Journal of Science Education and Technology (2022) 31:621–637
https://doi.org/10.1007/s10956-022-09982-z

Exploring the Effectiveness and Moderators of Augmented Reality on Science Learning: a Meta-analysis

Wen-Wen Xu[1] · Chien-Yuan Su[2] · Yue Hu[3] · Cheng-Huan Chen[4]

Accepted: 22 June 2022 / Published online: 5 July 2022

Abstract
The use of augmented reality (AR) technology in the science curriculum has the potential to assist students in comprehending abstract and complex concepts or unobservable phenomena, as well as to better explain knowledge regarding science content by superimposing virtual objects over genuine items or environments in a multidimensional approach. However, the overall effects on students' academic achievement of using AR technology in scientific courses and the key factors that influence such effects are still unclear. Therefore, we performed a meta-analysis in this work to systematically review 35 empiric trials (with 39 effect sizes) that used experimental or quasi-experimental approaches to determine the academic achievement of using AR techniques in science-related courses. In addition, we explored possible moderators such as differences in disciplines, educational stages, types of AR (marker-based, markerless-based, or location-based), display devices (mobiles, tablets, computers, or headsets), intervention duration, group size, and instructional strategies. The results revealed that the overall mean effect size (with AR into instruction vs without AR into instruction) was 0.737 under the random effects model, indicating a medium-to-large significant positive effect on students' academic achievement. The disciplines had significant moderating effects, types of AR had marginally significant effects, while educational stages, display devices, intervention duration, group size, and strategies used had insignificant influence. The impact of AR technology on scientific education was discussed in connection to the above seven moderators.

Keywords Augmented reality · Science learning · Meta-analysis · Moderators · Academic achievement

Introduction

Augmented reality (AR) technology superimposes virtual items into the actual environment, making it possible to convey complex scientific information in easier-to-understand ways, explain abstract science concepts, or demonstrate science phenomena that are difficult to observe firsthand (Cheng & Tsai, 2013; Sahin & Yilmaz, 2020; Walczak et al., 2006). Unlike virtual reality, which completely immerses a user in a synthetic environment, augmented reality allows users to see the real world with virtual features overlayed upon it in real time and is seen to have greater potential for science learning (Chang et al., 2020; Cheng & Tsai, 2013; Gnidovec et al., 2020). As a result, it has grabbed the attention of educators, practitioners, and academics, and has been applied to science education. A number of empirical studies have revealed that incorporating AR into the science curriculum (e.g., physics, chemistry, Earth science, biology, mathematics) can enhance student scientific learning, such as improving content understanding and interest in science (Radu, 2014), increasing science learning motivation and engagement (Cai et al., 2013; Diegmann et al., 2015; Goff et al., 2018), and improving academic achievement (Akçayır & Akçayır, 2017; Lu et al., 2020). For example, Hsiao et al. (2016) discovered that students who used a manipulative AR system (which included 3D interactive models and manipulative aids, technologies that enhanced the interactivity and utility of AR) had significantly better academic achievement and learning motivation than students who utilized

✉ Chien-Yuan Su
cysu@mail.nutn.edu.tw

[1] Department of Curriculum and Learning Sciences, Zhejiang University, Hangzhou, China

[2] Department of Education, National University of Tainan, Tainan, Taiwan

[3] Department of Educational Technology, Hangzhou Normal University, Hangzhou, China

[4] Department of M-Commerce and Multimedia Applications, Asia University, Taichung, Taiwan

multimedia resources to learn natural science. Akçayır et al. (2016) also explored the impact of AR technologies on university students' laboratory skills in science laboratories, discovering that AR technology greatly increased the growth of such skills. Chang and Hwang (2018) investigated the learning outcomes of elementary students in flipped-learning based experiences for physics, both with and without using AR. They discovered that students in AR-based flipped learning outperformed those without AR in academic achievement, motivation, critical thinking, and group self-efficiency. Lu et al. (2020) investigated the effect of an AR application on elementary school students' natural science learning achievement for understanding rocks and minerals, with results showing that the AR-based group scored substantially higher than the paper-based group. Sahin and Yilmaz (2020) used AR to visualize science concepts and phenomena in terms of the solar system and electromagnetism. This method helps students better comprehend science knowledge and promotes more engagement and a positive attitude.

Despite much previous research revealing that AR usage might increase student academic achievement in scientific learning, some research revealed no meaningful effect. For example, Erbas and Demirer (2019) employed AR in a middle school biology course and found no significant difference in academic achievement between those who used AR and those who did not. Thees et al. (2020) applied AR in a college physics laboratory experiment of heat conduction and compared their knowledge gains with those of a traditional group. The findings revealed that there was no difference in their academic achievements. Chien et al. (2019) found no significant difference in academic achievement between the groups using AR and herbarium specimens when learning in the introductory course entitled "Plant Stem." Dehghani et al. (2020) found no significant difference in academic achievement between senior students using static infographics and those using AR. In addition, the application of AR may vary among different age groups (Wu et al., 2013), while the use of different devices (such as desktop PCs, smartphones, and head-mounted displays) can result in different learning experiences (Garzón & Acevedo, 2019; Ozdemir et al., 2018; Radu, 2014); thus, student learning outcomes may be moderated in science learning. For example, Juan et al. (2010) argued that children rated the head-mounted AR system as less user-friendly and found themselves prone to dizziness, while tablets were deemed more suitable for teaching and were preferred by students for some AR activities (Fokides & Mastrokoukou, 2018). Garzón and Acevedo (2019) demonstrated that the effect size of AR for learning arts and humanities is larger than it is for natural sciences and mathematics; moreover, Ozdemir et al. (2018) found that the effect size of AR was relatively larger in natural sciences than in social sciences (e.g., economics, political sciences, psychology, and sociology). As for educational stages, some studies found no significant difference in educational levels (Garzón et al., 2019; Ozdemir et al., 2018). In addition, various forms of AR display, such as location-based and image-based displays (Wojciechowski & Cellary, 2013), have been shown to have different affordances for learning (Cheng & Tsai, 2013). The results above may imply that the impact of augmented reality on science learning is unclear due to discipline differences, educational stages, types of AR, display devices, and so on (Cai et al., 2014, 2017; Santos et al., 2014).

Furthermore, several instructional (e.g., learning strategy or teaching method) or experimental treatments (e.g., intervention duration, group size) (Chen & Yang, 2019; Sung et al., 2016) may have an influence on the effectiveness of science education, but have not been investigated in previous research. For example, Dehghani et al. (2020) reported that a combined teaching method using infographics and AR showed significant improvement compared to simply utilizing infographics or AR, suggesting that the combination of technology and learning strategies may have a more significant effect. Fidan and Tuncel (2019) examined students' achievement under diverse circumstances, including AR with problem-based learning, problem-based learning alone, and traditional teaching. They discovered that AR with problem-based learning was more effective than the other groups, demonstrating that integrating learning strategies into AR may improve learning effectiveness. Furthermore, nearly half of the studies on AR-assisted scientific learning were conducted in small groups; however, it is unclear if individuals or groups might moderate learning achievement.

Prior studies have indeed been conducted and reported on the benefits of employing AR technology to improve student learning. For example, Garzón and Acevedo (2019) performed a systematic review of 64 publications and a meta-analysis of 27 studies on AR applications for education across a variety of fields, including natural sciences, arts and humanities, social sciences, information and communication technologies, and health and welfare, while Ozdemir et al. (2018) analyzed 16 research projects from 2007 to 2017 to determine the impact of augmented reality applications on the learning process, including both natural science and social science. Despite the fact that the average effect of AR in a wide range of educational fields can be seen in the above two meta-analyses, the effect size of AR in various science-related disciplines may be significantly different. Thus, it is necessary to further investigate the effect of AR in student science learning. Moreover, previous research on moderator analysis has focused only on education areas, grade levels, display devices, and sample size, with little research on the various types of AR, intervention duration, group size, and learning strategies. Nevertheless, there has been no comprehensive evaluation of the application of AR in scientific education and its impact size on students' academic achievement, especially when different types of AR, group distribution, and teaching techniques used are taken into account.

As a result, this study attempts to systematically examine the previous research results concerning the influence of AR use in science education on students' academic accomplishments, as measured by grades or performance on educational achievement tests (Wigfield & Cambria, 2010). Furthermore, the study compares the size of such effects in various moderator variables, such as disciplines (domain subjects), educational stages, types of AR (e.g., marker-based AR must employ a marker as a trigger; markerless-based AR estimates the camera pose using visual or depth information from the captured natural scene; location-based AR provides AR features depending on the user's geographic location (Cheng & Tsai, 2013)), display devices (mobiles, tablets, computers, or headsets), intervention duration, group size, and learning strategies (e.g., self-directed learning, game-based learning, project-based learning, problem-based learning, or conventional methods). Accordingly, two research questions are posed as follows:

RQ1: What are the overall effects (i.e., overall weighted mean effect size) of using AR in science education on student academic achievement (such as grades or performance on educational achievement tests)?

RQ2: Do disciplines, educational stages, types of AR, display devices, intervention duration, group size, and learning strategies significantly influence the effects of AR on student academic achievement in science learning?

