials (Ndukwe & Daniel, 2020; Siemens, 2013).

One of the key benefits of LA is that it enables teachers to monitor student performance (Ustun et al., 2022). Teachers can gain a complete picture of how students are progressing and identify areas where they may need additional support by collecting data on students' activity, engagement, and achievement. Providing personalized instruction is another key benefit of LA. Teachers can tailor instruction to meet the specific needs of each student by analyzing data on how individual students learn (Schumacher & Ifenthaler, 2018). Therefore, students can more easily adapt to the content, pace, or style of the instruction, and this potentially leads to more effective and efficient learning. LA can also be used in the identification of at-risk students. By analyzing student engagement and performance data, teachers can identify students who may be at risk of falling behind and provide early interventions for these students according to their learning preferences and abilities to help them stay on track. (Gašević et al., 2016). Finally, LA can be utilized to enhance the design of learning environments and resources. The way students interact with learning environment and resources can be analyzed to identify areas where they can be improved to better support student learning. For instance, according to analyzing how students interact with a particular LMS, the interface of the LMS can be redesigned to make it more user-friendly or add features that students have found helpful (Ustun et al., 2021; Ustun & Tracey, 2020).

Assessment Analytics

Assessment analytics (AA) is a burgeoning research field and is considered a subset of LA. Economides (2009) states that 'like any other context-aware system, an AA procedure monitors, tracks and records data related to the context, interprets and maps the real current state of these data, organizes them (e.g., filter, classify, prioritize), uses them (e.g., decide adaptations, recommend, provide feedback, guide the learner) and predicts the future state of these data' (as cited in (Papamitsiou & Economides, 2016, p.118). In this sense, assessment analytics like LA is related to measuring, collecting, analyzing, and reporting data about students and environments in which learning occurs for the purposes of comprehension and optimization of the learning environments where data is extracted from assessment (Cooper, 2015). One of the major aims of assessment analytics is to support the assessment process in an effective and efficient manner (Papamitsiou & Economides, 2016) because assessment data has great potential for students to take advantage of them after meaningful results derive from analyses of assessment data (Ellis, 2013). The assessment analytics explicitly show what students need to invest their time to improve learning and lead teachers on what they need to modify and shape in the learning to improve learning processes. Assessment analytics can be used to predict student performance, improve the detection of students at risk and misconceptions, uncover gaps between what needs to be learned and what is already learned, and reveal students' behavior, cheating, and guessing.

Learning analytics and Formative assessment

LA and FA are closely related, as both involve the use of data to inform instruction and improve student learning. LA and formative assessment can provide an entire picture of student learning. Combining these two concepts informs pedagogical decisions and practices such as providing feedback to students (Taras, 2008). LA can be used to support formative assessment by providing data and insights that can inform instructional decisions and help teachers understand how their students are learning and make more informed decisions about instruction. LA offers opportunities for educational progress and gives formative guidance to students or teachers (Gašević et al., 2022). Specially, using analytical tools help teachers to provide LA based personalized feedback (Pardo et al., 2019). LA and formative assessment can be used to create a more data-driven and personalized approach to instruction, one that is continuously informed by student data and tailored to meet the needs of individual learners (Merikko, 2022). To gain a more comprehensive understanding of student learning, LA can be used in conjunction with formative assessment.

Purpose of the study

There are many studies on FA and LA in the literature. However, a gap exists in the literature in terms of reviews of research on applying LA in formative assessment. In order to fill this gap, the articles that were indexed in the Web of Science and Scopus databases and addressed the use of LA in the formative assessment were pinpointed and analyzed. The Web of Science and Scopus databases were chosen for the study because they provide access to the most relevant and prestigious publications in the related research area.

This review aimed to sketch the current landscape of published studies on LA for FA in a variety of learning environments through various delivery modes. The following questions guided our review and analyses.

1. Bibliometrics of the reviewed articles:
    - 1.1. What were the descriptive bibliometrics like?
    - 1.2. What journals were these studies published in?
    - 1.3 What disciplines or professional fields were these studies conducted in?
    - 1.4 What types of learning environments were these studies conducted in?
    - 1.5. What delivery modes were utilized in these studies?
2. Methodologies of the reviewed articles:
    - 2.1. What research methods were employed in these studies?
    - 2.2. What populations were studied with what types and sizes of participants?

Method

This review focused on the research on learning analytics in formative assessment.

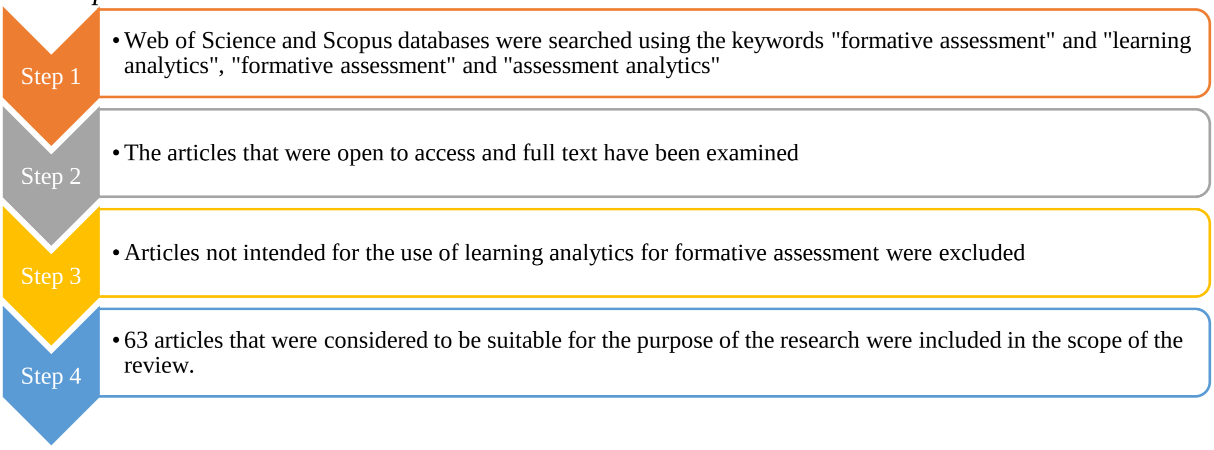
Search and Selection: Criteria and Processes

Multiple rounds of search were conducted. Web of Sciences and Scopus databases were identified and selected as the source databases to find related research publications on LA in FA, using the following keywords: “formative assessment” and “learning analytics”, “formative assessment” and “assessment analytics”.

