



The effectiveness of web-based Mathematics instruction (WBMI) on K-16 students' mathematics learning: a meta-analytic research

Ayça Akin¹

Received: 22 September 2021 / Accepted: 31 January 2022 / Published online: 28 February 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Given the increasing prevalence of web technology, web-based mathematics environments have been increasingly widely used in mathematics education for the past two decades. The COVID-19 pandemic has led to an urgent transition from traditional mathematics instruction (TMI) to web-based mathematics instruction (WBMI) at all levels of mathematics education. At this point, it is crucial to scrutinize the effects of WBMI on K-16 students' mathematics learning comprehensively. This meta-analysis research contained a total of 63 studies with 115 effect sizes, which aimed to investigate the effectiveness of WBMI on K-16 students' mathematics learning by incorporating potential moderators, namely mathematics topics, mathematical content standards, feedback status, type of instructional features, age (i.e., grade level), and assessment methods. Based on findings, WBMI has a significantly strong effect on K-16 students' mathematics learning ($g = 1.10$, $p = 0.01$, 95% CI [0.95, 1.27]). Moderator analyses reveal that the effect sizes of WBMI on K-16 students' mathematics learning varied significantly depending on all these potential moderators. Additionally, higher-level mathematical concepts, statistics and probability, WBMI with providing feedback, tutorial systems, undergraduate students, and traditional paper-pencil assessment are the strongest moderators in their context. The most notable results of this research are that WBMI is significantly more effective on students' mathematics learning than TMI, while even in the context of WBMI, traditional paper-pencil assessment is significantly more effective than online assessment. This meta-analytic research provides a comprehensive and up-to-date perspective on the effectiveness of WBMI on K-16 students' mathematics learning.

Keywords Web-based learning environments · Mathematics learning · K-16 students · meta-analysis · Moderator analysis · Web-based mathematics instruction

1 Introduction

Web-based mathematics instruction (WBMI) plays a crucial role in mathematics education as technological advancements have changed the way of learning and teaching mathematics (Cao et al., 2021). Although WBMI has been initially seen only as a preferred teaching option in graduate and undergraduate education, it has been widely used at all levels of K-16 education in recent years. The most important reason why WBMI has gained that popularity is the COVID-19 pandemic. This pandemic has led to an urgent transition from traditional mathematics instruction to WBMI at all levels of mathematics education. Thanks to today's advanced internet and technology, WBMI has helped to overcome educational and instructional challenges in the pandemic, as it provides the flexibility to learn and teach mathematics at anytime and anywhere (Misirli & Ergulec, 2021).

A well-planned WBMI is not a simple process; however, is considered a complex process that requires elaborated teaching design, practice, and evaluation to create a productive learning environment (Misirli & Ergulec, 2021; Palloff & Pratt, 2013). Many studies emphasize that WBMI provides many learning opportunities such as communicating with large individual groups involving students, teachers, parents, and school principals (Martindale et al., 2005), implementing various instructional features (i.e., drill-and-practice programs, simulations, tutorial, and ITS), applying different educational features (i.e., pacing, feedback, and guided task), and empowering individuals to access many digital resources associated with mathematics topics (Hillmayr et al., 2020; Lin, 2009; MacGregor & Lou, 2004). By providing students with a personalized learning resource, WBMI allows students to dynamically adjust the learning process based on their skills and previous knowledge, considering the level and type of teaching (Hillmayr et al., 2020). Moreover, WBMI can increase students' mathematics performance due to providing feedback to students in this environment (e.g., Guzeller & Akin, 2012; Nguyen & Kulm, 2005). Since WBMI gives individuals the opportunity to access the learning resource at any time, individuals can advance the learning process at any desired pace, cease their learning at any time, and repeat parts of the learning section (Aberson et al., 2000).

In the literature, it is emphasized that WBMI offers many learning opportunities to students, as well as many difficulties are encountered in this learning environment (Gu & Lee, 2019). Recent studies indicate that individuals usually give preference to face-to-face instruction over WBMI, and WBMI has a fairly high drop-out rate compared to other web-based courses due to the difficult and complex structure of mathematics (e.g., Jaggars, 2014; Smith & Ferguson, 2005). WBMI is problematic in terms of mathematics learning as students experience a sense of isolation and lack of social support due to the nature of this learning environment (Jaggars, 2014; Gu & Lee, 2019). Moreover, interesting sites on the web such as social media, games, advertisements are another factor that negatively affects mathematics learning in WBMI (Tsai & Shen, 2009). Although the COVID-19 pandemic has led to an automatic transition to WBMI, it is not possible to infer that students and mathematics teachers have adapted to WBMI (Baran, Correia, & Thompson, 2011). Face-to-face mathematics instruction has been the first preferred type of instruction by students and mathematics teachers for centuries, making it difficult for mathematics teachers

and students to switch to WBMI (Cao et al., 2021). This leads students and mathematics teachers to have the belief that face-to-face mathematics instruction is more effective than WBMI.

More research papers have been published in the context of WBMI in the past few years due to the COVID-19 pandemic (e.g., Misirli & Ergulec, 2021; Sun et al., 2021; Ulum, 2022). It has been revealed that there were many disparities among students due to socio-economic reasons, and many students had difficulty in learning mathematics within the context of WBMI during this pandemic (Agasisti et al., 2020; Dorn et al., 2020). This has made the need for strengthening mathematics education to be more acute (Sun et al., 2021). As WBMI has an increasingly key role in the post-pandemic world, it is necessary to provide clear findings on the effectiveness of WBMI to all stakeholders such as mathematics educators, mathematics teachers, students, and parents (Sun et al., 2021).

Numerous studies are investigating the effect of WBMI on mathematics achievement (e.g., Aberson et al., 2000; Baki & Guveli, 2008; Gu & Lee, 2019; Guzeller & Akin, 2012; Nguyen & Kulm, 2005). The results of many of these studies indicate that WBMI has been significantly more effective in mathematics learning compared to traditional mathematics instruction (TMI) (e.g., Gu & Lee, 2019; Guzeller & Akin, 2012; Lin, 2009; Nguyen & Kulm, 2005). On the contrary, the results of several studies reveal that no significant difference was obtained between mathematics scores of students in WBMI and TMI (e.g., Baki & Guveli, 2008; Martindale et al., 2005; Mman & Tudunkaya, 2019). Since there are ambiguous findings on the effect of WBMI on mathematics learning in the literature, previous research has not provided much insight into the effect of WBMI on mathematics learning.

This inconsistency is evident when we scrutinize research papers that examine the effectiveness of web-based instruction in experimental and quasi-experimental research. Unfortunately, even in the context of meta-analysis, the effectiveness of web-based instruction on academic or mathematics learning remains unclear, as many research papers vary in their results (Fang et al., 2019; Hillmayr et al., 2020; Sitzmann, Kraiger, Stewart, & Wisher, 2006; Sun et al., 2021; Ulum, 2022). One of the largest sources of variation is likely to be highly heterogeneous meta-analysis research conducted with the inclusion of all studies related to web-based instruction that focuses not only on mathematics but also on biology, English, ICT, science, and social science (e.g., Hillmayr et al., 2020; Ulum, 2022). Probably, another reason for this variability is that only research papers related to Assessment and Learning in Knowledge Spaces (ALEKS) in the context of online intelligent tutoring systems (ITS) are included in previous meta-analysis studies (Fang et al., 2019; Sun et al., 2021).

In order to obtain comprehensive and consistent findings related to the effectiveness of WBMI on mathematics learning, it is necessary to include research papers associated with the implementation of various instructional features such as drill-and-practice programs other than ALEKS in the future meta-analysis research. However, as far as we know, the effect of web-based instruction on mathematics topics, mathematical content standards, feedback status, and assessment methods have not been examined independently and investigated as potential moderators in any previous meta-analysis studies (Fang et al., 2019; Hillmayr et al., 2020; Sitzmann et al.,

2006; Sun et al., 2021; Ulum, 2022). Additionally, due to the impact of the COVID-19 pandemic, new research papers have been published in the context of WBMI. Therefore, even the most recent meta-analysis studies are limited in terms of research intensity (Hillmayr et al., 2020; Sun et al., 2021; Ulum, 2022). It is also found that unpublished dissertations have not been included in the most recent meta-analysis in the context of WBMI (e.g., Hillmayr et al., 2020; Ulum, 2022). Consequently, previous meta-analysis studies on web-based instruction have not provided comprehensive and detailed findings about mathematics learning and assessment. Given all the above, it is crucial to scrutinize the effects of WBMI on K-16 students' mathematics learning comprehensively by conducting a new meta-analysis study. In conclusion, a study that the effectiveness of WBMI on mathematics learning while incorporating potential moderators may provide comprehensive, updated, and valuable findings on this topic.

2 Potential moderators of the effectiveness of WBMI on mathematics learning

Based on the literature review, potential moderators of the effectiveness of WBMI on mathematics learning can be listed as mathematics topics, mathematical content standards, feedback status, type of instructional features, age (i.e., grade level), and assessment methods.