Method

We conducted a meta-analysis to evaluate the academic achievement of students when they use AR techniques to perform science learning activities, as well as following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) Statement (Page et al., 2021) to guide the search, screening, and extraction process, which are detailed below.

Study Searching

Articles were retrieved from an electronic search of educational databases in October 2020: namely, the Web of Science (WOS) Core Collection, EBSCOhost, Scopus, ProQuest, and ScienceDirect, with the publication time span restricted to 2020. The following search terms were used: ("AR" OR "augmented reality" OR "augmented reality system" OR "AR tools" OR "AR application") AND ("science" OR "science education" OR "scientific experiment" OR "physics" OR "biology" OR "chemistry" OR "mathematics" OR "Earth science") AND ("learn" OR "learning" OR "teaching" OR "learning outcomes" OR "education" OR "educational" OR "course" OR "instruction" OR "pedagogy"). As a result, the search yielded 803 papers (253 in the WOS Core Collection, 63 in EBSCOhost, 289 in Scopus, 139 in ProQuest, 59 in ScienceDirect), with 184 being duplicated and 619 remaining after removal.

Article Selection Process

Table 1 lists the inclusion and exclusion criteria for this systematic review. This effort determined the initial eligibility. Two educational researchers independently examined the titles and abstracts of 619 publications based on the inclusion criterion in the first round of screening, with an agreement rate of around 91.92% (Chen & Yang, 2019). Following this, 527 articles were eliminated, leaving 92 studies to be further examined. In the second stage of screening, these two researchers reviewed the whole text of the 92 papers based on the exclusion criterion and addressed articles that were contradictory in order to choose publications that satisfied the meta-analysis criterion. The degree to which the two researchers agreed was 91.30%. Finally, 57 publications were eliminated, leaving

Table 1 Inclusion and exclusion criteria

Inclusion criteria	It must be published in an English-language journal
	It must be on educational research with a focus on science (physics, biology, chemistry, earth science) or mathematics
	The use of AR apps, tools, or systems is required
Exclusion criteria	The full text is not available in the database(s)
	The study did not conduct experimental or quasi-experimental investigations
	AR technologies were not the primary independent variable in the study
	The study did not address learning achievement or describe academic achievement as a major dependent variable
	The amount of data in the study is insufficient to compute an effect size (sample sizes, means, standard deviations, or *t*, *F* values)

35 research papers suitable for meta-analysis. Figure 1 depicts a flow chart of the selection process.

Data Coding

The 35 studies that met the inclusion criteria were coded in three main parts: basic information, main content, and statistics for meta-analysis. Basic information includes the author's name, year of publication, and the title of the article. The main content includes research objectives, results, and the seven moderators. The research questions frame the seven moderators chosen for analysis: disciplines (e.g., physics, chemistry, Earth science, biology, mathematics [as mathematics and science education have a greater degree of overlap, "mathematics" was included in the search criteria and coded in this research], and some related science themes which are not specified), educational stages (e.g., preschools- and elementary schools, middle schools, high schools, colleges/universities), types of AR (e.g., marker-based AR, markerless-based AR, location-based AR), display devices of AR (e.g., mobile, tablet, computer combined with camera, and both tablet and mobile phone are used), intervention duration (e.g., no more than 24 h, 1 day to 1 week, 1 to 4 weeks, 1 to 4 months), group size of students (not specified, individual, two to four, more than four), and strategies employed (e.g., self-directed learning, game-based learning, project-based learning, problem-based learning, or conventional method). We defined self-directed learning as flipped class, learning without a teacher or teachers acting simply

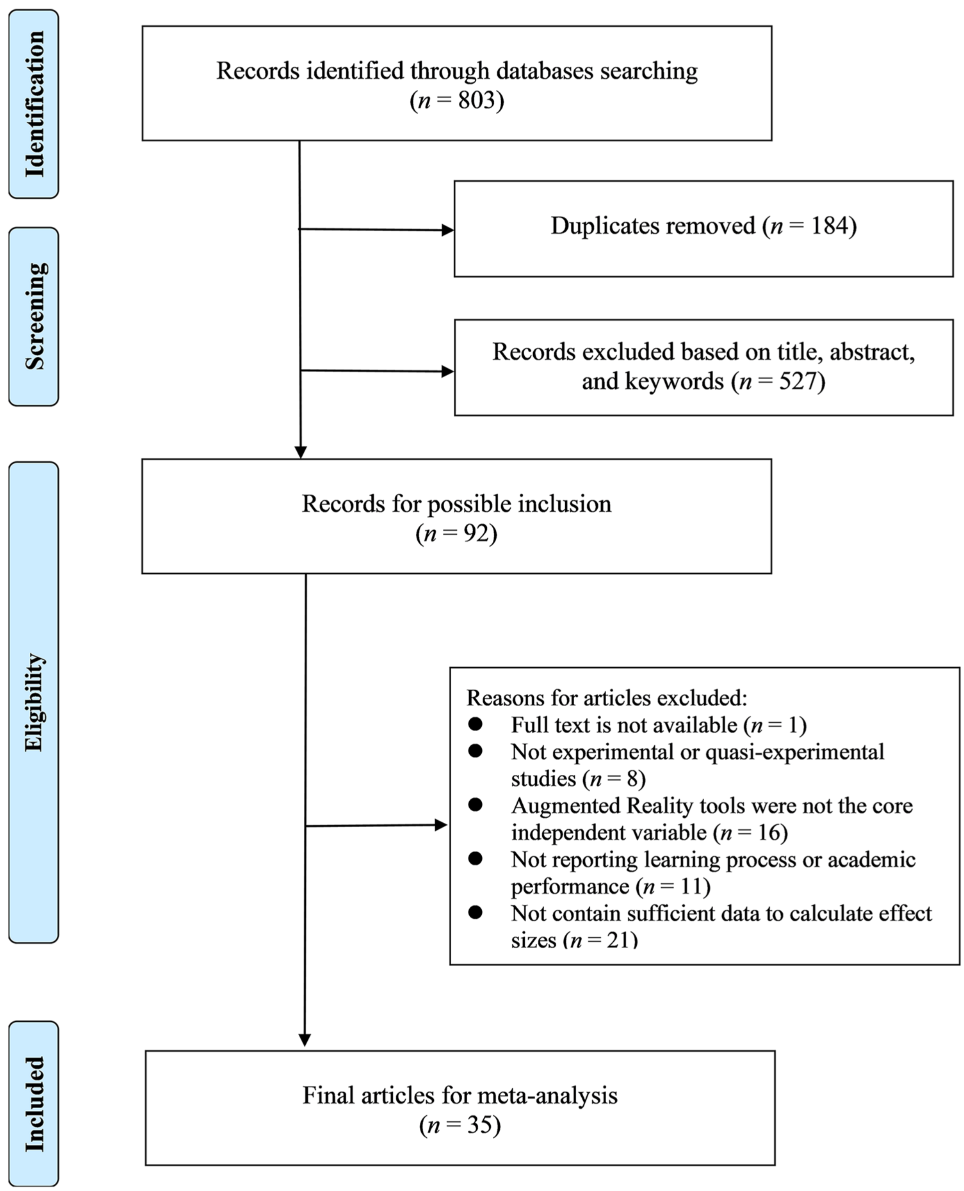

Fig. 1 A diagrammatic representation of the PRISMA review process

as facilitators (Silén & Uhlin, 2008), while the conventional method is best described as teacher-based instruction in which teachers introduce course content or explain concepts first, and then present AR or engage students to operate it (Dange, 2018).

Data Analysis

The meta-analytical procedure was to (a) calculate effect sizes of each study's outcome measure on a standard scale; (b) calculate overall effect sizes and test for heterogeneity; (c) investigate the moderating effects of study characteristics on the outcome measure; and (d) investigate publication bias (Hu et al., 2021). In this study, the Comprehensive Meta-Analysis (CMA) software (Version 3) (Borenstein et al., 2013) was used to calculate all statistical analyses.

The effect size of each study was calculated. Hedges's *g* was chosen as the standard predictor of mean weighted effect sizes since it has the optimal qualities for small samples (Borenstein et al., 2009). Furthermore, we used Cohen's criterion to quantify the effect size in our sample, with 0.2 indicating a minor effect, 0.5 indicating a moderate effect, and 0.8 indicating a large effect (Cohen, 1988). A 95% confidence interval (CI) for Hedges's *g* was used to test significant differences. After calculating the effect sizes for each individual study, the overall weighted mean effect size was calculated using the random effects model selected in the CMA software.

Heterogeneity was estimated using the Cochran's Q statistic and the I^2 statistic. The observed dispersion of impact sizes is represented by the Q statistic (Q_T). The I^2 coefficient determines how much variation between experiments is due to actual variance rather than sampling bias (Borenstein et al., 2009). If considerable variance existed, further moderator analyses were needed to determine if the moderator variables were responsible for the heterogeneity (Higgins et al., 2003; Lipsey & Wilson, 2001).