The article search process in the databases was carried out by searching the keywords throughout the entire paper. In the Web of Science database, 90 articles were found using the keywords "formative assessment" and "learning analytics" and three additional articles were found using the keywords "formative assessment" and "assessment analytics". In the Scopus database, 796 articles were found using the keywords "formative assessment" and "learning analytics" and 24 articles were found using the keywords "formative assessment" and "assessment analytics". Duplicates in the multiple search results were excluded. Retrieved articles were further screened by the researchers, in terms of suitability for the purpose of the study. As a result, 63 articles were included in the systematic review. The search process is shown in Figure 1.

Figure 1

Search process



Considering the aims of this review, 63 articles were selected for further analysis. Multiple analyses were conducted, including selected bibliometrics (Okubo, 1997; Thelwall, 2008), a categorical meta-trends analysis (e.g., Hung & Zhang, 2012; Thelwall, 2008; Zhang & Aslan, 2021), and inductive content analysis (e.g., Gao et al., 2012; Mogil et al., 2009; Zhang & Aslan, 2021).

In order to find answers to the research questions, criteria were determined, and a form was created in the Microsoft Excel program according to these criteria, and the data obtained by examining 63 articles were processed into this form. Graphics and visuals have been prepared to make the data more understandable. Microsoft Excel and VOSviewer programs were used for these processes.