2.1 Mathematics topics

School mathematics topics are a set of different components consisting of foundational concepts, higher-level mathematical concepts, and mathematical skills (Jitendra et al., 2018; Lin, 2011). Foundational concepts are associated with the basic content of mathematics such as numbers and operations, fractions, and decimals whereas higher-level mathematical concepts are associated with several complex mathematics topics such as trigonometry, functions, and equations (Jitendra et al., 2018). Moreover, mathematical skills are related to several skills such as spatial skills, mental rotation, visualization, and logical reasoning (Lin, 2011). Many studies indicate that WBMI helps students not only acquire foundational concepts but also assist them to comprehend higher-level mathematical concepts and master mathematical skills (e.g., Gu & Lee, 2019; Guzeller & Akin, 2012; Lin, 2009; Rafi et al., 2005). However, researchers (e.g., Moos & Azevedo, 2006; Smith & Ferguson, 2005) emphasize that when learning complex and difficult mathematics topics in WBMI, most students lose interest and barely adapt to this environment due to their inadequate previous knowledge in mathematics. It is also stated that poorly adapted web-based mathematics learning environments are problematic in learning high-level mathematical concepts (Gu & Lee, 2019; Smith & Ferguson, 2005). Therefore, the effects of WBMI on mathematics learning may differ depending on the components of school mathematics topics. Moreover, it remains unclear whether the effects of WBMI on mathematics learning vary depending on the components of school mathematics topics.

2.2 Mathematical content standards

In the literature, the mathematics performance of students is measured by mathematics tasks that span the domain of numbers & operations, algebra, geometry, statistics & probability, or mixed (i.e., all components of the content standards) in the context of WBMI. It is seen that the initial studies related to WBMI focused on students' mathematics learning in the domain of statistics and probability in the context of distance and remote learning technologies (Couch, 1997; Hurlburt, 2001). Additionally, many highly adaptive web-based learning environments have been created in the domain of statistics and probability (e.g., Hurlburt, 2001; Ozyurt et al., 2014; Muhanna & Abu-Al-Sha'r, 2010). It can be asserted that the effect of WBMI on students' mathematics learning is greater in the domain of statistics and probability compared to other content standards. Since the effects of WBMI on mathematics learning may differ depending on the components of content standards, more studies are needed to investigate this assertion.

2.3 Feedback status

One of the most crucial assets is feedback in the context of WBMI. Feedback not only allows students to check if their solution is correct but also provides guidance to students, helps students examine their own mistakes, and supports students to develop productive ways of thinking (Nguyen & Kulm, 2005). Many studies (e.g., Gu & Lee, 2019; Guzeller & Akin, 2012; Lin, 2009; Nguyen & Kulm, 2005) have been conducted to provide feedback to students in the web-based mathematics learning environment, whereas there are several studies (e.g., Babbitt et al., 2015; Kurtulus & Kilic, 2009; Taylor, 2008) that don't provide feedback to students in this environment. Therefore, examining the potential role of feedback in the effect of WBMI on learning mathematics may reveal crucial findings in this issue.

2.4 Instructional features

Research in the literature has examined the effectiveness of WBMI on mathematics learning by including the implementation of various instructional features namely drill-and-practice programs, simulations, tutorial systems, and intelligent tutoring systems (ITS) (Hillmayr et al., 2020; Lin, 2009). Accordingly, the type of instructional features may moderate the effects of WBMI on mathematics learning. Since the effects of WBMI on mathematics learning may differ depending on the types of instructional features, more studies are needed to handle this assertion.

2.5 Grade level

Before the COVID-19 pandemic, WBMI was widely used in distance education at the undergraduate level (Cao et al., 2021). There have only been a few research that examined the nature of the web-based mathematics learning environment (Cady & Rearden, 2009; Cao et al., 2021). The effectiveness of WBMI is not easy to interpret, as WBMI has not been widely used at the elementary, middle, and high school levels

compared to the undergraduate level until the post-pandemic world. It is emphasized that the web-based mathematics learning environments impede younger students' mathematics learning due to the overuse of unaided, uncontrolled, and undirected mathematical activities (Berger et al., 1994; Muhanna & Abu-Al-Sha'r, 2010). Since it is more difficult for younger children to maintain self-control compared to older children in the internet environment, an assisted and guided web-based mathematics learning environment needs to be designed, especially for elementary and middle school students. Due to the influence of grade level on web-based mathematics environments, more research exploring the effect sizes of WBMI across different age categories may be beneficial.

2.6 Assessment methods

It is seen that two different types of assessment methods are used in the context of WBMI. One of them is a traditional paper-pencil assessment and the other is an online assessment. In the traditional paper-pencil assessment under the scope of WBMI, the experimental and control groups take a paper-and-pencil test associated with a mathematical concept or topic. On the other hand, the control group students take a pre and post-test in paper-pencil form while the experimental group students take a pre and post-test on the web in online assessment under the scope of WBMI.

Online assessment has been increasingly widely used in STEM since the 2000s (e.g., Brouwer et al., 2009; Engelbrecht & Harding, 2004; Escudier et al., 2011; Jones & Long, 2013). In the internet age, it is asserted that traditional assessment is not the best way to evaluate students' learning outputs (Rane & MacKenzie, 2020). Moreover, many researchers argue that online assessment is more useful than traditional paper-pencil assessments since individuals may take their exams whenever and wherever they want using online assessments which give them more flexibility (e.g., Engelbrecht & Harding, 2004; Rane & MacKenzie, 2020). The most important reasons why students prefer online assessment can be listed as the absence of exam anxiety, providing immediate feedback on exam results, eligibility for formative assessment, the flexibility of the online setting, and access to the latest technology (Engelbrecht & Harding, 2004). In contrast, the biggest challenge with online assessment is that cheating, and plagiarism are high in online exams (Rane & MacKenzie, 2020). Therefore, Kennedy et al. (2000) have emphasized that as online assessment becomes more prevalent, academic dishonesty would rise. Although online assessment has been popular in recent years, it seems that most of the students prefer traditional paper-pencil assessment (e.g., Dandurand, Shultz, Onishi, 2008; Escudier et al., 2011; Rane & MacKenzie, 2020). The most crucial reasons why students prefer traditional paper-pencil assessment can be listed as marking online exams is a very time-consuming task, and the difficulty of adjusting to an unfamiliar way of testing (Engelbrecht & Harding, 2004).

Many studies compare students' academic achievements who were evaluated using traditional paper-pencil assessment with students' academic achievements evaluated with online assessment (e.g., Escudier et al., 2011; Jones & Long, 2013; Pennebaker, Gosling, Ferrell, 2013; Rane & MacKenzie, 2020; Stephens, 2001). Based on the literature, traditional paper-pencil assessment methods have shown to be more effective

than online assessment (e.g., Dandurand et al., 2008; Escudier et al., 2011; Jones & Long, 2013; Still & Still, 2015). In contrast, the results of several studies indicate that online assessment is more effective than traditional paper-pencil assessment methods in the context of mathematics (Ricketts & Wilks, 2002; Rane & MacKenzie, 2020; Pennebaker et al., 2013). However, it has numerous claims for the benefits of online assessment and determined many advantages, there is no evidence to indicate that if this approach would be as effective and strong as traditional paper-pencil assessment regarding students' achievement (Escudier et al., 2011). Therefore, more comprehensive research is needed to confirm or reject this claim.

3 The research question

The research problem of whether and to what extent WBMI is effective in K-16 students' mathematics learning has not been addressed in the literature. It has become clear that there is an increasing need for meta-analytic research that provides a comprehensive and up-to-date perspective on the effectiveness of WBMI on K-16 students' mathematics learning since educational stakeholders need to know to which extent the effects of WBMI on mathematics learning in the post-pandemic world (Sun et al., 2021). Therefore, this meta-analytic research scrutinizes the effectiveness of WBMI on K-16 students' mathematics learning by incorporating potential moderators, namely mathematics topics, mathematical content standards, feedback status, type of instructional features, age (i.e., grade level), and assessment methods. For this reason, the research questions of this research are:

- i) What is the overall effect of WBMI on K-16 students' mathematics learning?
- ii) Do the effect sizes of WBMI on K-16 students' mathematics learning vary depending on potential moderators?

4 Method

The present research is an appropriate context for the use of a meta-analysis since it makes it possible for a quantitative assessment of previous studies and provides more rigorous results due to a greater statistical power (Román-Caballero, Vadillo, Trainor, & Lupiáñez, 2021). Meta-analysis is an analysis of the analyses, allowing researchers to make statistically more accurate estimates of their research problem (Borenstein et al., 2009). Moreover, the meta-analysis design provides researchers with the opportunity to investigate the research problem in a comprehensive, in-depth, and systematic way (Guzeller & Celiker, 2019). The most important features of meta-analytical studies are that they provide numerical estimators of summary effect and between-research consistency, which makes it possible to evaluate the relevance of interventions/instruction (not just their statistical significance) and identify potential moderator variables (Román-Caballero et al., 2021).

Hence, the method of this study has been considered in the context of the meta-analytic design owing to the in-depth study of the effectiveness of WBMI on K-16

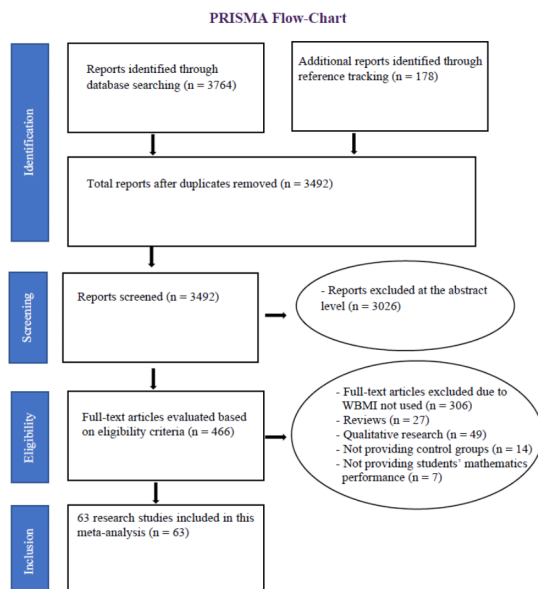
students' mathematics learning. Additionally, the present research is conducted in accordance with the meta-analysis design based on the research problem.