Furthermore, the random effects model was used to perform a moderator analysis to detect associations between moderator variables and impact sizes. Between-group homogeneity (Q_B) was used to investigate the moderators. Q_B explores the homogeneity of effect sizes between groups, and its significance level indicates the significant influence of the potential moderator on the variance across groups. Additionally, the trim-and-fill approach (Duval & Tweedie, 2000) was used to evaluate publication bias by measuring the number of missed studies, producing a modified mean effect size by adding the missing studies on the skewed side; moreover, the Egger's regression (Bowden et al., 2015) intercept was used to detect small study bias in this meta-analysis.

Results and Discussion

Effect Sizes of Each Selected Study

Table 2 summarizes the details of the 35 qualifying published studies, which included a total of 2625 participants and 39 comparisons (effect sizes). The effect sizes of the selected studies ranged from − 1.15 to 3.603, all of which were within three standard deviations of the total effect size; hence, no studies were eliminated (Lipsey & Wilson, 2001). Among them, 22 comparisons (56.4%) demonstrated statistically significant positive effects, 16 comparisons (41.0%) failed to reveal significant effects, and only 1 comparisons (2.6%) demonstrated statistically significant negative results.

Overall Effect Sizes and Testing for Heterogeneity

Table 3 shows the results of the overall effect size analysis and the homogeneity test. The random effects model analysis found a mean effect size of 0.737 (95% CI was 0.506–0.969). The results showed that AR had a significantly greater effect on students' academic achievement than teaching without AR technology, with a medium-to-large level, according to Cohen's criterion.

In addition, Duval and Tweedie's trim-and-fill approach was employed to count the number of missing trials and adjust the mean effect size (see in Table 4). Both models discovered 0 missing studies to the left of the mean. A fixed effects model found 7 missing studies and adjusted the overall effect size from 0.701 to 0.882 ($p < .05$); a random effects model to trim-and-fill found 11 missing studies and adjusted the overall effect size from 0.737 to 1.060 ($p < .05$). Overall, the adjusted overall effect sizes are greater than the observed values—in fact, they indicate a larger effect size. Furthermore, Egger's test for a regression intercept gave a *p*-value of 0.235 (1 tailed), indicating no evidence of publication bias in this meta-analysis.

These findings are consistent with those of meta-analysis results from Garzón and Acevedo (2019) (covering 27 studies published between 2010 and 2018, $d = 0.68$, $p < .001$) and Ozdemir et al. (2018) (covering 16 studies published between 2007 and 2017, $d = 0.517$, $p < .001$), revealing that AR had a medium effect on learning effectiveness in the learning process. To compare the results of these two research projects on the size of augmented reality (AR) effects on learning achievement in general education settings, this study discovered a significant medium-to-large effect size of AR technology on students' academic achievement in science learning. According to the findings, incorporating AR into scientific courses can improve students'

Table 2 Characteristics and effect sizes of selected studies

Authors (year)	Num	*N*	EG	CG	Discipline	Educational stage	Types of AR	Device types of AR	Duration	Group size	Strategy	Measurement of academic achievement	Hedges's *g*
Abdusselam and Karal (2020)	2	69	24	25	Physics	High school	Marker-based	Computer	> 1 week and ≤ 4 weeks	Two to four	Self-directed learning	Achievement test scores	0.511
				20	Physics	High school	Marker-based	Computer	> 1 week and ≤ 4 weeks	Two to four	Self-directed learning	Achievement test scores	0.331
Akçayır et al. (2016)	1	76	38	38	Physics	College/ university	Marker-based	Mobile	> 1 month and ≤ 4 months	More than four	Self-directed learning	The lab skills test scores	0.557*
Cai et al. (2013)	1	50	24	26	Physics	Middle school	Marker-based	Computer	Not mention	Not specified	Conventional method	A paper and pencil test scores	0.116
Cai et al. (2020)	1	68	33	35	Mathematics	Middle school	Marker-based	Tablet	> 1 month and ≤ 4 months	Two to four	Game-based learning	Test scores	0.025
Carrera and Asensio (2017)	1	123	63	60	Earth science	College/ university	Marker-based	Tablet	Not mention	Individual	Conventional method	Perspective-taking/ spatial orientation test scores	0.986*
Chang et al. (2016)	1	55	28	27	Biology	Elementary school	Marker-based	Computer	≤ 24 h	Individual	Conventional method	Achievement test scores	0.391
Chang and Yu (2018)	1	93	47	46	Biology	College/ university	Markerless-based	Mobile	> 1 week and ≤ 4 weeks	More than four	Self-directed learning	The lab skills test scores	0.378
Chang and Hwang (2018)	1	111	56	55	Physics	Elementary school	Markerless-based	Tablet	> 1 month and ≤ 4 months	Two to four	Project-based learning	Learning achievement test scores	0.257
Chen (2020)	2	74	25	25	Biology	Elementary school	Markerless-based	Tablet	> 1 week and ≤ 4 weeks	Individual	Game-based learning	Learning achievement test scores	0.216
			24		Biology	Elementary school	Markerless-based	Tablet	> 1 week and ≤ 4 weeks	Individual	Self-directed learning	Learning achievement test scores	0.573*
Chiang et al. (2014)	1	57	28	29	Biology	Elementary school	Markerless-based	Tablet	≤ 24 h	Individual	Self-directed learning	Learning achievement test scores	0.531*
Chien et al. (2019)	1	45	23	22	Biology	Elementary school	Markerless-based	Tablet	≤ 24 h	Individual	Conventional method	Academic achievement test scores	− 0.136
Dehghani et al. (2020)	1	50	25	25	Biology	High school	Markerless-based	Tablet	> 1 week and ≤ 4 weeks	More than four	Conventional method	Learning achievement test scores	− 0.147

Table 2 (continued)

Authors (year)	Num	*N*	EG	CG	Discipline	Educational stage	Types of AR	Device types of AR	Duration	Group size	Strategy	Measurement of academic achievement	Hedges's *g*
Demitriadou et al. (2020)	1	20	10	10	Mathematics	Elementary school	Marker-based	Tablet	Not mention	Individual	Game-based learning	Test scores	0.710
Erbas and Demirer (2019)	1	40	20	20	Biology	Middle school	Marker-based	Tablet	> 1 month and ≤ 4 months	Not specified	Conventional method	Achievement test scores	− 0.231
Faridi et al. (2021)	1	80	40	40	Physics	College/ university	Marker-based	Mobile	≤ 24 h	Individual	Conventional method	Knowledge test scores	1.36*
Fidan and Tuncel (2019)	1	60	30	30	Physics	Middle school	Marker-based	Tablet	> 1 month and ≤ 4 months	Two to four	Problem-based learning	Achievement test scores	1.498*
Fokides and Mastrokoukou (2018)	2	75	25	25	Biology	Elementary school	Markerless-based	Tablet	> 1 week and ≤ 4 weeks	Two to four	Self-directed learning	Achievement test scores	1.095*
				25	Biology	Elementary school	Markerless-based	Tablet	> 1 week and ≤ 4 weeks	Two to four	Self-directed learning	Achievement test scores	0.521
Gecu-Parmaksiz and Delialioglu (2019)	1	72	36	36	Mathematics	Preschool	Marker-based	Tablet	> 1 month and ≤ 4 months	Not specified	Conventional method	The Geometric Shape Recognition Task scores	1.155*
Gopalan et al. (2016)	1	140	70	70	Related science themes	Middle school	Marker-based	Computer	Not mention	Individual	Self-directed learning	Achievement test scores	0.548*
Hsiao et al. (2016)	1	64	31	33	Earth science	Elementary school	Markerless-based	Tablet	> 1 week and ≤ 4 weeks	Individual	Problem-based learning	Achievement test scores	0.835*
Huang and Lin (2017)	1	104	52	52	Earth science	Elementary school	location-based	Tablet	> 1 week and ≤ 4 weeks	Individual	Conventional method	Learning achievement test scores	1.082*
Karagozlu (2018)	1	147	77	70	Related science themes	Middle school	Marker-based	Tablet	> 1 month and ≤ 4 months	Not specified	Not mention	Learning achievement test scores	1.755*
Kirikkaya and Başgül (2019)	2	120	30	30	Earth science	Middle school	Marker-based	Mobile	> 1 month and ≤ 4 months	Not specified	Conventional method	Achievement test scores	0.856*
			30	30	Earth science	Middle school	Marker-based	Mobile	> 1 month and ≤ 4 months	Not specified	Conventional method	Achievement test scores	0.766*
Lin et al. (2013)	1	40	20	20	Physics	College/university	Marker-based	Mobile	≤ 24 h	Two to four	Game-based learning	Learning achievement test scores	0.732*

Table 2 (continued)