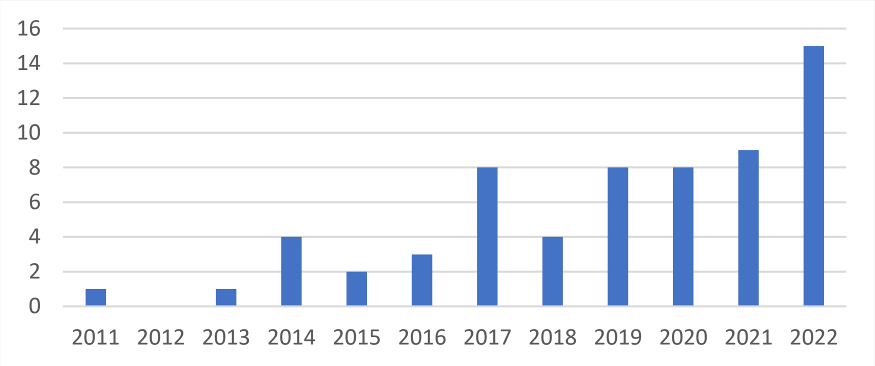
Findings

Descriptive Bibliometrics of the reviewed articles

LA in FA Research Article by Year

Figure 2

The Distribution of the articles by publication years

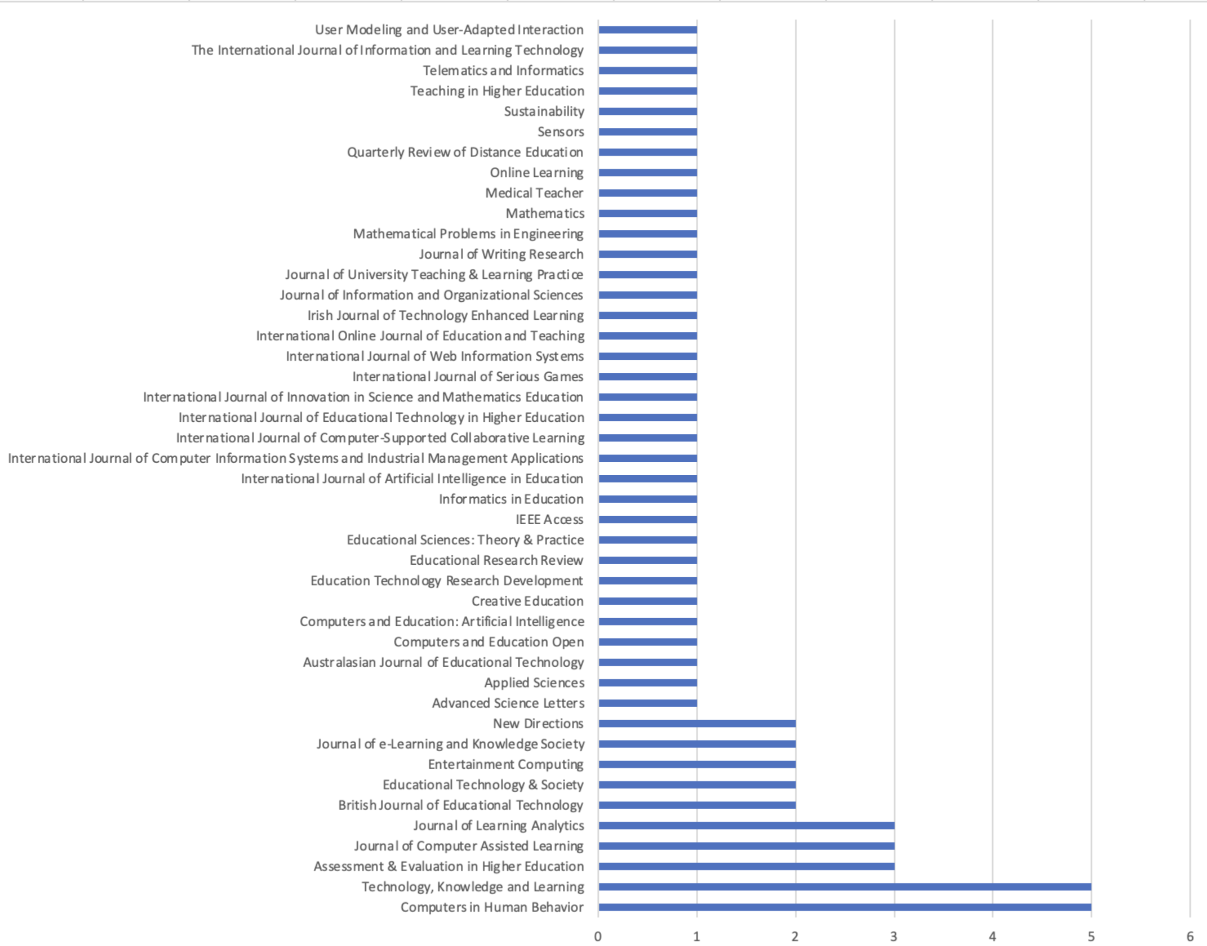


As illustrated in Figure 2, the first eligible research article in this review was published in 2011. Since then, in 12 years, the number of related research articles has increased from one in 2011 to 15 in 2022.

Journals publishing LA in FA research articles

Figure 3

The distribution of the articles according to journals

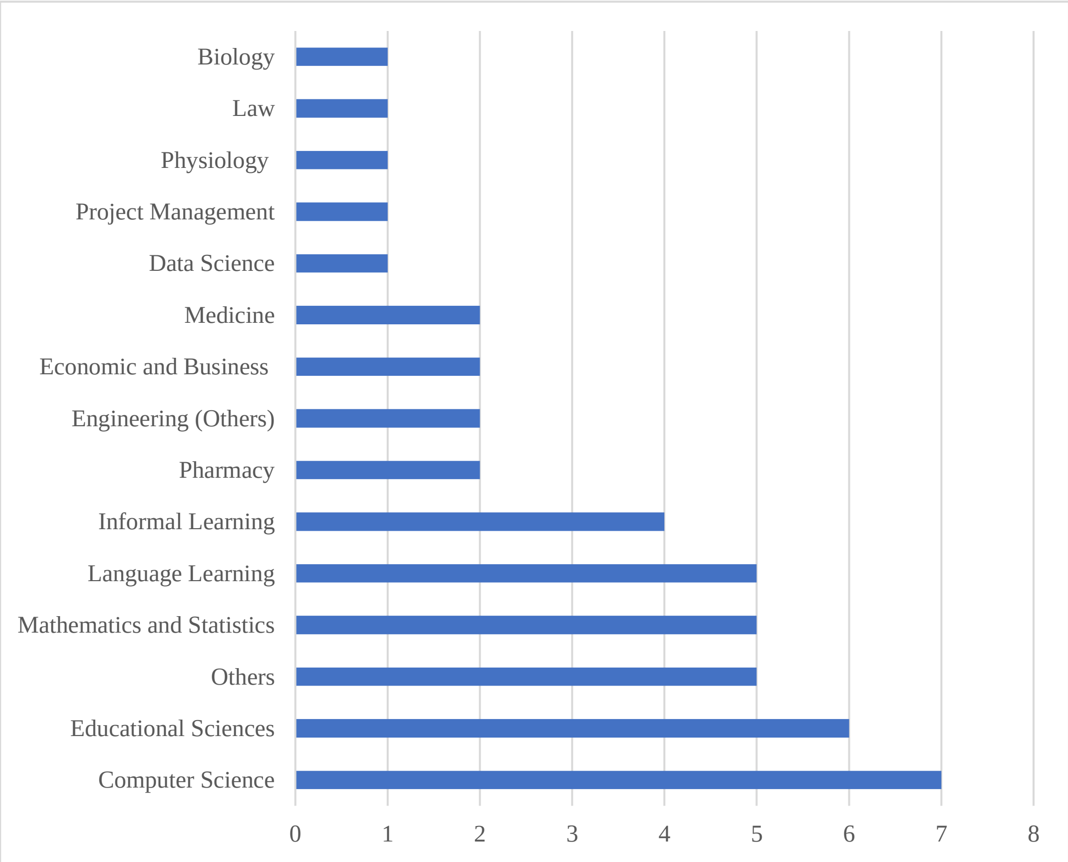


A total of 44 journals have published articles on LA in FA. Most of these journals have only published one or two of such studies so far, while the following journals have published a few more, *Computers in Human Behavior* (n=5), *Technology, Knowledge and Learning* (n=5), *Assessment & Evaluation in Higher Education* (n=3), *Journal of Computer Assisted Learning* (n=3), *Journal of Learning Analytics* (n=3).