4.1 Literature search and research identification

The literature search was carried out in six databases that are most commonly used by researchers including Web of Science, Scopus, Google Scholar, ERIC, EBSCOhost online, and ProQuest Dissertations & Theses. The combination of keywords “web-based mathematics instruction”, “WBMI”, “WMT”, “web-based learning”, “online learning”, “online mathematics”, web-based course”, web-based” and “online” were used in this literature search that confined to the research published between January 2000 and December 2020. It was necessary to fulfill the following six criteria for any research to be included in the meta-analytic research. As presented in Figs. 1 and 63 research studies that met the following criteria were deemed eligible for this meta-analytic research based on the PRISMA flow-chart (Moher et al., 2009).

- The research had to be published between 2000 and 2020.
- The research had to be the nature of an experimental or quasi-experimental design.
- For the calculation of the WBMI effect size, the research had to provide adequate information for meta-data.
- Research had to focus on K-16 students' mathematics learning in the context of WBMI.
- K-16 students' mathematics performance had to be one of the outcome variables in the research.

Fig. 1 The PRISMA flow-chart used in this meta-analytic research



- TMI or alternative teaching had to be considered for the control condition in the research.

4.2 Coding plan

To elucidate the effectiveness of WBMI on K-16 students’ mathematics learning, potential moderators emphasized in the literature were handled in this meta-analysis. A detailed coding plan was developed for coding procedures. Codes were generated to characterize the research studies involved in this meta-analytic research (see Fig. 2).

The coding form included the study identification tag, the categories of mathematics topics (i.e., foundational concepts, higher-level mathematical concepts, and mathematical skills), the components of mathematical content standards (i.e., numbers & operations, geometry, algebra, statistics and probability, and mixed), feedback status (i.e., WBMI with or without providing feedback), the types of instructional features (i.e., drill-and-practice programs, simulations, tutorial systems, and ITS), grade level (i.e., elementary, middle, high school, and undergraduate level), assessment methods (i.e., traditional paper-pencil assessment, and online assessment), and quantitative information to calculate the effect size of each study regarding the effectiveness of WBMI on K-16 students’ mathematics learning (e.g., *n*, *M*, and *SD*). Additionally, it is considered that each category of moderator variables is independent research when research contains multiple categories of moderator variables. Consequently, the present study carried out a meta-analysis of 63 research, with 115 effect sizes to examine the effectiveness of WBMI on K-16 students’ mathematics learning. As seen in Table 1, general features of these 63 studies with 115 effect sizes regarding poten-

Fig. 2 Codes for this meta-analytic research features. (adapted from Hu, Chen, Li, & Huang, 2021)

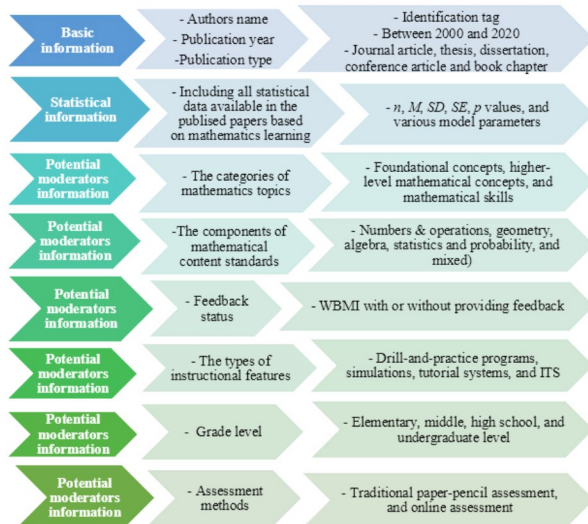


Table 1 General features regarding potential moderators in the meta-analysis

Potential moderators	<i>f</i>	%
Mathematics topics		
Foundational concepts	39	33.9
Higher-level mathematical concepts	64	55.7
Mathematical skills	12	10.4
Mathematical content standards		
Numbers and operations	35	30.5
Geometry	12	10.4
Algebra	30	26.1
Statistics and probability	19	16.5
Mixed	19	16.5
Feedback status		
WBMI with providing feedback	80	69.5
WBMI without providing feedback	35	30.5
Instructional features		
Drill-and-practice programs	32	27.8
Simulations	4	3.5
Tutorial	55	47.8
ITS	24	20.9
Grade level		
Elementary	11	9.6
Middle	43	37.4
High	29	25.2
Undergraduate	32	27.8
Assessment methods		
Traditional paper-pencil assessment	97	84.3
Online assessment	18	15.7

Note. *f* frequency

tial moderators are characterized in this meta-analytic research. Moreover, eligible studies were independently coded by two researchers, obtained Cohen's Kappa was 0.95 revealing a nearly perfect agreement (Landis & Koch, 1977).

4.3 Data analysis

Hedges's *g* was considered as an effect size measure for each study to demonstrate the effectiveness of WBMI on K-16 students' mathematics learning regarding this meta-analysis. The effect size is determined by dividing the mean difference between groups by the standard deviation of the control group in independent group designs (Sun et al., 2021). Since the nature of this study was within the scope of pretest-post-test-control groups design, the effect size was determined by subtracting the mean of the pre-post change in the experimental group from the mean of the pre-post change in the control group and dividing it into the pooled pre-test standard deviation (Borenstein et al., 2014).

The effect sizes for the experimental and control groups were computed utilizing the pre-and post-test dataset. The effect sizes were computed using the means and standard deviations reported by the majority of the research studies for outcomes based on students' mathematics learning. Several of the research studies simply

reported mean changes and p values for the change. In accordance with the thumb rule suggested by Thalheimer and Cook (2002), the effect size was characterized as follows. -0.15 to 0.15 as negligible, 0.15 to 0.40 as low, 0.40 to 0.75 as moderate, 0.75 to 1.10 as large, 1.10 to 1.45 as very large and above 1.45 as a huge effect size. Additionally, the negative sign before the effect sizes imply that the effect favors the control group, while the positive sign implies that the effect favors the experimental group (Koydemir, Sokmez, & Schutz, 2021).

The test for homogeneity was performed by computing Q , I^2 , and p to determine whether a fixed or random-effects model best fit the meta-data (Borenstein et al., 2014; Hillmayr et al., 2020). The statistical criteria provided by Higgins et al. (2003), I^2 value was clarified as follows 25–50% as low, 50–75% as moderate, above 75% as a highly heterogeneous distribution of effect sizes. Additionally, a random effect model is used if the p -value for Q -statistics is below 0.05 and the I^2 value above 60% (Warrier, 2018). Mixed-effects models were also performed to investigate whether moderator variables were responsible for the diversity of effect sizes in this meta-analysis (Sun et al., 2021). Moreover, the difference between related subgroups/moderators was analyzed using the Q -between-groups test (Q_B) that represented heterogeneity between groups and was equal to the F value in the analysis of variance (Hillmayr et al., 2020). To further assess the effects of multi-covariates on overall effect size, a meta-regression was performed.

In this meta-analysis, publication bias was examined using a funnel plot, a fail-safe N test (FSN), Orwin's fail-safe N (FSN) test, and Duval and Tweedie's trim-and-fill (DTEK) test (Borenstein et al., 2014; Duval & Tweedie, 2000; Rosenthal, 1979). Many researchers emphasize that the research is resistant to publication bias if the effect size of any research displays a symmetrical distribution along the vertical line using the funnel plot analysis (e.g., Juandi et al., 2021; Tang & Liu, 2000). If the funnel plot analysis indicates a reasonable symmetrical distribution, it is recommended to use FSN and Orwin's FSN test to investigate publication bias (Borenstein et al., 2009). For the research to be considered resistant to publication bias, the FSN value is expected to be above 1 (Mullen et al., 2001). The analysis of Orwin's FSN test reveals how many missing studies with a zero-effect size would be required to minimize the overall effect size to a trivial level (Koydemir et al., 2021). According to the DTEK test, if there is no difference between the observed values and the adjusted values, it is argued that the research is resistant to publication bias (Celiker, Ustunel, & Guzeller, 2019). All analyses were performed using the Comprehensive Meta-analysis (CMA) and JASP (Borenstein et al., 2014; JASP, 2021).

5 Results

5.1 The overall effect size of WBMI on students' mathematics learning

The sample of this meta-analysis research contained 115 individual effect sizes obtained from 30,207 students who took part in 63 research which examined the effectiveness of WBMI on K-16 students' mathematics learning. The number of students in each research ranged from 17 to 2499. Since the first research question of the

meta-analytic research was related to the overall effect of WBMI on K-16 students' mathematics learning, the findings of the first research question were clarified by considering the random-effects model. The Q statistic ($Q = 4137.22$, $df = 114$, $p < .01$) was statistically significant in the meta-analysis research, using the test of homogeneity (Sun et al., 2021). Additionally, there were actual variations in effect sizes across research studies reflecting 97.25% of the observed variance, according to the criteria of Higgins et al. (2003) associated with the I^2 value (see Table 2). The overall effect of WBMI on K-16 students' mathematics learning was large and statistically significant ($g = 1.10$, $SE = 0.08$, $p = .01$, 95% CI [0.95, 1.27]) based on Thalheimer and Cook's (2002) criteria. Consequently, the findings of the meta-analytic research indicated that there was a statistically significant difference in mathematics achievement between participants who used WBMI and participants who used TMI or alternative teaching. Moreover, a significant and strong effect size for WBMI allowed us to infer that WBMI was more effective on students' mathematics learning than TMI.