Authors (year)	Num	N	EG	CG	Discipline	Educational stage	Types of AR	Device types of AR	Duration	Group size	Strategy	Measurement of academic achievement	Hedges's g
Lu et al. (2020)	1	53	25	28	Earth science	Elementary school	Marker-based	Tablet	≤ 24 h	Individual	Game-based learning	Learning achievement test scores	3.681*
Özerbaş (2019)	1	45	24	21	Physics	Middle school	Marker-based	Tablet /mobile	> 1 month and ≤ 4 months	Not specified	Conventional method	Knowledge test scores	1.423*
Sahin and Yilmaz (2020)	1	100	50	50	Earth science	Middle school	Marker-based	Computer	> 1 week and ≤ 4 weeks	Individual	Conventional method	Learning achievement test scores	1.104*
Singh et al. (2019)	1	60	30	30	Physics	College/ university	Marker-based	Computer	≤ 24 h	Individual	Self-directed learning	The lab skills test scores	1.326*
Sirakaya and Cakmak (2018)	1	118	62	56	Earth science	Middle school	Marker-based	Tablet	> 1 month and ≤ 4 months	Individual	Conventional method	Learning achievement test scores	0.957*
Sun and Chen (2019)	1	61	30	31	Mathematics	Elementary school	Marker-based	Tablet	> 1 month and ≤ 4 months	Two to four	Self-directed learning	Academic achievement test scores	0.986*
Tarng et al. (2015)	1	60	30	30	Biology	Elementary school	Location-based	Tablet	> 1 month and ≤ 4 months	Individual	Game-based learning	Learning achievement test scores	− 1.15*
Thees et al., 2020)	1	74	40	34	Physics	College/ university	Marker-based	Headset	≤ 24 h	Two to four	Self-directed learning	Conceptual knowledge test scores	− 0.332
Turan et al. (2018)	1	95	40	55	Earth science	College/ university	Marker-based	Mobile	> 1 week and ≤ 4 weeks	Not specified	Conventional method	Academic achievement test scores	3.603*
Weng et al. (2020)	1	68	34	34	Biology	Middle school	Marker-based	Mobile	≤ 24 h	Individual	Self-directed learning	Knowledge test scores	0.183
Yoon et al. (2017)	1	58	29	29	Physics	Middle school	Marker-based	Computer	≤ 24 h	Two to four	Self-directed learning	Knowledge test scores	0.404
Sum	39	2625	1303	1322									

Num num of comparisons, *N* total sample size, *EG* experimental group, *CG* control group

*$p < .05$

Table 3 Overall effect size and the homogeneity test

N	k	Hedges's g	95% CI	Q_T	df	I^2
2625	39	0.737*	[0.506, 0.969]	317.828*	38	88.044%

N total number of participants, *k* total number of effect sizes, *CI* confidence interval

*$p < .05$

learning while also improving their knowledge and academic achievement. This might reflect why the bulk of the studies (40.6%) were done in the field of "Science" (Pellas et al., 2019), and indicate AR's potential influence in teaching certain abstract or complicated science ideas (Furió et al., 2013). Another probable explanation is that the benefits of AR in the demonstration, such as letting students engage with virtual items in learning activities and situations (Chen & Wang, 2015) by displaying information in 3D format, contribute to enhancing students' understanding of abstract concepts and promoting inquiry-based learning in science education. In addition, the homogeneity test yielded a Q statistic of 317.828, which was higher than the degree of freedom (df). With a p-value of less than 0.001, the null hypothesis test is statistically significant. According to I^2 statistics, 88.044% of the observed total variation is not attributable to sampling mistakes within the same population. Consequently, moderator analyses were performed in this study to investigate the possible influence of moderator factors on students' academic success in AR-aided scientific learning.

Moderator Analysis

Table 5 demonstrates that the only moderator that was significantly associated with variability in student academic achievement was discipline. In contrast, the educational stages, types of AR, AR devices, intervention duration, group size, and strategies were irrelevant. The following sections provide and discuss the results for each moderator.

Discipline

The findings show that heterogeneity between disciplines has a significant influence on academic achievement in AR-supported science learning ($Q_B = 19.322$, $df = 4$, $p < .05$), and there is considerable variance in effect sizes across all disciplines. The majority of the research projects considered were for physics ($k = 12$), biology ($k = 12$), and Earth sciences ($k = 9$). The results show that AR works best in the Earth science disciplines ($g = 1.451$) and related science themes ($g = 1.146$), with a large effect size. Furthermore, AR has a medium-to-large effect size in mathematics ($g = 0.716$) and physics ($g = 0.670$), but a small and non-significant effect size in biology ($g = 0.186$). The findings revealed that AR had much better benefits than traditional instruction in a variety of science-related disciplines. It is also worth noting that the effect size in Earth sciences was 1.451, indicating a rather large effect. Five of the Earth sciences were concerned with astronomy, such as solar systems and space, while the remaining four were concerned with geomorphology and landscapes. Both astronomy and landscapes are creative and ambiguous concepts (Sahin & Yilmaz, 2020) that are challenging for students to visualize and require spatial skills (Lu et al., 2020). With the affordances of AR, these Earth sciences learning materials may be more effective in its visualizations and multiple presentations through students manipulating 3D items (e.g., observing the galaxy and/or landforms), hence enhancing conceptual understanding in learning the Earth sciences (Cheng & Tsai, 2013; Elford et al., 2022; Linn, 2003). As a result, it is likely that employing AR is more successful in difficult-to-observe learning topics such as the galaxy and landforms. In addition, there are significant medium-to-large effect sizes in physics ($k = 12$, $g = 0.670$) and mathematics ($k = 4$, $g = 0.716$). A partial explanation for this finding is that AR has been shown to be effective when applied to physics experiments (Abdusselam & Karal, 2020; Ibáñez et al., 2014), mathematics and geometry (Cai et al., 2020), and in general, various occurrences and abstract concepts that students were unable to observe in real life, such as magnetic fields (Ibáñez et al., 2014), the atomic model (Suprapto

Table 4 Trim-and-fill results for publication bias analysis

Model	Side	No. of missing studies	Observed values	Adjusted values	Q_T
Fixed effects model	Left	0	0.701*	0.701*	317.828
	Right	7		0.882*	478.374
Random effects model	Left	0	0.737*	0.737*	317.828
	Right	11		1.060*	617.800

*$p < .05$

Table 5 Effect sizes by moderator variables on students' academic achievement

Moderator variables	*N*	*k*	Hedges's *g*	95% CI	Variance	Q_B
Discipline		39			0.072	19.322*
1. Physics	723	12	0.670	[0.292, 1.049]	0.037	($p = .001$)
2. Biology	617	12	0.186^{N}	[−0.193, 0.566]	0.037	
3. Earth sciences	777	9	1.451	[1.015, 1.886]	0.049	
4. Mathematics	221	4	0.716	[0.047, 1.384]	0.116	
5. Related science themes	287	2	1.146	[0.261, 2.030]	0.204	
Educational stage		39			0.025	2.788
1. Preschool and elementary school	851	15	0.681	[0.297, 1.065]	0.038	($p = .425$)
2. Junior high school	1014	13	0.726	[0.321, 1.132]	0.043	
3. Senior high school	119	3	0.230^{N}	[−0.628, 1.088]	0.192	
4. College/university	641	8	1.045	[0.528, 1.563]	0.070	
Type of AR		39			0.072	5.758
1. Marker-based	1892	27	0.915	[0.641, 1.189]	0.020	($p = .056$)
2. Markerless-based	569	10	0.411^{N}	[−0.037, 0.860]	0.052	
3. Location-based	164	2	0.001^{N}	[−0.991, 0.993]	0.256	
Device of AR		39			0.038	4.482
1. Computer (laptop/desktop) with a camera	532	8	0.594	[0.081, 1.108]	0.069	($p = .345$)
2. Mobile	572	8	1.027	[0.512, 1.542]	0.069	
3. Tablet	1402	21	0.704	[0.384, 1.024]	0.027	
4. Headset	74	1	-0.332^{N}	[−1.767, 1.103]	0.536	
5. Mixed tablet and mobile	45	1	1.423^{N}	[−0.083, 2.930]	0.591	
Duration		39			0.015	0.390
1. ≤ 24 h	590	10	0.759	[0.277, 1.242]	0.061	($p = .942$)
2. > 1 week and ≤ 4 weeks	724	12	0.831	[0.394, 1.268]	0.050	
3. > 1 month and ≤ 4 months	978	13	0.683	[0.266, 1.101]	0.045	
4. Not mentioned	333	4	0.592^{N}	[−0.168, 1.351]	0.150	
Group size		39			0.036	5.039
1. Individual	1221	17	0.748	[0.405, 1.090]	0.030	($p = .169$)
2. Two to four	711	11	0.538	[0.112, 0.963]	0.047	
3. More than four	219	3	0.270^{N}	[−0.532, 1.072]	0.167	
4. Not specified	474	8	1.172	[0.670, 1.673]	0.065	
Strategy		39			0.034	5.291
1. Self-directed learning	855	14	0.538	[0.166, 0.911]	0.036	($p = .381$)
2. Game-based learning	291	6	0.593^{N}	[−0.001, 1.188]	0.092	
3. Conventional method	1097	15	0.881	[0.520, 1.241]	0.034	
4. Problem-based learning	124	2	1.161	[0.168, 2.154]	0.257	
5. Project-based learning	111	1	0.257^{N}	[−1.093, 1.606]	0.474	
6. Not specified	147	1	1.755	[0.403, 3.107]	0.476	

N total number of participants, *k* number of effect sizes, *Hedges's g* weighted mean effect size, in which.^{N} indicates that the effect size is nonsignificant at 95% confidence interval, *CI* confidence interval