LA in FA Research Article by discipline

Figure 4

The distribution of the articles according to the educational fields

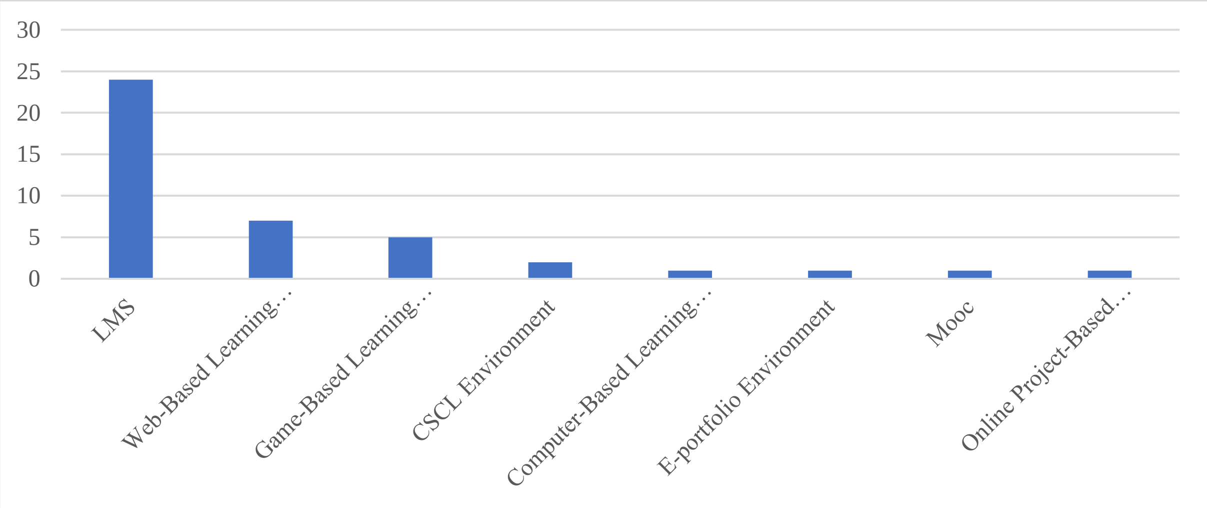


As seen in Figure 4, the articles on the use of learning analytics in formative assessment are mostly prepared on computer science (n=7), educational science (n=6), mathematics and statistics(n=5), and foreign language learning (n=5). Some studies have been conducted to include more than one discipline (e.g., Knight et al., 2020) or not to include any discipline (e.g., Barana et al., 2019). Therefore, the number of disciplines in which the research is conducted may differ in this respect.

LA in FA Research by technological learning environment

Figure 5

The distribution of the articles according to the learning environment

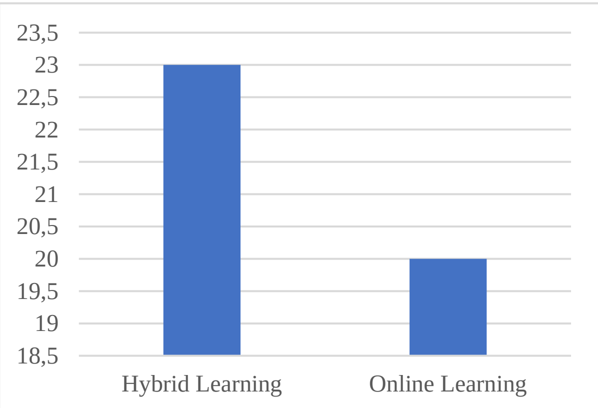


As seen in Figure 5, the studies on the use of learning analytics in formative assessment were mostly carried out using LMS. In addition to LMS, it is seen that web-based learning environments, game-based learning environments and CSCL environments were also used in the studies. The descriptions of the learning environments expressed in Figure 5 are as follows: Learning Management Systems (LMS) are utilized for educational purposes, exemplified by platforms like Moodle. Web-based learning environments encompass dynamic or static web pages designed for educational purposes. Computer Supported Collaborative Learning (CSCL) environments are utilized for computer-supported collaborative learning activities. Computer-based learning environments operate without an internet connection. E-portfolio environments allow students to create e-portfolios, upload content, and share them with their peers. Massive Open Online Course (MOOC) environments offer a wide range of courses to a large number of participants such as Khan Academy. Online project-based learning environments allow students to plan, collaborate and structure project products online.

LA in FA Research by delivery mode

Figure 6

The distribution of the articles according to delivery modes



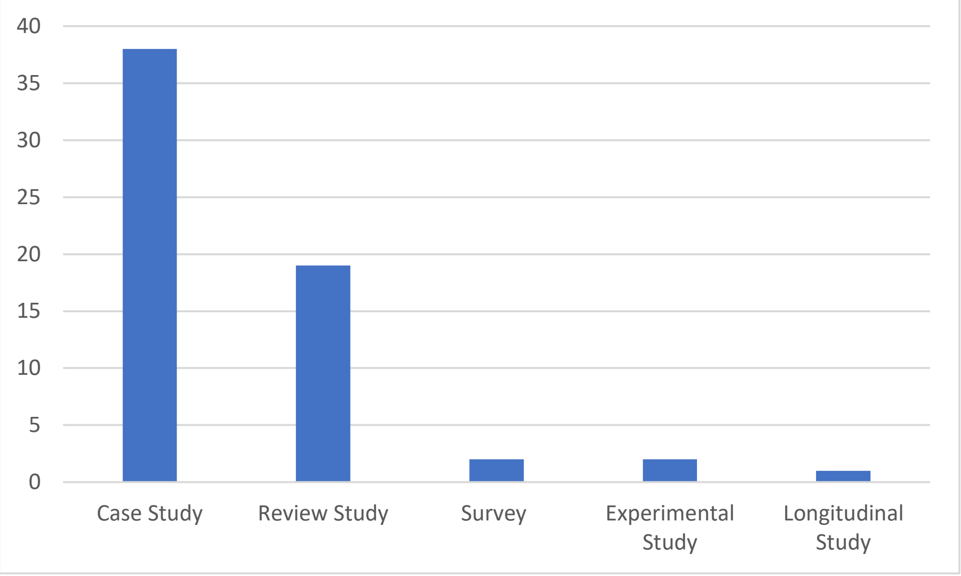
As seen in Figure 6, the studies on the use of learning analytics in formative assessment were mostly carried out in the modes of Hybrid Learning (n=23) and Online Learning (n=20).

Methodologies

Methodologies used in LA in FA research

Figure 7

The distribution of the articles according to the research method

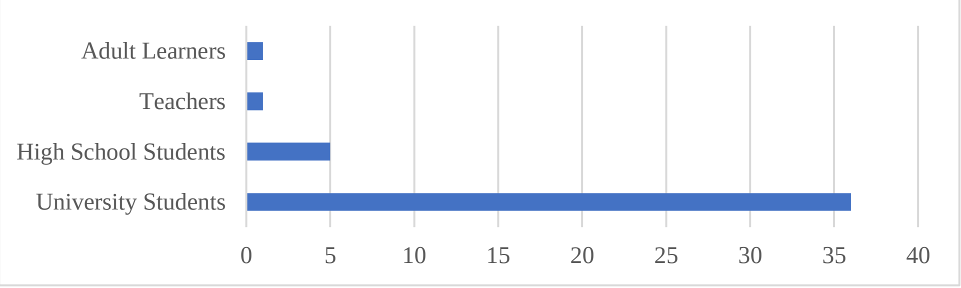


As seen in Figure 7, the case study method was mostly used in studies on the use of learning analytics in formative assessment, with 38 out of 63 studies. In addition, there were also studies conducted using the methods of Review Study (n=19), Survey (n=2), Experimental Study (n=2) and Longitudinal Study (n=1).