5.2 Moderator analyses

For moderator analysis, it is emphasized that a statistically significant heterogeneous distribution of effect sizes is necessary (Lipsey & Wilson, 2001). Moderator analyses were performed for the six moderator variables (mathematics topics, mathematical content standards, feedback status, instructional features, grade level, and assessment methods) in this research since this criterion was provided. First, the effect sizes of WBMI on K-16 students' mathematics learning varied significantly depending on the categories of mathematical skills ($Q_b(2) = 6.49$, $p < .05$). Mean effect sizes for all categories of mathematics topics were significantly positive. A list of all mean effect sizes is given in Table 3. Higher-level mathematical concepts had the greatest mean effect size ($g = 1.28$, 95% CI [1.07, 1.50], $p < .01$), pursued by mathematical skills ($g = 1.20$, 95% CI [0.70, 1.70] $p < .01$), and foundational concepts ($g = 0.83$ 95% CI [0.55, 1.11], $p < .01$). Meta-regression analyses revealed that the mean effect size for foundational concepts was significantly lower than for higher-level mathematical concepts ($\beta = 0.43$, $z = 2.51$, $p < .05$), and that for mathematical skills ($\beta = 0.31$, $z = 2.29$, $p < .01$). Moreover, mean effect sizes for higher-level mathematical concepts and mathematical skills were comparable and large based on Thalheimer and Cook's (2002) criteria.

Second, effect sizes differed significantly by the components of mathematical content standards ($Q_b(4) = 92.65$, $p < .01$). The findings indicated that statistics and probability yielded the greatest and significant effect size ($g = 2.86$, 95% CI [2.47, 3.27], $p < .01$) followed by numbers and operations ($g = 1.00$, 95% CI [0.71, 1.29], $p < .01$), and algebra ($g = 0.73$, 95% CI [0.43, 1.03], $p < .01$), indicating a large effect

Table 2 The overall effect of WBMI on K-16 students' mathematics learning

	% 95 CI						ToH		
	<i>k</i>	<i>n</i>	<i>g</i>	LL	UP	<i>p</i>	<i>Q</i>	<i>p</i>	<i>I</i> ²
Overall effect of WBMI	115	30,207	1.10**	0.95	1.27	0.00	4137.22**	0.00	97.25

Note. CI confidence interval; ToH test of homogeneity ** $p < .01$

size, whereas geometry ($g = 0.55$, 95% CI [0.07, 1.04], $p < .05$) and mixed ($g = 0.67$, 95% CI [0.29, 1.04], $p < .01$) yielded a moderate and significant effect size. Additionally, meta-regression analyses demonstrated that the mean effect size for statistics and probability was significantly larger than numbers and operations ($\beta = -2.26$, $z = -4.35$, $p < .01$), geometry ($\beta = -2.74$, $z = -4.09$, $p < .01$), algebra ($\beta = -2.58$, $z = -4.84$, $p < .01$), and mixed ($\beta = -2.64$, $z = -4.48$, $p < .01$).

Third, the findings revealed that the mean effect sizes differed significantly depending on the feedback status ($Q_b(1) = 15.40$, $p < .05$). The mean effect size for studies within the context of WBMI with providing feedback ($g = 1.33$, 95% CI [1.14, 1.52], $p < .01$) was significantly greater than studies within the context of WBMI without providing feedback ($g = 0.64$, 95% CI [0.35, 0.93], $p < .01$) based on meta-regression analysis ($\beta = -0.82$, $z = -2.03$, $p < .05$). Fourth, a significant between-level variance was observed as a result of the type of instructional features ($Q_b(3) = 19.32$, $p < .01$). The findings indicated that tutorial systems ($g = 1.45$, 95% CI [1.22, 1.68], $p < .01$), and ITS ($g = 1.10$, 95% CI [0.76, 1.44], $p < .01$) produced a large and comparable mean effect size, while drill-and-practice programs ($g = 0.64$, 95% CI [0.34, 0.93], $p < .01$), and simulations ($g = 0.62$, 95% CI [-0.25, 1.49], $p > .05$) yielded a moderate and comparable mean effect size. Moreover, meta-regression analyses revealed the mean effect size for the tutorial systems was significantly larger than for drill-and-practice programs ($\beta = -0.48$, $z = -2.17$, $p < .05$), and that for simulations ($\beta = -0.52$, $z = -5.90$, $p < .05$). Although ITS produced a large mean effect size, there were no significant differences in the mean effect size of ITS between drill-and-practice programs ($\beta = -0.46$, $z = -0.87$, $p > .05$), and simulations ($\beta = -0.36$, $z = -1.56$, $p > .05$), individually.

Fifth, WBMI showed a significantly positive effect on mathematics learning across all grade levels ($Q_b(3) = 37.99$, $p < .01$), including elementary students ($g = 0.59$, 95% CI [0.07, 1.11], $p < .01$), middle school students ($g = 0.98$, 95% CI [0.72, 1.24], $p < .01$), high school students ($g = 0.69$, 95% CI [0.37, 1.00], $p < .01$), and undergraduate students ($g = 1.91$, 95% CI [1.61, 2.22], $p < .01$). Studies including elementary and high school students both yielded a moderate and comparable mean effect size, whereas studies including middle school students produced a large effect size, and studies including undergraduate students yielded a huge effect size based on Thalheimer and Cook's (2002) criteria. Additionally, meta-regression analyses revealed that the mean effect size for studies including undergraduate students was significantly larger than for studies including elementary ($\beta = -0.90$, $z = -7.5$, $p < .01$), middle ($\beta = -0.58$, $z = -6.21$, $p < .01$), and high school students ($\beta = -0.76$, $z = -6.33$, $p < .01$).

Sixth, the results indicated that the mean effect sizes differed significantly depending on the assessment methods ($Q_b(1) = 5.76$, $p < .01$). Within the scope of WBMI, the mean effect size for traditional paper-pencil assessment ($g = 1.20$, 95% CI [1.03, 1.37], $p < .01$) was significantly greater than online assessment ($g = 0.68$, 95% CI [0.28, 1.07], $p < .01$) based on meta-regression analysis ($\beta = -0.65$, $z = -6.26$, $p < .05$). Studies including traditional paper-pencil assessment yielded a very large mean effect size, whereas studies including online assessment produced a moderate effect size based on Thalheimer and Cook's (2002) criteria.

Table 3 The effect sizes of WBMI on K-16 students' mathematics learning based on moderator analyses

Potential moderators	<i>k</i>	<i>g</i>	<i>g</i> (95% CI)	<i>Q_b</i> (df)
Mathematics topics				6.49*(2)
Foundational concepts	39	0.83**	[0.55, 1.11]	
Higher-level mathematical concepts	64	1.28**	[1.07, 1.50]	
Mathematical skills	12	1.20**	[0.70, 1.70]	
Mathematical content standards				92.65**(4)
Numbers and operations	35	1.00**	[0.71, 1.29]	
Geometry	12	0.55*	[0.07, 1.04]	
Algebra	30	0.73**	[0.43, 1.03]	
Statistics and probability	19	2.86**	[2.47, 3.27]	
Mixed	19	0.67**	[0.29, 1.04]	
Feedback status				15.40**(1)
WBMI with providing feedback	80	1.33**	[1.14, 1.52]	
WBMI without providing feedback	35	0.64**	[0.35, 0.93]	
Instructional features				19.32** (3)
Drill-and-practice programs	32	0.64**	[0.34, 0.93]	
Simulations	4	0.62	[-0.25, 1.49]	
Tutorial	55	1.45**	[1.22, 1.68]	
ITS	24	1.10**	[0.76, 1.44]	
Grade level				37.99**(3)
Elementary	11	0.59**	[0.57, 0.72]	
Middle	43	0.98**	[0.72, 1.24]	
High	29	0.69**	[0.37, 1.00]	
Undergraduate	32	1.91**	[1.61, 2.22]	
Assessment methods				5.76**(1)
Traditional paper-pencil assessment	97	1.20**	[1.03, 1.37]	
Online assessment	18	0.68**	[0.28, 1.07]	

* $p < 0.05$; ** $p < 0.01$

5.3 Publication bias

For this meta-analysis research, publication bias analyses were performed to evaluate if the research data was affected by publication bias or not. Figure 3 showed a funnel plot analysis that revealed a relatively symmetrical distribution. Consequently, the FSN and Orwin's FSN tests were considered to evaluate the likelihood of publication. To eliminate the significant effect at $p > .05$, 11.87 studies were required according to the findings of the FSN test (Mullen et al., 2001). Depending on Orwin's FSN analysis with a trivial g dataset at 0.01 level, 548 missing studies with a trivial effect size would be necessary to reduce the overall effect size to an almost zero (i.e., trivial) effect. Based on Duval and Tweedie's trim-and-fill test, findings demonstrated no difference between the observed value of the random effects model ($g = 1.10$) and the adjusted value of the random effects model ($g = 1.10$). Koydemir, Sokmez, and

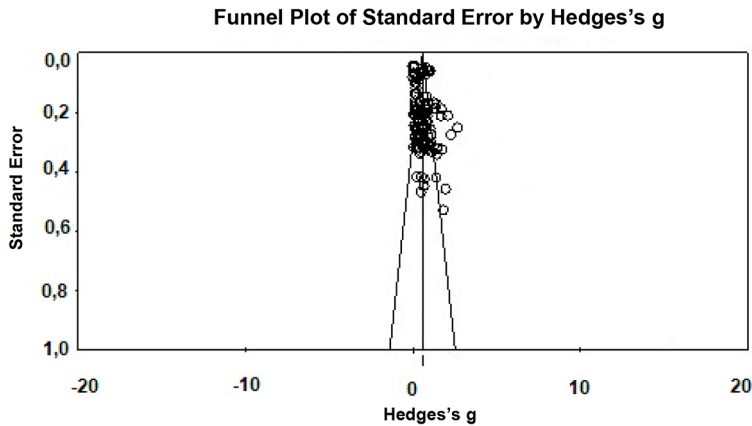


Fig. 3 Funnel chart according to the studies analyzed in this research

Schutz (2021) emphasize that the nature of publication bias can be subjectively interpreted, but this meta-analysis research appears to be immune and resistant to publication bias based on these estimates of publication bias analyses.