*$p < .05$

et al., 2020), force, net force, friction (Enyedy et al., 2012), solid geometry (Lin et al., 2015; Rossano et al., 2020), and algebra (Saundarajan et al., 2020). These abstract topics requiring math and physics might be translated into simple principles or relationships, or perhaps observed phenomena, using AR assistance. In contrast, this study showed that AR has a small and nonsignificant impact size in biology ($k = 12$, $g = 0.186$), implying that using AR to teach students about biology has no impact on their academic achievement. We further inspect these studies, with nine of the research topics addressing the microscopic ecological system or the growth of plants and insects, while three are connected with human body systems. We speculate that biology-related topics may require systematic/methodical integration and actual physical experience or observation through microscopes and other equipment in order to feel its slight shifts more authentically. In contrast, AR adoption may be more focused on visualizing creatures and phenomena while ignoring systematic integration and learning material arrangement, resulting in students' misunderstanding in using AR animations or complex operations. This is partially consistent with the findings of Dehghani et al. (2020), who argued that while AR has the ability to offer various aspects of a phenomenon or function to improve learners' comprehension, it may also impose a cognitive load on students owing to insufficient AR representation or inadequate design. They also observed that utilizing AR on its own had little influence on learning results. Students' learning was enhanced dramatically when infographics were combined with augmented reality. Another point to consider is that 8 of the 12 articles are for students to study biology information outside of the classroom. Outdoor inquiry learning may help students pay greater attention to biological issues, in contrast to simply using AR. For example, Chien et al. (2019) discovered no significant difference in academic achievement between the groups using AR and herbarium specimens when learning in the introductory course entitled "plant stem." Furthermore, as for the unspecified science themes, we cannot conclude that AR is more effective, due to the limited amount of research.

Educational Stage

The homogeneity test reveals that there is no significant difference between the different weighted effect sizes of educational stages ($Q_B = 2.788$, $df = 3$, $p > .05$), and the value of Q_B is small. The results indicated that the effects of AR in science learning are not different across educational stages. Among those studies, most were conducted at preschool and elementary school ($k = 15$, $g = 0.681$) or junior high school ($k = 13$, $g = 0.726$) levels, with a significant medium-to-large effect size. In contrast, the result showed that the largest effect size was found at the college/university level ($k = 8$, $g = 1.045$). This finding is consistent with the findings of Ozdemir et al. (2018) and Garzón and Acevedo (2019), who found a greater impact size among undergraduates. One possibility is that when students are expected to multitask in AR environments (Wu et al., 2013), their understanding and operational abilities are tested. College students have a reasonably high degree of technology adaption and precise operation, and they may conduct experimental exercises in engineering, physics, and other subjects using AR or other interactive simulation equipment. In addition, five of the eight studies involve college/university students learning physics and engineering concepts, which necessitates a high level of inquiry skills. Compared to students of other ages, students in college/university may be more autonomous and capable of inquiring, allowing them to perceive AR-assisted scientific learning more efficiently. Furthermore, the effect size for senior high school students is small and non-significant ($k = 3$, $g = 0.230$). This might be due to increased academic pressure and a smaller sample size, a possibility that should be further investigated.

Types of AR

Homogeneity statistics Q_B with a value of 5.758 ($df = 2$, $p = .056$) show that there is a marginally significant difference between the effect size of marker-based, markerless-based, and location-based AR. To compare these three types further, marker-based ($k = 27$, $g = 0.915$) is the most frequently used and beneficial in improving academic achievement in scientific learning, which is consistent with the findings of Arici et al. (2019) which posit that marker-based type of AR were utilized more in science learning. It is possible that marker-based AR may have a lower technical threshold and be easier to apply to the display of scientific materials or with other learning strategies. For example, learning exercises for the majority of scientific topics employed marker-based AR, which may be accomplished by scanning QR codes on the paper and displaying the AR effect via the preconfigured app. Furthermore, the results show that markerless-based ($k = 10$, $g = 0.411$) has a small-to-medium effect size but is nonsignificant. We can see that eight of the markerless-based ARs are utilized in a biology discipline, which represents the relevant features of this type of AR; this smaller effect size may be due to the combined moderate effect of disciplines (see the discussion above) and AR types. Another possible reason is that we found that markerless-based AR is mostly used for outdoor scientific learning and provides more opportunities for learners to conduct scientific inquiries, such as investigating flora and insects or visiting exhibitions on human body systems. Students' scientific learning may be readily exposed to venue constraints or inclined toward informal learning methods, which result in instructional issues (e.g., roaming, ineffective teamwork) that affect the learning benefits of adopting markerless-based AR. The study also found that the location-based effect size ($k = 2$, $g = 0.001$) is small but not significant, with one possible explanation being that the number of studies is limited.

Display Devices of AR

As for the display devices, the homogeneity test reveals no significant difference ($Q_B = 4.482$, $df = 4$, $p > .05$). According to the findings, AR using mobile devices ($k = 8$, $g = 1.027$) had a large effect size, whereas tablets ($k = 21$, $g = 0.704$) and computers (laptop/desktop) combined with a video camera ($k = 8$, $g = 0.594$) had a medium-to-large effect on students' academic achievement in AR-assisted science learning. However, the effect size of the headset ($k = 1$, $g = -0.332$)

and mixed tablet and mobile devices ($k=1$, $g=1.423$) is not significant. The findings are similar to the meta-analysis findings of Ozdemir et al. (2018), who showed that mobile devices had the largest effect size, followed by tablets and desktop computers. One potential advantage of mobile devices is that presenting virtual objects on tablets and smartphones is handier and more suitable (Al-Mashaqbeh & Al Shurman, 2015) and may enhance students' academic achievement (Huang et al., 2014), as well as their involvement. Juan et al. (2014) made a similar argument, claiming that incorporating augmented reality with mobile devices might result in more flexibility and self-operation for students worldwide. When employing computers (laptop/desktop) in conjunction with a video camera, it is cumbersome for teachers to present AR or for students to operate AR, which may reduce learning efficiency and impact its effect. Furthermore, due to the small number of trials, the effect sizes of "headsets" and "mixed tablet and mobile devices" are not included in the moderator analysis.

Intervention Duration

When the "not mentioned" category and days 1 to 7 ($k=0$) are excluded, there is no significant difference in the effect size of intervention lengths ($Q_B=0.390$, $df=3$, $p=.942$), and the Q_B is quite small. The duration length of ">1 week and ≤ 4 weeks" ($k=12$, $g=0.831$) has a large effect size on students' academic achievement in AR-aided science learning, whereas "$=24$ h" ($k=10$, $g=0.759$) and ">1 month and ≤ 4 months" ($k=13$, $g=0.683$) both have medium-to-large effect sizes, implying that the effect size of medium intervention time tends to be larger, while the effect of long intervention time seems to be inferior to that of shorter time. This conclusion is consistent with the meta-analysis results of Garzón et al. (2020) and echoes the finding that technology-based instruction, such as computers and mobile phones, had a greater effect when the duration was shorter (Cheung & Slavin, 2013; Kulik & Kulik, 1991). Sung et al. (2016) argued that long-term trials did not always meet the needs of the teaching process, in particular learning topics, and did not always identify which teaching approaches to utilize in order to accomplish certain educational goals. In many short-term studies, researchers would select or use the best appropriate software and design more elaborate learning activities to minimize for confounding factors. However, according to the findings of this study, if the intervention is too brief (less than 24 h), the effect size is also undesirable. As a result, a medium intervention period (>1 week and ≤ 4 weeks) tends to have a larger effect size on students' academic achievement in AR-aided science learning.

Long-term educational interventions, on the other hand, are necessary for excluding "novelty effect" and achieving reliable results (Garzón & Acevedo, 2019). It is possible that the negative long-term consequences are due to the AR tools utilized in learning being unappealing to the participants or most likely misaligned with the learning objectives, or possibly that students are only engaged in AR for a short period of time and are hesitant to use it long term. Undoubtedly, it is important to perform AR with learning content and activities, and to guarantee that students go through the adoption and adaptation process before using it (Sung et al., 2016).

Group Size

The homogeneity test revealed no significant difference in the effect size of group sizes ($Q_B=5.039$, $df=3$, $p=.169$). Eight of the 39 studies do not describe the grouping status. Aside from the "not specified" grouping, "individual" (one person) is the most common group size ($k=17$), followed by groups of "two to four" ($k=11$), and "more than four" ($k=3$). As shown in Table 4, the effect size of "individual" ($g=0.748$) is larger than that of "two to four" ($g=0.538$), and the effect size of "more than four" is nonsignificant ($g=0.270$, $p>.05$). The results indicated that allowing individual students to operate and learn independently is the most popular approach when compared to small group operations, and it is worth noting that the greater the group size, the worse the effect.

Some researchers have argued that AR applications can stimulate debate, problem-solving, and communication, which can in turn promote cooperative learning (Ozdemir et al., 2018). It has had its greatest impact in education when the collaborative pedagogical approach was employed (Garzón et al., 2020). But according to the findings of this study, the effect size of group size "two to four" was less than that of "individual." AR-assisted scientific instruction frequently necessitated devices like computers, smartphones, and tablets, which can increase student engagement and social cohesion. However, increasing interaction may not directly increase students' academic achievement; in fact, excessive social communication may cause students to pay less attention to learning content (Sung et al., 2016). Although autonomous learning with a single device per student may be more beneficial to academic achievement, it is possible that students' grouping and collaboration are unsupervised and unmanaged, which may result in poor academic achievement. It is also worthwhile to continue researching the impact of AR on collaboration and communication in science learning.