LA in FA research participants

Figure 8

The distribution of the articles according to the types of participants



As seen in Figure 8, studies on the use of learning analytics in formative assessment were mostly conducted with university students (n=36). In addition, there were also studies conducted on high school students (n=5), teachers (n=1), and adults (n=1), albeit a small number.

Discussion

Since the first LA in FA study was published in a Web of Science journal in 2011, the number of such publications has increased tremendously in the past decade. The 63 articles analyzed in this review represent the current state of research on LA in FA, through a range of disciplines, journals, research methods, learning environments, and delivery modes.

LA in FA has been applied in various fields, including computer science (e.g., Yan et al., 2021), law (e.g., Knight et al., 2020), education (e.g., Merikko et al., 2022), engineering (e.g., Gasevic et al., 2017), pharmacy (e.g., Liu et al., 2021) and many more. Thus, journals that have published such studies are diverse as well. A total of 44 journals have published research on LA in FA since 2011. The following journals including *Computers in Human Behavior*, *Technology, Knowledge and Learning*, *Assessment & Evaluation in Higher Education, Journal of Computer Assisted Learning* and the *Journal of Learning Analytics* have published a few more such studies than other Web of Science journals.

A wide range of applications of LA in FA was reported in these studies. For instance, LA is used to monitor the learning progress and learner engagement (e.g., Koc, 2017; O'Dowd, 2022; Nguyen et al., 2016), identify learners at risk (e.g., Choi et al., 2018), generate adaptive testing (e.g., Yilmaz et al., 2021), provide feedback for instructors and learners (e.g., Banihashem et al., 2022; Krull & Leijen, 2015), predict academic performance (e.g., Bulut et al., 2023; Martin & Ndoye, 2016), detect learning strategies (e.g., Gasevic et al., 2017), facilitate peer assessment (e.g., Er et al., 2021) and provide early warning for potential dropouts (e.g. Choi et al., 2018).

Most of the studies were conducted in either fully online or a hybrid delivery mode, which generate rich digital data ready for LA. More specifically, the research on LA in FA was implemented in learning management systems (LMS), web-based, game-based or CSCL learning environments. LA in FA was employed in higher education (e.g., O'Dowd, 2022) and high schools (e.g., Gomez et al., 2021). Thus, varied participants were recruited in related research, including teachers (e.g., Admiraal et al., 2020), high school students (e.g., Tempelaar et al., 2015), college students (e.g., Karaoglan Yilmaz, & Yilmaz, 2020), and adult learners (e.g., Serrano-Laguna et al., 2014) to explore the effects and user experiences of LA in FA.

A few different research methods are employed in these studies. Case study is the most often applied method in LA in FA studies, which allows deeply contextualized analysis of the practice. At the same time, the methodological limitation of such methods may also significantly limit the generalizability of the research findings. It is noteworthy though, that a small number of longitudinal (e.g., Martinez-Maldonado, 2019) and experimental (e.g., Tan et al., 2017) studies are also available.

LA techniques used in these studies include data mining, predictive modeling, and visualizations. The types of data analyzed in the studies include clickstream, log, and assessment data. The impact of LA in FA has been examined in terms of student learning outcomes, student engagement, and instructor feedback.

The review has found that LA in FA increases the capacity of digital learning by providing timely and actionable feedback to students and instructors. These studies investigate LA in FA for different purposes, such as generating feedback for students, providing feedback for instructors, creating student profiling, facilitating peer assessment, monitoring student performance, detecting learning strategies, offering automatic instant corrections, and more.

LA has become an essential area of education research. These reviewed studies provide further evidence for the educational benefits of LA in FA. LA provides instructors with data-driven insights into student learning (Karaoglan Yilmaz & Yilmaz, 2020). By leveraging LA, instructors can make informed choices about best supporting their students' learning progress. It can enhance student learning and engagement by providing personalized feedback and support while supporting instructor decision-making and promoting metacognitive development (Harindranathan & Folkestad, 2019). Besides, it can assist in identifying students who may be at risk of falling behind or encountering difficulties. By analyzing data related to behavior, participation, and student performance, instructors are able to intervene in advance and provide additional support to these students during the FA process. This proactive approach allows for timely interventions and can prevent academic setbacks.

The surge of AI technologies calls for creative ways to transform education and extend the educational landscape for more equitable and accessible education (Üstün, 2021; Zhang & Aslan, 2021). With the emergence of learning engineering as a new, interdisciplinary field (Zhang & Zhu, 2022), LA in FA becomes even more important as educators, educational technologies and educational researchers collaborate to transform digital learning. While LA focuses on the analysis of data to improve teaching and learning, learning engineering is concerned with the design, development and research of effective learning systems and technologies (Zhang & Zhu, 2022). LA in FA research can inform learning engineering by providing insights into student learning behaviors, preferences, and needs (Zhang & Zhu, 2022). LA makes it possible to provide immediate feedback to both students and instructors during the FA process. Through data analysis, instructors are able to identify areas where students may be struggling or excelling and provide relevant and constructive feedback to guide their learning (Ustun et al., 2022). Students can also receive personalized feedback during the FA process and LA-based feedback enables them to understand their strengths and weaknesses and make necessary improvements.

LA can help teachers tailor instruction to individual student needs. By analyzing learner-generated data, LA can identify patterns and trends (Hung & Zhang, 2008) that can be used to optimize learning design and delivery. By analyzing student data, instructors can identify knowledge gaps, learning styles, and preferences, allowing them to adapt their teaching strategies accordingly. For example, LA can be used to identify which instructional strategies are most effective for different types of learners, or which types of learning content are most engaging. This personalized approach enhances the effectiveness of FA by addressing specific student needs and promoting a more profound understanding. On the other hand, learning engineering can inform LA for FA by providing guidance on the design and development of effective LA tools (Zhang & Zhu, 2022). By designing tools that are tailored to the needs of learners and instructors, learning engineering can help to ensure that LA is actionable and scalable. For example, learning engineering can help to design LA tools that provide personalized feedback to learners (Ustun et al., 2022).