6 Discussion

6.1 Summary of the overall effect size of WBMI on K-16 students' mathematics learning

This meta-analytic research contained a total of 63 studies with 115 effect sizes, which aimed to investigate the effectiveness of WBMI on K-16 students' mathematics learning by incorporating potential moderators, namely mathematics topics, mathematical content standards, feedback status, type of instructional features, age (i.e., grade level), and assessment methods. The findings of this research reveal that WBMI yields a statistically significant and large effect size on K-16 students' mathematics learning ($g = 1.10$) based on Thalheimer and Cook's (2002) criteria. This effect size value of 1.10 implies that nearly 86% of the individuals in the experimental group performed above the mean of the individuals in the control group (Coe, 2002).

6.2 Comparison of findings regarding the overall effect size of WBMI on K-16 students' mathematics learning with previous research

The result of the overall positive effect size of WBMI on K-16 students' mathematics learning is largely in accordance with studies of the effectiveness of mobile learning (Guler et al., 2021), Assessment and Learning in Knowledge Spaces (ALEKS) (Fang et al., 2019; Sun et al., 2021), and learning with digital tools (Hillmayr et al., 2020) on students' mathematics achievement. It reveals that the overall effect size of this

research is quite larger than compared to the effect sizes of previous meta-analysis studies such as Sun et al. (2021) with $g = 0.05$, Guler et al. (2021) with $g = 0.48$, and Hillmayr et al. (2020) with $g = 0.65$. Moreover, it is seen that these previous meta-analysis studies produced small to negligible, small to moderate, or moderate effect sizes. Therefore, this result of the present research is distinguishable from the results of past meta-analyses. The reason for this difference can be explained by the focus only on learning mathematics in this meta-analytic study, and the rapid growth of web-based tools and digital learning in the last few years (Hillmayr et al., 2020).

Based on the results of this meta-analytic research, WBMI has a significantly stronger effect on K-16 students' mathematics learning than TMI. The result is consistent with many previous studies (e.g., Gu & Lee, 2019; Guzeller & Akin, 2012; Lin, 2009; Nguyen & Kulm, 2005) indicating that the mean mathematics achievement of WBMI students is significantly higher than the mean mathematics achievement of TMI students. In contrast, the result is inconsistent with several studies (e.g., Baki & Guveli, 2008; Martindale et al., 2005; Mman & Tudunkaya, 2019) indicating no significant difference between WBMI students and TMI students regarding mathematics learning. This study reveals that WBMI is more effective on students' mathematics learning than TMI in terms of meta-analysis findings. The reason for this finding may be related to the fact that the use of high-quality and well-planned web-based math learning environments leads to higher mathematical performance (Misirli & Ergulec, 2021; Palloff & Pratt, 2013). Therefore, it can be argued that most of the previous studies included in this meta-analysis have high-quality web-based mathematics learning environments.

Although this study indicates that WBMI is more effective on students' mathematics learning than TMI, other studies in the literature have reported that compared to online classes, students' mathematics performance in face-to-face classes is better (e.g., Amro, Mundy, Kupczynski, 2015; Flanagan, 2012; Heppen et al., 2017; Li, Uvah, Amin, & Hemasinha, 2009). This situation can be explained by several possible explanations. Since most of the online mathematics classes are still in their technological infancy, they do not meet the needs of students and fail to adapt to advanced technology (Flanagan, 2012). The experience of participating in online courses also affects the mathematics performance of the students (Amro et al., 2015). A previous study revealed that students who had previously taken and passed online courses had a higher achievement score in subsequent online courses (Beyrer, 2010). Additionally, many students are still unfamiliar with online courses due to learning habits related to face-to-face courses (Cao et al., 2021). It has been revealed in many studies that the rate of dropout and satisfaction from online mathematics courses is higher than face-to-face courses. (e.g., Jagggers, 2014; Smith & Ferguson, 2005; Summers, Waigandt; Whittaker, 2005; Zavarella & Ignash, 2009). On the other hand, recent studies have indicated that students have significantly higher academic performance in STEM-related online courses compared to face-to-face courses (AbdelSalam, Pilotti, & El-Moussa, 2021; Gonzalez et al., 2020; Iglesias-Pradas et al., 2021). The reason for this situation is that the COVID-19 pandemic has improved students' digital skills and changed their learning strategies from discontinuous habits to continuous habits (Gonzalez et al., 2020). The COVID 19 pandemic has also ensured that students do not become unfamiliar with online courses and can easily adapt to these

courses. Therefore, it is argued that online learning environments can be used as a crucial tool to improve STEM-related performance in the post-pandemic world (AbdelSalam et al., 2021).

6.3 Summary of findings regarding potential moderators

Six types of potential moderators namely mathematics topics, mathematical content standards, feedback status, type of instructional features, age (i.e., grade level), and assessment methods were investigated in this meta-analysis. The findings of the research clarified that the effect sizes of WBMI on K-16 students' mathematics learning varied significantly depending on all these potential moderators.

6.4 Comparison of findings regarding potential moderators with previous research

6.4.1 Moderating effect of mathematics topics

Regarding mathematics topics, the effect sizes found for higher-level mathematical concepts and mathematical skills were significantly larger than for foundational concepts. This result can be explained by the complicated nature of school mathematics topics. Previous studies emphasize that WBMI assists students to comprehend higher-level mathematical concepts such as function and trigonometry, which are viewed as difficult to learn in TMI (e.g., Gu & Lee, 2019; Guler et al., 2021; Guzeller & Akin, 2012; Lin, 2009). Moreover, web-based mathematics learning environments are beneficial for allowing individuals to master complicated skills such as spatial skills through visualization (Rafi et al., 2005). On the other hand, poorly adapted web-based mathematics learning environments are problematic, especially in the learning of higher-level mathematical concepts, and it is not easy for students to adapt and maintain their interest in this environment (Gu & Lee, 2019; Moos & Azevedo, 2006; Smith & Ferguson, 2005). At this point, the overall effect size of WBMI on K-16 students' mathematics learning may be influenced by the quality of the web-based mathematics learning environment. Therefore, it can be argued that the use of well-planned web-based mathematics learning environments in the context of high-level mathematical concepts and mathematical skills leads to higher mathematical performance, hence producing larger effect sizes in this meta-analysis research.

6.4.2 Moderating effect of mathematical content standards

With respect to mathematical content standards, all components of mathematical content standards were significantly positive moderators. The findings indicated that the effect size found for statistics and probability was significantly larger than for other components of mathematical content standards, which was in accordance with the research hypothesis. The findings of a recent meta-analysis study have revealed that the domain of statistics and probability produced the greatest effect size compared to other mathematical content standards in the context of the mobile learning environment (Guler et al., 2021). One possible reason for this result is the fact that many

highly adaptive web-based learning environments have been created in the domain of statistics and probability (e.g., Hurlburt, 2001; Ozyurt et al., 2014; Muhanna & Abu-Al-Sha'r, 2010). Moreover, the initial research associated with WBMI focused on students' mathematics learning in the domain of statistics and probability in the context of distance and remote learning technologies (Couch, 1997; Hurlburt, 2001). Thanks to the flexibility and freedom provided by web-based mathematics learning environments, it can be stated that both the moderator of mathematics topics and mathematical content standards produced significantly positive effect sizes in this meta-analysis study (Misirli & Ergulec, 2021).

6.4.3 Moderating effect of feedback status

Regarding feedback status, the effect size found for studies in the context of WBMI with providing feedback was significantly larger than for studies in the context of WBMI without providing feedback. Previous studies highlight the benefits of feedback even in TMI but emphasize that feedback plays a much more key role in WBMI, as it is one of the most valuable assets of the WBMI (Nguyen & Kulm, 2005). Consequently, this is not a surprising finding given that the use of feedback in WBMI provides numerous benefits, such as helping students develop productive ways of thinking, and contributing to self-regulation skills (Gu & Lee, 2019; Nguyen & Kulm, 2005).

6.4.4 Moderating effect of instructional features

Concerning various types of instructional features, this research revealed that tutorial systems and ITS yielded significantly large effect sizes, whereas drill-and-practice programs and simulations produced significantly moderate effect sizes. This finding conforms with the findings of previous studies that indicating the effect sizes of ITS and tutorial systems are significantly larger than for drill-and-practice programs (e.g., Bayraktar, 2001, Hillmayr et al., 2020), and revealing simulations in the context of virtual reality yield moderate effect sizes (Hillmayr et al., 2020). The synergistic characteristics of ITS and tutorial systems, including feedback, engagement of necessary knowledge, and adjustment of learning mathematics topics to prerequisite skills and concepts, might account for the greater effect sizes of ITS and tutorial systems (Hillmayr et al., 2020). The reason why drill-and-practice programs produce a moderate effect size could be related to the negative features inherent. These are the fact that drill-and-practice programs have difficulty in adapting to the previous knowledge, do not allow the construction of new knowledge, only aim to strengthen previously learned knowledge and concepts (e.g., Bayraktar, 2001, Hillmayr et al., 2020).