Strategies

The homogeneity test results for the strategy used are not significant, indicating that there is no significant difference between the different instructional strategies used during the experiment ($Q_B=5.291$, $df=5$, $p=.381$). The conventional approach ($k=15$) and self-directed learning ($k=14$) were the most commonly utilized methods in scientific learning using AR, whereas game-based learning ($k = 6$), problem-based learning ($k = 2$), and project-based learning ($k = 1$) were used less frequently in AR-aided science learning. The effect size of the conventional method ($g = 0.881$) was greater than that of game-based learning ($g = 0.593$, $p = .05$) and self-directed learning ($g = 0.538$). Many researchers noted that while AR has the potential to improve science learning, it must be used in conjunction with appropriate learning design and instructional guidance (Wu et al., 2013). In this study, self-directed learning, which is student-centered and facilitated by the teacher, has a medium effect size on students' academic achievement. In comparison to the conventional method, self-directed learning may encourage students to develop critical thought and research abilities in addition to academic achievement (Gerard et al., 2011). Surprisingly, there has been little research on the usage of game mechanisms in AR-aided scientific learning. For instance, Chen (2020) developed an AR game to assist students in recognizing insects by merging AR technology with a digital game, observing that using the game method alone greatly enhanced students' learning achievements. Also, the findings of this study showed that the effect size of game-based learning is medium-to-high, implying that appropriate gaming strategies may improve students' scientific academic achievement. In terms of other strategies, the number of research projects that have used a problem- and project-based strategy is very limited; thus, it is unlikely to be typical for AR-supported science learning.

Conclusion

This meta-analysis was carried out to synthesize the effectiveness of the adoption of AR technology on student academic achievement in scientific learning. Our findings revealed a medium-to-large significant positive effect on students' academic achievement in science-related courses; additionally, discipline and AR types served as a significant moderator and a marginally significant moderator, implying that various domains in science education moderated the effect of using AR on students' science academic achievement, and notes that the types of AR may also have potential to affect student achievement in some cases. In particular, AR technologies are utilized in the Earth sciences or provided through marker-based AR, where they have the most positive effect on students' learning achievements among the "discipline" and "types of AR" moderators. AR technology has advanced in recent years and is now widely utilized in educational settings to increase student learning, particularly in science education. Although prior studies have shown the benefits of employing augmented reality to enhance student learning, our findings advance the study by providing additional and recent empirical data to support the claims, particularly in investigating the wide-ranging efforts of moderators (such as disciplines, educational stages, types of AR, display devices of AR, intervention duration, group size, and strategies used) that may influence the effectiveness of AR in science learning. According to the findings of this study, the application of augmented reality in scientific learning was verified to be beneficial for enhancing learning outcomes and could be used as a feasible alternative to traditional teaching. The most impactful benefits on scientific academic accomplishment in AR-assisted science learning are usually aligned with conceptual comprehension, spatial abilities, or science inquiry skills. Furthermore, marker-based AR was shown to be more relevant to the classroom and favorable to students' scientific academic achievement. Markerless AR was shown to be preferable for developing inquiry-based activities in which students interact with one another and with the actual environment (Cheng & Tsai, 2013).

These findings have practical implications for educational institutions, practitioners, and policymakers, urging them to focus more on promoting and supporting the use of AR in scientific education. In the future, we predict that as barriers to entry for augmented reality technology and/or its costs steadily drop, there will be an increase in the number of teaching methods incorporating augmented reality into science-related curricula. For instance, the advancement of augmented reality technology enables the presentation of experimental data graphically, as well as more clearly and frequently, along with the avoidance of superfluous experimental equipment expenditures. Thus, we encourage more science instructors to adopt AR technologies into their teaching, taking into account various special topic aspects and learning strategies utilized to maximize the effect of AR in order to enhance students' scientific learning through the proper application of various types of AR and its appropriate equipment. For example, encouraging science education with the aid of AR or to be carried out with scientific inquiry activities for abstract or conceptual learning content such as galaxies, ecosystems/biological structure, or geometry may have a greater effect in improving upon students' academic achievement levels. Furthermore, while using AR, we must analyze specific learning subject features or apply instructional approaches to advise or give scaffolding support to avoid misunderstandings or decrease the cognitive load produced by AR technology to the greatest possible degree. Depending on the type of AR, marker-based AR is more suited for use in classroom instruction, whereas markerless AR may allow students to engage in outdoor or inquiry learning. In addition, students should be encouraged to learn and explore science through self-directed or game-based learning and the use of mobile devices (smartphones or tablets) to exhibit AR. Alternately, science instructors can first explain the topic to be studied, then utilize AR to boost students' scientific accomplishment.

Regardless of the fact that this meta-analysis adds to current field of study on the effectiveness of AR in science learning contexts, it does have certain limitations. First, it should be noted that this work only includes journal articles (which are normally of great quality) and ignores degree theses or unpublished papers (of which quality varies), which might inflate the total effect size (but under control, as shown in the publication bias assessment), as studies with significant results are more likely to be published. Moreover, due to the severe criteria of meta-analysis on empirical data, other research involving non-experimental/quasi-experimental designs, studies with small sample sizes, and qualitative studies were excluded from this study; however, these studies may contain valuable information or experiences about the effectiveness of AR for us to learn from and should be considered in future studies. Also, this study examined students' academic achievement in science learning with and without the use of AR, rather than other affective factors such as motivation, attitude, and interests, which could be examined in future meta-analyses (if enough studies are conducted) to better understand the impact of AR applications on the science learning process. Furthermore, because the findings indicating the effect size of group learning was not significant and surpassed our expectations, further research into the influence of AR on communication and collaboration, as well as cooperation strategy assistance, is required. Second, several moderator variables that may influence the efficacy of AR-supported scientific education—such as gender, learning styles, and students' acceptance of technology—were not explored in this study due to the limited number of studies that report enough statistics; these variables may be considered in future studies to perform a comprehensive literature search for meta-analysis and to further examine the practical use of AR in scientific education.

Acknowledgements The authors would like to thank Mr. En Hu for his support with data analysis.

Author Contribution Wen-Wen Xu: conceptualization, methodology, investigation, data analysis, draft writing. Chien-Yuan Su: conceptualization, writing, editing and reviewing, supervision. Yue Hu: methodology, validation, supervision. Cheng-Huan Chen: methodology, reviewing.

Funding This research was supported by the Startup Research Grant Scheme of Hangzhou Normal University (No. RWSK20201015) and the Provincial Advantageous and Characteristic Discipline Cultivation Project of Hangzhou Normal University (No. 20JYXK030).

Declarations

Informed Consent Not applicable.

Conflict of Interest The authors declare no competing interests.

Research Involving Human Participants and/or Animals Not applicable.

References

References marked with an asterisk indicate studies included in the meta-analysis.

*Abdusselam, M. S., & Karal, H. (2020). The effect of using augmented reality and sensing technology to teach magnetism in high school physics. *Technology, Pedagogy and Education, 29*(4), 407–424. https://doi.org/10.1080/1475939X.2020.1766550

Akçayır, M., & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review, 20*, 1–11. https://doi.org/10.1016/j.edurev.2016.11.002

*Akçayır, M., Akçayır, G., Pektaş, H. M., & Ocak, M. A. (2016). Augmented reality in science laboratories: The effects of augmented reality on university students' laboratory skills and attitudes toward science laboratories. *Computers in Human Behavior, 57*, 334–342. https://doi.org/10.1016/j.chb.2015.12.054

Al-Mashaqbeh, I., & Al Shurman, M. (2015). The adoption of tablet and e-textbooks: First grade core curriculum and school administration attitude. *Journal of Education and Practice, 6*(21), 188–194. https://doi.org/10.1016/j.compedu.2019.103647

Arici, F., Yildirim, P., Caliklar, Ş., & Yilmaz, R. M. (2019). Research trends in the use of augmented reality in science education: Content and bibliometric mapping analysis. *Computers & Education*. https://doi.org/10.1016/j.compedu.2019.103647

Borenstein, M., Hedges, L. V., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. John Wiley & Sons.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2013). *Comprehensive meta-analysis version 3* [Computer software]. Biostat Inc.

Bowden, J., Davey Smith, G., & Burgess, S. (2015). Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *International Journal of Epidemiology, 44*(2), 512–525. https://doi.org/10.1093/ije/dyv080

Cai, S., Chiang, F.-K., Sun, Y., Lin, C., & Lee, J. J. (2017). Applications of augmented reality-based natural interactive learning in magnetic field instruction. *Interactive Learning Environments, 25*(6), 778–791. https://doi.org/10.1080/10494820.2016.1181094

*Cai, S., Chiang, F.-K., & Wang, X. (2013). Using the augmented reality 3D technique for a convex imaging experiment in a physics course. *International Journal of Engineering Education, 29*(4), 856–865.