For LA to be used effectively, it is crucial to integrate it into learning environments. Cavus Ezin and Yilmaz (2022) indicate that LA must be integrated into the learning environments to benefit from the potential of LA in both online and hybrid learning. While different strategies and approaches can be followed, integration can be planned with the following steps by considering the educational goals of LA in general: a) Setting learning goals, b) Monitoring the learning process, c) Personalizing the learning experience, and d) Improving the learning experience.

a) Setting learning goals: LA can help instructors to set students' learning goals. Instructors can use student performance data to determine which areas students struggle or excel in. This information can assist them in setting goals and choosing appropriate activities to support students' learning. They should clarify the goals of lessons and what they want students to learn. These goals will help determine which data types and metrics to use for LA. For example, time-related results can be obtained from log data to increase students' attendance time in online courses. It is also essential to determine which data will be analyzed by LA. Various methods can be used to collect data such as exam grades, assignment performance, online interactions, and class attendance. LMS log data, surveys, quizzes, and other digital tools are some of the tools that can be used to collect data. LA can help instructors better understand students' learning and teach them more effectively. This tool can contribute to developing students' self-efficacy, and students who have developed self-efficacy increase their active participation in the lesson and feel their learning is more exciting and effective (Karaoglan-Yilmaz et al., 2023).

b) Monitoring the learning process: LA can be used to monitor students' learning progress. Instructors can use student performance data to track how students' learning progresses over time. This information can help instructors identify their needs and offer them sufficient support. Appropriate tools can be used to analyze the collected data. At this stage, the most frequently used data mining algorithms are decision trees, support vector machines, Naive Bayes, artificial neural networks, and regression methods (Tosunoğlu et al., 2021). At this stage, LA tools (dashboards, etc.) or data analysis software (R, Python, etc.) can be used to visualize data, identify trends, and understand student performance. Students' progress, strengths and weaknesses, interactions, and other important factors should be considered when analyzing data in this process.

c) Personalizing the learning experience: LA can be used to personalize students' learning experiences. Instructors can use student performance data to select activities that suit students' interests and needs. Hence, students' learning can be made more engaging and effective. Individualized feedback can be provided to students using the information obtained with LA. Suggestions can be made for students to improve their weaknesses while appreciating their strengths. This feedback can help them make their learning process more effective.

d) Improving the learning experience: LA can be used to improve the learning experience. Using student performance data, instructors can improve their teaching methods and materials. This can make students' learning more effective. Each item mentioned above regarding the integration of LA into the learning process constitutes a stage of the formative evaluation process. Therefore, through the formative assessment process, LA enables data collection, reporting on the acquired information, and facilitating interventions based on these reports. This way, LA can effectively support the advancement of the learning process. The learning process can be adjusted based on the information LA provides. For instance, if LA results show that students have difficulty understanding a particular subject, they may devote more time to it or offer them additional learning materials. Interactive activities or discussions can be organized to increase student interest and participation.

The above processes can be followed in integrating LA into the lessons. One of the essential points to be considered in this process is data privacy and ethical processes (Çetintav et al., 2022). Since LA results contain students' personal data, it is essential to pay attention to ethical processes in obtaining and using this data. LA results of a student should not be shared with other students in the class in a way that makes it identifiable to whom they belong.

Limitations of this review

This review is limited in its scope, as defined and specified in the method section. The selection of the source database and the specific search engines used in this review have also contributed to its methodological limitation. Research publications that do not include the selected keywords/terms as a descriptor in their title, abstract, or keyword list, as well as those not indexed in the source database are not included in this review.

Suggestions for future reviews

Future reviews may extend the search scope to include other reputable databases, specialized journals, or peer-reviewed conference proceedings. In addition, applying different search strategies, keywords, selection criteria, and exclusion criteria may retrieve more relevant research publications for a broader review.

Conclusion

Research has explored some of the powerful potentials of LA in renovating FA practices in digital learning. Dynamic LA empowers educators by providing critical insights into students' learning progress (e.g., Koc, 2017; O'Dowd, 2022; Nguyen et al., 2016), identifying struggling students (e.g., Choi et al., 2018; Saqr et al., 2017), and generating adaptive materials accordingly (e.g., Yilmaz et al., 2021). Thus, effective implementation of LA in FA could result in increased learner engagement, improved learning outcomes, boosted teaching efficiency, and better retention rates. The potential benefits of LA for FA make it a worthwhile investment for educational institutions, together with technology advancement.

Through a systematic review of empirical studies published in Web of Science and Scopus databases, this article portrays the trends of LA in FA research in the recent decade, since the first study was published in 2011. It has also explored the learning environments, delivery modes, disciplines, and participants of these studies, to develop a macro, as well as a micro-view of LA in FA research. The findings provide a preliminary foundation for more, historical, or meta-analyses of the increasing body of research literature on LA in FA.

To build a deeper understanding of the benefits as well as the challenges and issues of using LA in FA in digital education, more research is necessary. Different research methods are essential, and a larger number of participants are required for research on the scalable practice of LA in FA.

Declarations

Author Contribution: All authors equally contributed to conceptualization, methodology, analysis, writing, reviewing & editing, and visualization.

Conflict of Interest: No potential conflict of interest was reported by the authors.

Ethical Approval: Ethical rules were followed in this research. Ethical approval is not required, because data from WoS and Scopus was used in this research.

Funding: The authors declare that no funds, grants, or other support were received during the preparation of this manuscript

Consent to Publish: Written consent was sought from each author to publish the manuscript.

Competing Interests: The authors have no relevant financial or non-financial interests to disclose.