6.4.5 Moderating effect of grade level

The findings of the meta-analytic research indicated that WBMI was positively effective at all levels of schooling. The studies conducted at elementary and high school levels showed moderate effect sizes, whereas studies conducted middle school level produced a large effect size, and studies conducted at the undergraduate level demon-

strated a huge effect size. Moreover, the effect size of studies conducted at the undergraduate level was significantly greater than the studies conducted at the elementary, middle, and high school levels. This result conforms with the results of Hillmayr et al. (2020) and the hypothesis of Steenbergen-Hu and Cooper (2014) that WBMI could be more useful for adults who have better self-control abilities, digital skills, previous mathematical knowledge, and motivational beliefs than younger children. One of the reasons for the smallest effect size concerning studies conducted at the elementary level might be that WBMI hinders younger students' mathematics learning owing to the overuse of unaided, uncontrolled, and undirected mathematical activities (Berger et al., 1994; Muhanna & Abu-Al-Sha'r, 2010). Consequently, it can be referred that this result is not surprising based on previous studies.

6.4.6 Moderating effect of assessment methods

Concerning assessment methods, the effect size found for studies regarding WBMI using traditional paper-pencil assessment was significantly larger than for the studies regarding WBMI using online assessment. The results of this meta-analytic research imply that students who are assessed via traditional paper-pencil methods earn higher mathematics achievement test scores than the students who are assessed via the online testing modules/methods. Due to the nature of mathematics, it requires students to use cognitive skills such as reasoning, problem-solving, and calculation. From the earliest ages, it has been seen that students use paper and pencil to solve mathematical problems. Thus, with the help of paper and pencil, students can reveal their mental processes and strategies related to a mathematical problem. On the other hand, solving mathematical problems and performing calculations in online exams are the biggest challenges addressed by students due to the inadequacy of the online system in terms of recording mathematical calculations and ease of use of the calculator (Ilgaz & Afacan-Adanir, 2020; Laine et al., 2016). It is also emphasized that most of the students have difficulty in time management in online mathematics exams, which negatively affects their mathematics performance (Ilgaz & Afacan-Adanir, 2020).

The result is consistent with many previous studies (e.g., Dandurand et al., 2008; Escudier et al., 2011; Jones & Long, 2013; Still & Still, 2015) indicating that traditional paper-pencil assessment methods have shown to be more effective than online assessment. In contrast, the result is inconsistent with several studies (Ricketts & Wilks, 2002; Rane & MacKenzie, 2020; Pennebaker et al., 2013) indicating online assessment is more effective than traditional paper-pencil assessment in the context of mathematics. Although it has numerous claims for the benefits of online assessment and determined many advantages, online assessment is not as effective and strong as traditional paper-pencil assessment regarding students' mathematics achievement based on the results of this study. Additionally, this result does not conform with the hypothesis of Charman and Elmes (1998) that online assessment can increase students' academic achievement and thus, students' learning.

The present study carried out a meta-analysis of 63 research with 115 effect sizes in the context of WBMI, but only 18 of the 115 effect sizes were used in online assessment. This finding can be explained by one possible explanation that designing effective online exams is not an easy task. It is difficult to implement formative assess-

ment procedures in online courses because of the large number of students they serve concurrently (Ilgaz & Afacan-Adanir, 2020). Moreover, it is revealed that it is more tiring for individuals to read the text on the computer screen than to read the same text in print (Mourant, Lakshmanan, & Chantadisai, 1981). Specifically, in online mathematics examinations that require computational skills, it is also observed that students use paper-pen and computer simultaneously to solve mathematical problems (Ilgaz & Afacan-Adanir, 2020). Additionally, marking online exams is a very time-consuming task (Engelbrecht & Harding, 2004). Therefore, students have difficulty in terms of time management, especially in online exams (Ilgaz & Afacan-Adanir, 2020). For many reasons mentioned above, it is thought that traditional paper-pencil assessment is preferred more than online assessment even in the context of WBMI based on previous studies.

7 Limitations and directions for further studies

The present meta-analytic research includes some limitations. The most important of these limitations may be due to the search strategy. The combination of keywords “web-based mathematics instruction”, “WBMI”, “WMT”, “web-based learning”, “online learning”, “online mathematics”, “web-based course”, “web-based” and “online” were considered in this research. However, in some publications, authors have used proper or special names such as WbVE and UZWEBMAT (Ozyurt et al., 2014; Rafi et al., 2005) instead of a web-based mathematics learning environment or WBMI. It has also been noted that some publications focused on specific skills such as spatial rotation rather than mathematical concepts. Therefore, such situations might weaken the power of search strategy in the context of this meta-analysis research. Another limitation is that a limited number of effect sizes for simulations makes estimations regarding the moderator effect of simulations in terms of the effectiveness of WBMI on K-16 students’ mathematics learning less certain. As a result, more effect sizes for simulations might be beneficial for further studies to elucidate this moderator effect. Additionally, various types of feedback could not be included as a potential moderator variable since many publications did not include the necessary data for this meta-analysis research. Further limitations of this research are that potential moderator variables such as gender, the type of publication, learning approaches, learning materials, duration of implementation, and year were not included in this meta-analysis. Therefore, future meta-analysis studies that include adequate data for moderators such as the type of feedback, learning approaches, learning materials, and duration of implementation are required to shed light on the effectiveness of WBMI on K-16 students’ mathematics learning by integrating other moderators. In further meta-analysis studies, it may be illuminating to investigate whether the effectiveness of WBMI varies by socioeconomic characteristics to understand the equitable use of web technology for all students.

8 Conclusions

In this post-pandemic world, the usage of WBMI and online assessment raises many questions that have not been completely investigated yet. This meta-analytic research can shed light on the impact of WBMI on students' mathematics learning comprehensively. The findings of this meta-analytic research suggest that WBMI is promising and has the potential to improve mathematics learning and instruction across all levels of education since its efficiency is proven to be superior to TMI. Mathematics topics, mathematical content standards, feedback status, type of instructional features, age, and assessment methods significantly moderate the effects of WBMI on mathematics learning. Moderation analyses demonstrate that higher-level mathematical concepts, statistics and probability, WBMI with providing feedback, tutorial systems, undergraduate students, and traditional paper-pencil assessment are the strongest moderators in their context. In contrast, foundational concepts, geometry, WBMI without providing feedback, drill-and-practice programs, elementary students, and online assessment are the weakest moderators in their context compared to all these potential moderators. The most notable results of this study are that WBMI is more effective on students' mathematics learning than TMI, while even in the context of WBMI, traditional paper-pencil assessment is more effective than online assessment. At this point, this study can shed light on the effects of WBMI and online assessment on students' mathematics learning in a comprehensive and detailed way.

The results of the present research have crucial educational implications. Although this meta-analysis research shows that WBMI is superior and more effective than TMI, the reason for the large effect size of this study may be associated with the quality of the web-based mathematics learning environments. Because WBMI is not a simple process, but rather a complicated process that needs detailed instructional design cycles, practice, and assessment to create a fruitful learning environment within the context of a well-planned WBMI (Misirli & Ergulec, 2021; Palloff & Pratt, 2013). Therefore, it can be argued that the use of well-planned web-based mathematics learning environments leads to higher mathematical performance, hence producing larger effect sizes in this meta-analysis research. Given the fact that web-based mathematics learning environments will be used more frequently and widely in the post-pandemic world, it would be useful for educational policymakers and mathematics educators to offer mathematics teachers opportunities for well-planned learning environments in terms of WBMI, rather than simply encouraging mathematics teachers to use the WBMI (Hillmayr et al., 2020).

The results of this meta-analytic research shed light on what kind of assessment methods are more useful and stronger than others in the context of WBMI. The current research indicates that online mathematics assessment is not as effective and strong as traditional paper-pencil assessment. The reason for this finding may be related to difficulties encountered in online assessment. These are factors such as students having difficulty in time management, and the inadequacy of the online system in terms of recording mathematical calculations, and ease of use of the calculator (Ilgaz & Afacan-Adanir, 2020; Laine et al., 2016). At this point, it is necessary to develop online systems that allow students to overcome the difficulties experienced in online mathematics exams. If such online systems are developed in the context of

online mathematics assessment, it can be claimed that online assessment will be as effective and strong as traditional paper-pencil assessment in future meta-analytic studies. Consequently, this meta-analytic research provides a comprehensive and up-to-date perspective on the effectiveness of WBMI and assessment methods on K-16 students' mathematics learning since educational stakeholders need to know to which extent the effects of WBMI on mathematics learning in the post-pandemic world (Sun et al., 2021).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