*Cai, S., Liu, E., Shen, Y., Liu, C., Li, S., & Shen, Y. (2020). Probability learning in mathematics using augmented reality: Impact on student's learning gains and attitudes. *Interactive Learning Environments, 28*(5), 560–573. https://doi.org/10.1080/10494820.2019.1696839

Cai, S., Wang, X., & Chiang, F.-K. (2014). A case study of augmented reality simulation system application in a chemistry course. *Computers in Human Behavior, 37*, 31–40. https://doi.org/10.1016/j.chb.2014.04.018

*Carrera, C. C., & Asensio, L. A. B. (2017). Landscape interpretation with augmented reality and maps to improve spatial orientation skill. *Journal of Geography in Higher Education, 41*(1), 119–133. https://doi.org/10.1080/03098265.2016.1260530

Chang, H.-Y., Liang, J.-C., & Tsai, C.-C. (2020). Students' context-specific epistemic justifications, prior knowledge, engagement, and socioscientific reasoning in a mobile augmented reality learning environment. *Journal of Science Education and Technology, 29*(3), 399–408. https://doi.org/10.1007/s10956-020-09825-9

*Chang, R.-C., Chung, L.-Y., & Huang, Y.-M. (2016). Developing an interactive augmented reality system as a complement to plant education and comparing its effectiveness with video learning. *Interactive Learning Environments, 24*(6), 1245–1264. https://doi.org/10.1080/10494820.2014.982131

*Chang, R.-C., & Yu, Z.-S. (2018). Using augmented reality technologies to enhance students' engagement and achievement in science laboratories. *International Journal of Distance Education Technologies, 16*(4), 54–72. https://doi.org/10.4018/IJDET.2018100104

*Chang, S.-C., & Hwang, G.-J. (2018). Impacts of an augmented reality-based flipped learning guiding approach on students' scientific project performance and perceptions. *Computers & Education, 125*, 226–239. https://doi.org/10.1016/j.compedu.2018.06.007

*Chen, C.-H. (2020). Impacts of augmented reality and a digital game on students' science learning with reflection prompts in multimedia learning. *Educational Technology Research and Development, 68*(6), 3057–3076. https://doi.org/10.1007/s11423-020-09834-w

Chen, C.-H., & Yang, Y.-C. (2019). Revisiting the effects of project-based learning on students' academic achievement: A meta-analysis investigating moderators. *Educational Research Review, 26*, 71–81. https://doi.org/10.1016/j.edurev.2018.11.001

Chen, C., & Wang, C.-H. (2015). Employing augmented-reality-embedded instruction to disperse the imparities of individual differences in Earth science learning. *Journal of Science Education and Technology, 24*(6), 835–847. https://doi.org/10.1007/s10956-015-9567-3

Cheng, K.-H., & Tsai, C.-C. (2013). Affordances of augmented reality in science learning: Suggestions for future research. *Journal of Science Education and Technology, 22*(4), 449–462. https://doi.org/10.1007/s10956-012-9405-9

Cheung, A. C. K., & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review, 9*, 88–113. https://doi.org/10.1016/j.edurev.2013.01.001

*Chiang, T. H. C., Yang, S. J. H., & Hwang, G.-J. (2014). An augmented reality-based mobile learning system to improve students' learning achievements and motivations in natural science inquiry activities. *Educational Technology and Society*, *17*(4), 352–365. https://www.jstor.org/stable/pdf/jeductechsoci.17.4.352.pdf

*Chien, Y.-C., Su, Y.-N., Wu, T.-T., & Huang, Y.-M. (2019). Enhancing students' botanical learning by using augmented reality. *Universal Access in the Information Society, 18*(2), 231–241. https://doi.org/10.1007/s10209-017-0590-4

Cohen, J. D. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates.

Dange, J. K. (2018). Mobile-assisted-learning approach in enhancing the student teacher's vocabulary and usage of mobile phone. In J. Keengwe (Eds.), *Handbook of research on mobile technology, constructivism, and meaningful learning* (pp. 316–330). IGI Global. https://doi.org/10.4018/978-1-5225-3949-0.ch017

*Dehghani, M., Mohammadhasani, N., Ghalevandi, M. H., & Azimi, E. (2020). Applying AR-based infographics to enhance learning of the heart and cardiac cycle in biology class. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2020.1765394

*Demitriadou, E., Stavroulia, K. E., & Lanitis, A. (2020). Comparative evaluation of virtual and augmented reality for teaching mathematics in primary education. *Education and Information Technologies, 25*(1), 381–401. https://doi.org/10.1007/s10639-019-09973-5

Diegmann, P., Schmidt-Kraepelin, M., Van den Eynden, S., & Basten, D. (2015). Benefits of augmented reality in educational environments – A systematic literature review. *Proceedings of the 12th International Conference on Wirtschaftsinformatik*, article 103 (pp. 1542–1556).

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

Elford, D., Lancaster, S. J., & Jones, G. A. (2022). Exploring the effect of augmented reality on cognitive load, attitude, spatial ability, and stereochemical perception. *Journal of Science Education and Technology, 31*(3), 322–339. https://doi.org/10.1007/s10956-022-09957-0

Enyedy, N., Danish, J. A., Delacruz, G., & Kumar, M. (2012). Learning physics through play in an augmented reality environment. *International Journal of Computer Supported Collaborative Learning, 7*, 347–378. https://doi.org/10.1007/s11412-012-9150-3

*Erbas, C., & Demirer, V. (2019). The effects of augmented reality on students' academic achievement and motivation in a biology course. *Journal of Computer Assisted Learning, 35*(3), 450–458. https://doi.org/10.1111/jcal.12350

*Faridi, H., Tuli, N., Mantri, A., Singh, G., & Gargrish, S. (2021). A framework utilizing augmented reality to improve critical thinking ability and learning gain of the students in Physics. *Computer Applications in Engineering Education, 29*(1), 258–273. https://doi.org/10.1002/cae.22342

*Fidan, M., & Tuncel, M. (2019). Integrating augmented reality into problem based learning: The effects on learning achievement and attitude in physics education. *Computers & Education*. https://doi.org/10.1016/j.compedu.2019.103635

*Fokides, E., & Mastrokoukou, A. (2018). Results from a study for teaching human body systems to primary school students using tablets. *Contemporary Educational Technology*, *9*(2), 154–170. https://doi.org/10.30935/cet.414808

Furió, D., González-Gancedo, S., Juan, M.-C., Seguí, I., & Rando, N. (2013). Evaluation of learning outcomes using an educational iPhone game vs. traditional game. *Computers & Education, 64*, 1–23. https://doi.org/10.1016/j.compedu.2012.12.001

Garzón, J., & Acevedo, J. (2019). Meta-analysis of the impact of augmented reality on students' learning gains. *Educational Research Review*, *27*, 244–260. https://doi.org/10.1016/j.edurev.2019.04.001

Garzón, J., Kinshuk, Baldiris, S., Gutiérrez, J., & Pavón, J. (2020). How do pedagogical approaches affect the impact of augmented reality on education? A meta-analysis and research synthesis. *Educational Research Review*, *31*, 100334. https://doi.org/10.1016/j.edurev.2020.100334

*Gecu-Parmaksiz, Z., & Delialioglu, O. (2019). Augmented reality-based virtual manipulatives versus physical manipulatives for teaching geometric shapes to preschool children. *British Journal of Educational Technology, 50*(6), 3376–3390. https://doi.org/10.1111/bjet.12740

Gerard, L. F., Varma, K., Corliss, S. B., & Linn, M. C. (2011). Professional development for technology-enhanced inquiry science. *Review of Educational Research, 81*(3), 408–448. https://doi.org/10.3102/0034654311415121

Gnidovec, T., Žemlja, M., Dolenec, A., & Torkar, G. (2020). Using augmented reality and the structure–behavior–function model to teach lower secondary school students about the human circulatory system. *Journal of Science Education and Technology, 29*(6), 774–784. https://doi.org/10.1007/s10956-020-09850-8

Goff, E. E., Mulvey, K. L., Irvin, M. J., & Hartstone-Rose, A. (2018). Applications of augmented reality in informal science learning sites: A review. *Journal of Science Education and Technology, 27*(5), 433–447. https://doi.org/10.1007/s10956-018-9734-4

*Gopalan, V., Zulkifli, A. N., & Bakar, J. A. A. (2016). Conventional approach vs augmented reality textbook on learning performance:

A study in science learning among secondary school students. *Revista de La Facultad de Ingeniería*, *31*(5), 19–26. https://doi.org/10.21311/002.31.5.03

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ, 327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

*Hsiao, H.-S., Chang, C.-S., Lin, C.-Y., & Wang, Y.-Z. (2016). Weather observers: A manipulative augmented reality system for weather simulations at home, in the classroom, and at a museum. *Interactive Learning Environments, 24*(1), 205–223. https://doi.org/10.1080/10494820.2013.834829

Hu, Y., Chen, C.-H., & Su, C.-Y. (2021). Exploring the effectiveness and moderators of block-based visual programming on student learning: A meta-analysis. *Journal of Educational Computing Research, 58*(8), 1467–1493. https://doi.org/10.1177/0735633120945935