References

Admiraal, W., Vermeulen, J., & Bulterman-Bos, J. (2020). Teaching with learning analytics: how to connect computer-based assessment data with classroom instruction?. *Technology, Pedagogy and Education, 29*(5), 577-591.

Ausubel, D. P. (1968). Educational psychology: A cognitive view. New York, NY: Holt, Rinehart & Winston

Banihashem, S. K., Noroozi, O., van Ginkel, S., Macfadyen, L. P., & Biemans, H. J. (2022). A systematic review of the role of learning analytics in enhancing feedback practices in higher education. *Educational Research Review,* 100489.

Barana, A., Conte, A., Fissore, C., Marchisio, M., & Rabellino, S. (2019). Learning analytics to improve formative assessment strategies. *Journal of e-learning and knowledge society, 15*(3), 75-88.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. Assessment in Education: *Principles, Policy & Practice, 5*(1), 7–74.

Bulut, O., Gorgun, G., Yildirim-Erbasli, S. N., Wongvorachan, T., Daniels, L. M., Gao, Y., ... & Shin, J. (2023). Standing on the shoulders of giants: Online formative assessments as the foundation for predictive learning analytics models. *British Journal of Educational Technology, 54*(1), 19-39.

Cavus Ezin, C., & Yilmaz, R. (2022). The effect of learning analytics-based interventions in mobile learning on students' academic achievements, self-regulated learning skills, and motivations. *Universal Access in the Information Society,* 1-16.

Çetintav, G., Düzenli Çil, B. & Yılmaz, R. (2022). Eğitsel veri madenciliği ve öğrenme analitikleri araştirmalarinda veri gizliliği ve etik meseleler: araştirmalar üzerine bir inceleme. *Eğitim Teknolojisi Kuram ve Uygulama , 12*(1) , 113-146

Cooper, A. (2015). Assessment analytics. In eunis elearning task force workshop. https://dspacecris. eurocris.org/bitstream/11366/559/1/12%20Adam%20Cooper%20EUNIS%202015%20ELTF% 20Workshop.pdf. Accessed 14 September 2022.

Ellis, C. (2013). Broadening the scope and increasing the usefulness of learning analytics: The case for assessment analytics. *British Journal of Educational Technology, 44*(4), 662-664.

Gao, F., Luo, T., & Zhang, K. (2012). Tweeting for learning: A critical analysis of research on microblogging in education published in 2008–2011. *British Journal of Educational Technology, 43*(5), 783-801.

Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education, 28*(2016), 68–84.

Gašević, D., Greiff, S., & Shaffer, D. W. (2022). Towards Strengthening Links between Learning Analytics and Assessment: Challenges and Potentials of a Promising New Bond. *Computers in Human Behavior*, 107304.

Gasevic, D., Jovanovic, J., Pardo, A., & Dawson, S. (2017). Detecting learning strategies with analytics: Links with self-reported measures and academic performance. *Journal of Learning Analytics, 4*(2), 113-128.

Gomez, M. J., Ruipérez-Valiente, J. A., Martínez, P. A., & Kim, Y. J. (2021). Applying learning analytics to detect sequences of actions and common errors in a geometry game. *Sensors*, *21*(4), 1025.

Harindranathan, P., & Folkestad, J. (2019). Learning analytics to inform the learning design: Supporting instructors' inquiry into student learning in unsupervised technology-enhanced platforms. *Online Learning*, *23(*3), 34-55.

Harlen, W. & Deakin Crick, R. (2002). *A systematic review of the impact of summative assessment and tests on students' motivation for learning* (EPPI-Centre Review, version 1.1). In Research Evidence in Education Library London, EPPI-Centre, Social Science Research Unit, Institute of Education.

Harlen, W., & James, M. (1997). Assessment and learning: differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy & Practice, 4*(3), 365-379.

Hung, J. L., & Zhang, K. (2008). Revealing online learning behaviors and activity patterns and making predictions with data mining techniques in online teaching. *MERLOT Journal of Online Learning and Teaching, 4*(4), 426-437.

Hung, J. L., & Zhang, K. (2012). Examining mobile learning trends 2003–2008: A categorical meta-trend analysis using text mining techniques. *Journal of Computing in Higher education*, *24*, 1-17.

Karaoglan Yilmaz, F. G., & Yilmaz, R. (2020). Student opinions about personalized recommendation and feedback based on learning analytics. *Technology, knowledge and learning, 25*(4), 753-768.

Karaoglan-Yilmaz, F. G., Ustun, A. B., & Yilmaz, R. (2020). Investigation of pre-service teachers' opinions on advantages and disadvantages of online formative assessment: an example of online multiple-choice exam. *Journal of Teacher Education and Lifelong Learning, 2*(1), 1-8.

Karaoglan-Yilmaz, F. G., Ustun, A. B., Zhang, K., & Yilmaz, R. (2023). Metacognitive awareness, reflective thinking, problem solving, and community of inquiry as predictors of academic self-efficacy in blended learning: a correlational study. *Turkish Online Journal of Distance Education, 24*(1), 20-36.

Karaoglan-Yilmaz, F. G., Zhang, K., Ustun, A. B., & Yilmaz, R. (2022). Transactional distance perceptions, student engagement, and course satisfaction in flipped learning: a correlational study. *Interactive Learning Environments*, 1-16.

Knight, S., Shibani, A., Abel, S., Gibson, A., Ryan, P., Sutton, N., Wight, R., Lucas, C., Sándor, Á., Kitto, K., Liu, M., Vijay Mogarkar, R., & Buckingham Shum, S. (2020). Acawriter: A learning analytics tool for formative feedback on academic writing. *Journal of Writing Research*, *12*(1), 141-186.

Krull, E., & Leijen, Ä. (2015). Perspectives for defining student teacher performance-based teaching skills indicators to provide formative feedback through learning analytics. *Creative Education*, *6*(10), 914.

Liu, M., Kitto, K., & Shum, S. B. (2021). Combining factor analysis with writing analytics for the formative assessment of written reflection. *Computers in Human Behavior, 120,* 106733.