- AbdelSalam, H. M., Pilotti, M. A., & El-Moussa, O. J. (2021). Sustainable math education of female students during a pandemic: Online versus face-to-face instruction. *Sustainability*, 13(21), 1–12. <https://doi.org/10.3390/su132112248>
- Abersson, C. L., Berger, D. E., Healy, M. R., Kyle, D. J., & Romero, V. L. (2000). Evaluation of an interactive tutorial for teaching the central limit theorem. *Teaching of Psychology*, 27(4), 289–291. https://doi.org/10.1207/S15328023TOP2704_08
- Agasisti, T., Gil-Izquierdo, M., & Han, S. W. (2020). ICT use at home for school-related tasks: What is the effect on a student's achievement? Empirical evidence from OECD PISA data. *Education Economics*, 28(6), 601–620. <https://doi.org/10.1080/09645292.2020.1822787>
- Amro, H., Mundy, M., & Kupeczynski, L. (2015). The effects of age and gender on student achievement in face-to-face and online college algebra classes. *Research in Higher Education Journal*, 27, 1–22
- Babbitt, W., Lachney, M., Bulley, E., & Eglash, R. (2015). Adinkra mathematics: A study of ethnocomputing in Ghana. *Multidisciplinary Journal of Educational Research*, 5(2), 110–135. <https://doi.org/10.17583/remie.2015.1399>
- Baki, A., & Guveli, E. (2008). Evaluation of a web-based mathematics teaching material on the subject of functions. *Computers & Education*, 51(2), 854–863. <https://doi.org/10.1016/j.compedu.2007.09.003>
- Baran, E., Correia, A. P., & Thompson, A. (2011). Transforming online teaching practice: Critical analysis of the literature on the roles and competencies of online teachers. *Distance Education*, 32(3), 421–439. <https://doi.org/10.1080/01587919.2011.610293>
- Bayraktar, S. (2001). A meta-analysis of the effectiveness of computer-assisted instruction in science education. *Journal of Research on Technology in Education*, 34(2), 173–188. <https://doi.org/10.1080/15391523.2001.10782344>
- Berger, C. F., Lu, C. R., Beltzer, S. J., & Voss, B. E. (1994). Research on the uses of technology in science education. In D. L. Gabel (Ed.), *Handbook of research on science teaching and learning* (4th ed., pp. 466–490). Macmillan
- Beyrer, G. M. (2010). Online student success: Making a difference. *Journal of Online Learning and Teaching*, 6(1), 89–109
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley-Blackwell
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2014). *Comprehensive meta-analysis (Version 3.3.070)*. Englewood, NJ: Biostat
- Brouwer, N., Ekimova, L., Jasinska, M., van Gastel, L., & Virgailaitė-Mečauskaitė, E. (2009). Enhancing mathematics by online assessments: Two cases of remedial education. *Industry and Higher Education*, 23(4), 277–283. <https://doi.org/10.5367/000000009789346095>
- Cady, J., & Rearden, K. (2009). Delivering online professional development in mathematics to rural educators. *Journal of Technology and Teacher Education*, 17(3), 281–298
- Cao, Y., Zhang, S., Chan, M. C. E., & Kang, Y. (2021). Post-pandemic reflections: lessons from Chinese mathematics teachers about online mathematics instruction. *Asia Pacific Education Review*, 22(2), 157–168. <https://doi.org/10.1007/s12564-021-09694-w>

- Celiker, N., Ustunel, M. F., & Guzeller, C. O. (2019). The relationship between emotional labour and burn-out: A meta-analysis. *Anatolia*, 30(3), 328–345. <https://doi.org/10.1080/13032917.2019.1581625>
- Charman, D., & Elmes, A. (1998). Formative assessment in a basic geographical statistics module. *Computer Based Assessment*, 2, 17–20
- Coe, R. (2002, September 12). *It's the effect size, stupid*. Paper presented at the Annual Conference of the British Educational Research Association, University of Exeter, Exeter, UK. Retrieved November 22, 2021, from <http://www.leeds.ac.uk/educol/documents/00002182.htm>
- Couch, J. V. (1997). Using the internet in instruction: A homepage for statistics. *Psychological Reports*, 81(3), 999–1003. <https://doi.org/10.2466/pr0.1997.81.3.999>
- Dandurand, F., Shultz, T. R., & Onishi, K. H. (2008). Comparing online and lab methods in a problem-solving experiment. *Behavior Research Methods*, 40(2), 428–434. <https://doi.org/10.3758/BRM.40.2.428>
- Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (2020, December 8). *COVID-19 and learning loss: Disparities grow and students need help*. New York, NY: McKinsey. Retrieved November 22, 2021, from <https://www.mckinsey.com/industries/public-and-social-sector/our-insights/covid-19-and-learning-loss-disparities-grow-and-students-need-help>
- Duval, S., & Tweedie, R. (2000). Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2), 455–463. <https://doi.org/10.1111/j.0006-341X.2000.00455.x>
- Engelbrecht, J., & Harding, A. (2004). Combing online and paper assessment in a web-based course in undergraduate mathematics. *Journal of Computers in Mathematics and Science Teaching*, 23(3), 217–231
- Escudier, M. P., Newton, T. J., Cox, M. J., Reynolds, P. A., & Odell, E. W. (2011). University students' attainment and perceptions of computer delivered assessment; a comparison between computer-based and traditional tests in a high-stakes examination. *Journal of Computer Assisted Learning*, 27(5), 440–447. <https://doi.org/10.1111/j.1365-2729.2011.00409.x>
- Fang, Y., Ren, Z., Hu, X., & Graesser, A. C. (2019). A meta-analysis of the effectiveness of ALEKS on learning. *Educational Psychology*, 39(10), 1278–1292. <https://doi.org/10.1080/01443410.2018.1495829>
- Flanagan, J. (2012). Online versus face-to-face instruction: Analysis of gender and course format in undergraduate business statistics courses. *Academy of Business Research*, 2, 93–101
- Gonzalez, T., De La Rubia, M. A., Hincz, K. P., Comas-Lopez, M., Subirats, L., Fort, S., & Sacha, G. M. (2020). Influence of COVID-19 confinement on students' performance in higher education. *PLoS One*, 15(10), 1–23. <https://doi.org/10.1371/journal.pone.0239490>
- Gu, P., & Lee, Y. (2019). Promoting students' motivation and use of SRL strategies in the web-based mathematics learning environment. *Journal of Educational Technology Systems*, 47(3), 391–410. <https://doi.org/10.1177/0047239518808522>
- Guler, M., Butuner, S. O., Danisman, S., & Gursoy, K. (2021). A meta-analysis of the impact of mobile learning on mathematics achievement. *Education and Information Technologies*, 1–21. <https://doi.org/10.1007/s10639-021-10640-x>
- Guzeller, C. O., & Akin, A. (2012). The effect of web-based mathematics instruction on mathematics achievement, attitudes, anxiety, and self-efficacy of 6th grade students. *International Journal of Academic Research in Progressive Education and Development*, 1(2), 42–54
- Guzeller, C. O., & Celiker, N. (2019). Examining the relationship between organizational commitment and turnover intention via a meta-analysis. *International Journal of Culture, Tourism and Hospitality Research*, 14(1), 102–120. <https://doi.org/10.1108/IJCTHR-05-2019-0094>
- Heppen, J. B., Sorensen, N., Allensworth, E., Walters, K., Rickles, J., Taylor, S. S., & Michelman, V. (2017). The struggle to pass algebra: Online vs. face-to-face credit recovery for at-risk urban students. *Journal of Research on Educational Effectiveness*, 10(2), 272–296. <https://doi.org/10.1080/19345747.2016.1168500>
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, 327(7414), 557–560. <https://doi.org/10.1136/bmj.327.7414.557>
- Hillmayr, D., Ziernwald, L., Reinhold, F., Hofer, S. I., & Reiss, K. M. (2020). The potential of digital tools to enhance mathematics and science learning in secondary schools: A context-specific meta-analysis. *Computers & Education*, 153, 1–25. <https://doi.org/10.1016/j.compedu.2020.103897>
- Hu, L., Chen, G., Li, P., & Huang, J. (2021). Multimedia effect in problem solving: A meta-analysis. *Educational Psychology Review*, 33, 1717–1747. <https://doi.org/10.1007/s10648-019-09490-4>
- Hurlburt, R. T. (2001). Lectlets deliver content at a distance: Introductory statistics as a case study. *Teaching of Psychology*, 28(1), 15–20. https://doi.org/10.1207/S15328023TOP2801_04