Huang, K.-L., Chen, K.-H., & Ho, C.-H. (2014). Enhancing learning outcomes through new e-textbooks: A desirable combination of presentation methods and concept maps. *Australasian Journal of Educational Technology*, *30*(5), 600–618. https://doi.org/10.14742/ajet.538

*Huang, Y.-M., & Lin, P.-H. (2017). Evaluating students' learning achievement and flow experience with tablet PCs based on AR and tangible technology in u-learning. *Library Hi Tech, 35*(4), 602–614. https://doi.org/10.1108/LHT-01-2017-0023

Ibáñez, M. B., Di Serio, Á., Villarán, D., & Delgado Kloos, C. (2014). Experimenting with electromagnetism using augmented reality: Impact on flow student experience and educational effectiveness. *Computers & Education, 71*, 1–13. https://doi.org/10.1016/j.compedu.2013.09.004

Juan, C. M., Toffetti, G., Abad, F., & Cano, J. (2010). Tangible cubes used as the user interface in an augmented reality game for edutainment. *2010 10th IEEE International Conference on Advanced Learning Technologies*, pp 599–603. https://doi.org/10.1109/ICALT.2010.170

Juan, M.-C., Mendez-Lopez, M., Perez-Hernandez, E., & Albiol-Perez, S. (2014). *Augmented Reality for the Assessment of Children's Spatial Memory in Real Settings.* https://doi.org/10.1371/journal.pone.0113751

*Karagozlu, D. (2018). Determination of the impact of augmented reality application on the success and problem-solving skills of students. *Quality and Quantity, 52*(5), 2393–2402. https://doi.org/10.1007/s11135-017-0674-5

*Kirikkaya, E. B., & Başgül, M. Ş. (2019). The effect of the use of augmented reality applications on the academic success and motivation of 7th grade students. *Journal of Baltic Science Education*, *18*(3), 362–378. https://doi.org/10.33225/jbse/19.18.362

Kulik, C. L. C., & Kulik, J. A. (1991). Effectiveness of computer-based instruction: An updated analysis. *Computers in Human Behavior, 7*(1), 75–94. https://doi.org/10.1016/0747-5632(91)90030-5

Lin, H.-C.K., Chen, M.-C., & Chang, C.-K. (2015). Assessing the effectiveness of learning solid geometry by using an augmented reality-assisted learning system. *Interactive Learning Environments, 23*(6), 799–810. https://doi.org/10.1080/10494820.2013.817435

*Lin, T.-J., Duh, H.-B.L., Li, N., Wang, H.-Y., & Tsai, C.-C. (2013). An investigation of learners' collaborative knowledge construction performances and behavior patterns in an augmented reality simulation system. *Computers & Education, 68*, 314–321. https://doi.org/10.1016/j.compedu.2013.05.011

Linn, M. (2003). Technology and science education: Starting points, research programs, and trends. *International Journal of Science Education, 25*(6), 727–758. https://doi.org/10.1080/09500690305017

Lipsey, M. W., & Wilson, D. B. (2001). *Applied social research methods series*. Sage Publications Inc.

*Lu, S.-J., Liu, Y.-C., Chen, P.-J., & Hsieh, M.-R. (2020). Evaluation of AR embedded physical puzzle game on students' learning achievement and motivation on elementary natural science. *Interactive Learning Environments, 28*(4), 451–463. https://doi.org/10.1080/10494820.2018.1541908

Ozdemir, M., Sahin, C., Arcagok, S., & Demir, M. K. (2018). The effect of augmented reality applications in the learning process: A meta-analysis Study. *Eurasian Journal of Educational Research*, *18*(74), 165–186. https://doi.org/10.14689/ejer.2018.74.9

*Özerbaş, D. S. (2019). The effect of marker-based augmented reality (MBAR) applications on academic achievement and permanence. *Universal Journal of Educational Research*, *7*(9), 1926–1932. https://doi.org/10.13189/ujer.2019.070911

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., & Moher, D. (2021). The PRISMA 2020 Statement: An updated guideline for reporting systematic reviews. *BMJ*. https://doi.org/10.1136/bmj.n71

Pellas, N., Fotaris, P., Kazanidis, I., & Wells, D. (2019). Augmenting the learning experience in primary and secondary school education: A systematic review of recent trends in augmented reality game-based learning. *Virtual Reality, 23*(4), 329–346. https://doi.org/10.1007/s10055-018-0347-2

Radu, I. (2014). Augmented reality in education: A meta-review and cross-media analysis. *Personal and Ubiquitous Computing, 18*(6), 1533–1543. https://doi.org/10.1007/s00779-013-0747-y

Rossano, V., Lanzilotti, R., Cazzolla, A., & Roselli, T. (2020). Augmented reality to support geometry learning. *IEEE Access, 8*, 107772–107780. https://doi.org/10.1109/ACCESS.2020.3000990

*Sahin, D., & Yilmaz, R. M. (2020). The effect of augmented reality technology on middle school students' achievements and attitudes towards science education. *Computers & Education*. https://doi.org/10.1016/j.compedu.2019.103710

Santos, M. E. C., Chen, A., Taketomi, T., Yamamoto, G., Miyazaki, J., & Kato, H. (2014). Augmented reality learning experiences: Survey of prototype design and evaluation. *IEEE Transactions on Learning Technologies, 7*(1), 38–56. https://doi.org/10.1109/TLT.2013.37

Saundarajan, K., Osman, S., Kumar, J., Daud, M., Abu, M., & Pairan, M. (2020). Learning algebra using augmented reality: A preliminary investigation on the application of photomath for lower secondary education. *International Journal of Emerging Technologies in Learning, 15*(16), 123–133.

Silén, C., & Uhlin, L. (2008). Self-directed learning – A learning issue for students and faculty! *Teaching in Higher Education, 13*(4), 461–475. https://doi.org/10.1080/13562510802169756

*Singh, G., Mantri, A., Sharma, O., Dutta, R., & Kaur, R. (2019). Evaluating the impact of the augmented reality learning environment on electronics laboratory skills of engineering students. *Computer Applications in Engineering Education, 27*(6), 1361–1375. https://doi.org/10.1002/cae.22156

*Sirakaya, M., & Cakmak, E. K. (2018). The effect of augmented reality use on achievement, misconception and course engagement. *Contemporary Educational Technology*, *9*(3), 297–314. https://doi.org/10.30935/cet.444119

*Sun, K. T., & Chen, M. H. (2019). Utilizing free augmented reality app for learning geometry at elementary school in Taiwan: Take volumetric measurement of compound body for example. *International Journal of Distance Education Technologies, 17*(4), 36–53. https://doi.org/10.4018/IJDET.2019100103

Sung, Y.-T., Chang, K.-E., & Liu, T.-C. (2016). The effects of integrating mobile devices with teaching and learning on students' learning performance: A meta-analysis and research synthesis. *Computers & Education, 94*, 252–275. https://doi.org/10.1016/j.compedu.2015.11.008

Suprapto, N., Nandyansah, W. & Mubarok, H. (2020). An evaluation of the "picsar" research project: An augmented reality in physics

learning. *International Journal of Emerging Technologies in Learning, 15*(10), 113–125. https://www.learntechlib.org/p/217045/
*Tarng, W., Ou, K.-L., Yu, C.-S., Liou, F.-L., & Liou, H.-H. (2015). Development of a virtual butterfly ecological system based on augmented reality and mobile learning technologies. *Virtual Reality, 19*(3–4), 253–266. https://doi.org/10.1007/s10055-015-0265-5
*Thees, M., Kapp, S., Strzys, M. P., Beil, F., Lukowicz, P., & Kuhn, J. (2020). Effects of augmented reality on learning and cognitive load in university physics laboratory courses. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2020.106316
*Turan, Z., Meral, E., & Sahin, I. F. (2018). The impact of mobile augmented reality in geography education: Achievements, cognitive loads and views of university students. *Journal of Geography in Higher Education, 42*(3), 427–441. https://doi.org/10.1080/03098265.2018.1455174
Walczak, K., Wojciechowski, R., & Cellary, W. (2006). Dynamic interactive VR network services for education. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, 277–286.
*Weng, C., Otanga, S., Christianto, S. M., & Chu, R. J. C. (2020). Enhancing students' biology learning by using augmented reality as a learning supplement. *Journal of Educational Computing Research, 58*(4), 747–770. https://doi.org/10.1177/0735633119884213
Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review, 30*(1), 1–35. https://doi.org/10.1016/j.dr.2009.12.001
Wojciechowski, R., & Cellary, W. (2013). Evaluation of learners' attitude toward learning in ARIES augmented reality environments. *Computers & Education, 68*, 570–585. https://doi.org/10.1016/j.compedu.2013.02.014
Wu, H.-K., Lee, S.W.-Y., Chang, H.-Y., & Liang, J.-C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education, 62*, 41–49. https://doi.org/10.1016/j.compedu.2012.10.024
*Yoon, S., Anderson, E., Lin, J., & Elinich, K. (2017). How augmented reality enables conceptual understanding of challenging science content. *Educational Technology and Society, 20*(1), 156–168.

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