Lubinescu, E. S., Ratcliff, J. L., & Gaffney, M. A. (2001). Two continuums collide: Accreditation and assessment. *New directions for higher education, 2001*(113), 5-21.

Mah, D. K. (2016). Learning analytics and digital badges: Potential impact on student retention in higher education. *Technology, Knowledge and Learning, 21,* 285-305.

Martin, F., & Ndoye, A. (2016). Using learning analytics to assess student learning in online courses. *Journal of University Teaching & Learning Practice*, *13*(3), 7.

Martinez-Maldonado, R. (2019). A handheld classroom dashboard: Teachers' perspectives on the use of real-time collaborative learning analytics. *International Journal of Computer-Supported Collaborative Learning, 14*(3), 383-411.

Merikko, J., Ng, K., Saqr, M., & Ihantola, P. (2022). To Opt in or to Opt Out? Predicting Student Preference for Learning Analytics-Based Formative Feedback. *IEEE Access, 10,* 99195-99204.

Mogil, J. S., Simmonds, K., & Simmonds, M. J. (2009). Pain research from 1975 to 2007: a categorical and bibliometric meta-trend analysis of every Research Paper published in the journal, Pain. *Pain*, *142*(1-2), 48-58.

Ndukwe, I. G., & Daniel, B. K. (2020). Teaching analytics, value and tools for teacher data literacy: A systematic and tripartite approach. *International Journal of Educational Technology in Higher Education, 17*(1), 1–31.

O'Dowd, I. (2022). Using learning analytics to improve online formative quiz engagement.Irish *Journal of Technology Enhanced Learning, 3*(1),28-41.

Okubo, Y. (1997). *Bibliometric Indicators and Analysis of Research Systems: Methods and Examples STI Working Papers 1997/1*. OECD.

Papamitsiou, Z., & Economides, A. A. (2016). An Assessment Analytics Framework (AAF) for enhancing students' progress. In Formative assessment, learning data analytics and gamification (pp. 117-133). Academic Press.

Pardo, A., Jovanovic, J., Dawson, S., Gašević, D., & Mirriahi, N. (2019). Using learning analytics to scale the provision of personalised feedback. *British Journal of Educational Technology, 50*(1), 128-138.

Saqr, M., Fors, U., & Tedre, M. (2017). How learning analytics can early predict under-achieving students in a blended medical education course. *Medical teacher, 39*(7), 757-767.

Schumacher, C., & Ifenthaler, D. (2018). Features students really expect from learning analytics. *Computers in Human Behavior, 78,* 397–407.

Serrano-Laguna, Á., Torrente, J., Moreno-Ger, P., & Fernández-Manjón, B. (2014). Application of learning analytics in educational videogames. *Entertainment Computing, 5*(4), 313-322.

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist, 57*(10), 1380–1400.

Stobart, G. (2008). *Testing times: The uses and abuses of assessment*. Routledge.

Tan, J. P. L., Koh, E., Jonathan, C., & Yang, S. (2017). Learner dashboards a double- edged sword? Students' sense-making of a collaborative critical reading and learning analytics environment for fostering 21st-century literacies. *Journal of Learning Analytics, 4*(1), 117-140.

Taras, M. (2008). Assessment for learning: Sectarian divisions of terminology and concepts. *Journal of Further and Higher Education, 32*(4), 389–397.

Tempelaar, D. T., Rienties, B., & Giesbers, B. (2015). In search for the most informative data for feedback generation: Learning analytics in a data-rich context. *Computers in Human Behavior, 47,* 157-167.

Thelwall, M. (2008). Bibliometrics to webometrics. *Journal of information science*, *34*(4), 605-621.

Tosunoğlu, E., Yilmaz, R., Özeren, E., & Sağlam, Z. (2021). Eğitimde makine öğrenmesi: Araştırmalardaki güncel eğilimler üzerine inceleme. *Ahmet Keleşoğlu Eğitim Fakültesi Dergisi, 3*(2), 178-199.

Üstün, A. B. (2021). The power of using emerging technologies in MOOCs: Accelerating globalization in higher education. *Journal of Learning and Teaching in Digital Age, 6*(2), 141-148.

Ustun, A. B., & Tracey, M. W. (2020). An effective way of designing blended learning: A three phase design-based research approach. *Education and Information Technologies, 25*(3), 1529-1552.

Ustun, A. B., & Tracey, M. W. (2021). An innovative way of designing blended learning through design-based research in higher education. *Turkish Online Journal of Distance Education, 22*(2), 126-146.

Ustun, A. B., Karaoglan Yilmaz, F. G., & Yilmaz, R. (2021). Investigating the role of accepting learning management system on students' engagement and sense of community in blended learning. *Education and Information Technologies, 26*(4), 4751-4769.

Ustun, A. B., Zhang, K., Karaoğlan-Yilmaz, F. G., & Yilmaz, R. (2022). Learning analytics based feedback and recommendations in flipped classrooms: an experimental study in higher education. *Journal of Research on Technology in Education,* 1-17.

Xing, W., Guo, R., Petakovic, E., & Goggins, S. (2015). Participation-based student final performance prediction model through interpretable Genetic Programming: Integrating learning analytics, educational data mining and theory. *Computers in Human Behavior, 47,* 168–181.

Yan, H., Lin, F., & Kinshuk (2021). Including learning analytics in the loop of self-paced online course learning design. *International Journal of Artificial Intelligence in Education, 31*(4), 878-895.

Yan, Z., Li, Z., Panadero, E., Yang, M., Yang, L., & Lao, H. (2021). A systematic review on factors influencing teachers' intentions and implementations regarding formative assessment. *Assessment in Education: Principles, Policy & Practice, 28*(3), 228-260.

Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. *Computers and Education: Artificial Intelligence*, *2*, 100025.

Zhang, K., & Zhu, M. (2022). Define "Learning Engineering" with the TRAP Framework. *Engineering Technology Open Access Journal, 4*(2): 555635. DOI: 10.19080/ETOAJ.2022.03.555635