- Iglesias-Pradas, S., Hernández-García, Á., Chaparro-Peláez, J., & Prieto, J. L. (2021). Emergency remote teaching and students' academic performance in higher education during the COVID-19 pandemic: A case study. *Computers in Human Behavior*, 119, 1–18. <https://doi.org/10.1016/j.chb.2021.106713>
- Ilgaz, H., & Afacan-Adanir, G. (2020). Providing online exams for online learners: Does it really matter for them? *Education and Information Technologies*, 25(2), 1255–1269. <https://doi.org/10.1007/s10639-019-10020-6>
- Jagers, S. S. (2014). Choosing between online and face-to-face courses: Community college student voices. *British Journal of Educational Technology*, 28(1), 27–38. <https://doi.org/10.1080/08923647.2014.867697>
- JASP, T. (2021, November 16). *JASP* (Version 0.16). Retrieved November 22, 2021, from <https://jasp-stats.org>
- Jitendra, A. K., Lein, A. E., Im, S. H., Alghamdi, A. A., Hefte, S. B., & Mouanoutoua, J. (2018). Mathematical interventions for secondary students with learning disabilities and mathematics difficulties: A meta-analysis. *Exceptional Children*, 84(2), 177–196. <https://doi.org/10.1177/0014402917737467>
- Jones, S. J., & Long, V. M. (2013). Learning equity between online and on-site mathematics courses. *Journal of Online Learning & Teaching*, 9(1), 1–12
- Juandi, D., Kusumah, Y. S., Tamur, M., Perbowo, K. S., & Wijaya, T. T. (2021). A meta-analysis of Geogebra software decade of assisted mathematics learning: what to learn and where to go? *Heliyon*, 7(5), 1–8. <https://doi.org/10.1016/j.heliyon.2021.e06953>
- Kennedy, K., Nowak, S., Raghuraman, R., Thomas, J., & Davis, S. F. (2000). Academic dishonesty and distance learning: Student and faculty views. *College Student Journal*, 34(2), 309–314
- Koydemir, S., Sokmez, A. B., & Schutz, A. (2021). A meta-analysis of the effectiveness of randomized controlled positive psychological interventions on subjective and psychological well-being. *Applied Research in Quality of Life*, 16(3), 1145–1185. <https://doi.org/10.1007/s11482-019-09788-z>
- Kurtulus, A., & Kilic, R. (2009). The effects of webquest assisted cooperative learning method on the achievement towards mathematics lesson. *Education Sciences*, 4(1), 62–70
- Laine, K., Sipilä, E., Anderson, M., & Sydänheimo, L. (2016, September 12). *Electronic exam in electronics studies*. Paper presented at the SEFI Annual Conference 2016: Engineering Education on Top of the World: Industry University Cooperation. Tampere, Finland. Retrieved November 22, 2021, from <https://www.sefi.be/wp-content/uploads/2017/09/laine-electronic-exam-in-electronics-studies-9.pdf>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>
- Li, K., Uvah, J., Amin, R., & Hemasinha, R. (2009). A study of non-traditional instruction on qualitative reasoning and problem solving in general studies mathematics courses. *Journal of Mathematical Sciences and Mathematical Education*, 4(1), 37–49
- Lin, C. (2009). A comparison study of web-based and traditional instruction on preservice teachers' knowledge of fractions. *Contemporary Issues in Technology and Teacher Education*, 9(3), 257–279
- Lin, C. (2011). *The development of intelligence and mathematical learning*. China Light Industry Press
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE
- MacGregor, S., & Lou, Y. (2004). Web-based learning: How task scaffolding and web site design support knowledge acquisition. *Journal of Research on Technology in Education*, 37(2), 161–175. <https://doi.org/10.1080/15391523.2004.10782431>
- Martindale, T., Pearson, C., Curda, L. K., & Pilcher, J. (2005). Effects of an online instructional application on reading and mathematics standardized test scores. *Journal of Research on Technology in Education*, 37(4), 349–360. <https://doi.org/10.1080/15391523.2005.10782442>
- Mman, S. S., & Tudunkaya, M. S. (2019). Effects of web-based practice on academic performance in coordinate geometry among colleges of education students in north-west zone, Nigeria. *Abacus Mathematics Education Series*, 44(1), 90–95
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Reprint—preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Physical Therapy*, 89(9), 873–880. <https://doi.org/10.1093/ptj/89.9.873>
- Moos, D. C., & Azevedo, R. (2006). The role of goal structure in undergraduates' use of self-regulatory processes in two hypermedia learning tasks. *Journal of Educational Multimedia and Hypermedia*, 15(1), 49–86
- Mourant, R. R., Lakshmanan, R., & Chantadisai, R. (1981). Visual fatigue and cathode ray tube display terminals. *Human Factors*, 23(5), 529–540

- Misirli, O., & Ergulec, F. (2021). Emergency remote teaching during the COVID-19 pandemic: Parents experiences and perspectives. *Education and Information Technologies*, 26, 6699–6718. <https://doi.org/10.1007/s10639-021-10520-4>
- Muhanna, W. N. (2010). & Abu-Al-Sha'r, A. M. The impact of using web-based curricula on ninth grade students' achievement in mathematics. In P. Kommers, T. Issa & P. Isaias (Eds.), *Internet Technologies & Society 2010* (1st ed., pp. 164–171). IADIS Press
- Mullen, B., Muellerleile, P., & Bryant, B. (2001). Cumulative meta-analysis: A consideration of indicators of sufficiency and stability. *Personality and Social Psychology Bulletin*, 27(11), 1450–1462. <https://doi.org/10.1177/01461672012711006>
- Nguyen, D. M., & Kulm, G. (2005). Using web-based practice to enhance mathematics learning and achievement. *Journal of Interactive Online Learning*, 3(3), 1–16
- Ozyurt, O., Ozyurt, H., Guven, B., & Baki, A. (2014). The effects of UZWEBMAT on the probability unit achievement of Turkish eleventh grade students and the reasons for such effects. *Computers & Education*, 75, 1–18. <https://doi.org/10.1016/j.compedu.2014.02.005>
- Palloff, R. M., & Pratt, K. (2013). *Lessons from the virtual classroom*. Jossey-Bass
- Pennebaker, J. W., Gosling, S. D., & Ferrell, J. D. (2013). Daily online testing in large classes: Boosting college performance while reducing achievement gaps. *PLoS One*, 8(11), 1–6. <https://doi.org/10.1371/journal.pone.0079774>
- Rafi, A., Anuar, K., Samad, A., Hayati, M., & Mahadzir, M. (2005). Improving spatial ability using a web-based virtual environment (WbVE). *Automation in Construction*, 14(6), 707–715. <https://doi.org/10.1016/j.autcon.2004.12.003>
- Rane, V., & MacKenzie, C. A. (2020). Evaluating students with online testing modules in engineering economics: A comparison of student performance with online testing and with traditional assessments. *The Engineering Economist*, 65(3), 213–235. <https://doi.org/10.1080/0013791X.2020.1784336>
- Ricketts, C., & Wilks, S. J. (2002). Improving student performance through computer-based assessment: Insights from recent research. *Assessment & Evaluation in Higher Education*, 27(5), 475–479. <https://doi.org/10.1080/026029302200009348>
- Román-Caballero, R., Vadillo, M. A., Trainor, L., & Lupiáñez, J. (2021). Please don't stop the music: A meta-analysis of the benefits of learning to play an instrument on cognitive and academic skills. *PsyArXiv*. <https://doi.org/10.31234/osf.io/4bm8v>
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86, 638–641. <https://doi.org/10.1037/0033-2909.86.3.638>
- Sitzmann, T., Kraiger, K., Stewart, D., & Wisher, R. (2006). The comparative effectiveness of web-based and classroom instruction: A meta-analysis. *Personnel Psychology*, 59(3), 623–664. <https://doi.org/10.1111/j.1744-6570.2006.00049.x>
- Smith, G. G., & Ferguson, D. (2005). Student attrition in mathematics e-learning. *Australasian Journal of Educational Technology*, 21(3), 323–334. <https://doi.org/10.14742/ajet.1323>
- Steenbergen-Hu, S., & Cooper, H. (2014). A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. *Journal of Educational Psychology*, 106(2), 331–347. <https://doi.org/10.1037/a0034752>
- Stephens, D. (2001). Use of computer assisted assessment: Benefits to students and staff. *Education for Information*, 19(4), 265–275
- Still, M. L., & Still, J. D. (2015). Contrasting traditional in-class exams with frequent online testing. *Journal of Teaching and Learning with Technology*, 4(2), 30–40. <https://doi.org/10.1080/15393101.2015.1055555>
- Summers, J. J., Waigandt, A., & Whittaker, T. A. (2005). A comparison of student achievement and satisfaction in an online versus a traditional face-to-face statistics class. *Innovative Higher Education*, 29(3), 233–250. <https://doi.org/10.1007/s10755-005-1938-x>
- Sun, S., Else-Quest, N. M., Hodges, L. C., French, A. M., & Dowling, R. (2021). The effects of ALEKS on mathematics learning in K-12 and higher education: A meta-analysis. *Investigations in Mathematics Learning*, 13(3), 182–196. <https://doi.org/10.1080/19477503.2021.1926194>
- Tang, J. L., & Liu, J. L. (2000). Misleading funnel plot for detection of bias in meta-analysis. *Journal of Clinical Epidemiology*, 53(5), 477–484. [https://doi.org/10.1016/S0895-4356\(99\)00204-8](https://doi.org/10.1016/S0895-4356(99)00204-8)
- Taylor, J. M. (2008). The effects of a computerized-algebra program on mathematics achievement of college and university freshmen enrolled in a developmental mathematics course. *Journal of College Reading, and Learning*, 39(1), 35–53. <https://doi.org/10.1080/10790195.2008.10850311>
- Thalheimer, W., & Cook, S. (2002, August 1). *How to calculate effect sizes from published research articles: A simplified methodology*. Work-Learning Research. Retrieved November 22, 2021, from http://work-learning.com/effect_sizes.htm

- Tsai, C., & Shen, P. (2009). Applying web-enabled self-regulated learning and problem-based learning with initiation to involve low-achieving students in learning. *Computers in Human Behavior*, 25(6), 1189–1194. <https://doi.org/10.1016/j.chb.2009.05.013>
- Ulum, H. (2022). The effects of online education on academic success: A meta-analysis study. *Education and Information Technologies*, 27(1), 429–450 <https://doi.org/10.1007/s10639-021-10740-8>
- Warrier, V. (2018). *The genetics of autism and related traits*. Unpublished Doctoral Dissertation. University of Cambridge
- Zavarella, C. A., & Ignash, J. M. (2009). Instructional delivery in developmental mathematics: Impact on retention. *Journal of Developmental Education*, 32(3), 2–13

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Ayça Akin¹

✉ Ayça Akin
aycaakin07@gmail.com

¹ Alumni Association, Yunusemre Campus Student Central, First Floor, Anadolu University, 26470 Tepebaşı/Eskişehir, Turkey