Original article

doi: 10.1111/jcal.12184

Journal of Computer Assisted Learning

Effects of mobile devices on K–12 students' achievement: a meta-analysis

S. Tingir,* B. Cavlazoglu,† O. Caliskan,‡ O. Koklu§ & S. Intepe-Tingir¶

*Department of Educational Psychology and Learning Systems, Florida State University, FL, USA
†Department of Science Education, Karadeniz Technical University, KTU, Turkey
‡Department of Educational Administration and Planning, Middle East Technical University, Turkey
§Department of Mathematics and Science Education, University of Georgia, Georgia
¶Department of School of Teacher Education, Florida State University, FL, USA

Abstract

In this meta-analytic study, we investigated the effects of mobile devices on student achievement in science, mathematics and reading in grades K–12. Based on our inclusion criteria, we searched the ERIC and PsycINFO databases and identified 14 peer-reviewed research articles published between 2010 and 2014. We identified the device type, subject area, intervention language, grade level, study design and implementer (i.e., of the intervention) as potential moderator variables that may influence student achievement in the targeted content areas. We followed a three-level meta-analytic procedure to estimate the overall effect of these variables and explain the variation in outcomes. The results suggest that use of mobile devices in teaching yielded higher achievement scores than traditional teaching in all subject areas. With regard to the analysis of moderator variables, the results suggest that using mobile devices in reading is significantly more effective than doing so in mathematics.

Keywords education technology, meta-analysis, mobile devices, mobile learning, student achievement.

Introduction

There is a broad consensus among stakeholders in K–12 education that students need opportunities to develop essential skills that promote creativity, critical thinking, productivity and collaboration in order to become technologically literate individuals in this digital age (NETS, 2007). Accordingly, appropriate technology integration in teaching and learning across various content areas (e.g., science, mathematics and reading) is important to enhance students' technological literacy. This integration is also essential for students to gain 21st-century skills that have been identified as necessary for success in 21st-century work and life (Partnership for 21st Century Skills, 2011). Therefore, educators have used various technological devices, including desktop computers, laptops and mobile devices. Among these devices, students in the current generation use mobile devices more often than any other technological devices (Sahin, Cavlazoglu, & Zeytuncu, 2015) owing to their advantageous features such as portability, touch sensitivity, wireless network capability, multifunctional camera ability and longer battery life.

A mobile device, also called a handheld device or handheld computer, is a small computing device that usually comes with a touch screen, wireless network capability and sometimes a mini keyboard (Viswanathan, 2015). Research on the use of mobile devices in teaching and learning in grades K–12 has been conducted in reference to such products as iPads (Riconscente, 2013), tablets (Huang, Chen, & Ho, 2014), smartphones (de-Marcos et al., 2010; Yang, Hwang, Hung, & Tseng, 2013b) and personal digital assistants (PDAs; Huang, Lin, & Cheng, 2010; Liu, 2009; Varma, 2014). A variety

Accepted: 15 January 2017
Correspondence: Seyfullah Tingir, Department of Educational Psychology and Learning Systems, Florida State University, 1114 W Call Street, Tallahassee, FL 32306, USA. Email: seyfullah.tingir@gmail.com

© 2017 John Wiley & Sons Ltd

of studies with differing focuses on the use of mobile devices also exist in the relevant literature. While some of these studies investigated student attitudes towards mobile learning and the implementation of new learning programmes with these mobile devices, most of them focused on the effects of these devices on student learning and achievement in various content areas. Furthermore, several systematic reviews on mobile learning documented that studies on mobile learning have mainly concentrated on the effects of the mobile devices on learning and achievement of K–12 students (Hung & Zhang, 2012; Hwang & Tsai, 2011; Wu et al., 2012). Moreover, the previous reviews (e.g., Hwang & Tsai, 2011; Wu et al., 2012) reviewed research studies that were published before 2010. Therefore, an up-to-date meta-analysis investigating effects of mobile devices on student achievement is critical to inform researchers and guide future studies in this research field. In addition, exploring current mobile devices' up-to-date effects on student achievement in the major content areas of K–12 education is essential in order to create better learning environments for students. Hence, the purpose of this study is to investigate the effects of mobile devices on student achievement in science, mathematics and reading content areas in grades K–12 through a meta-analysis.

Literature review

The rapid growth of mobile technology has influenced traditional practices in a number of areas such as communication, education and digital games. The use of mobile devices for educational purposes has become a new research area because mobile technology offers new opportunities for student learning, engagement and motivation (Lin, Fulford, Ho, Iyoda, & Ackerman, 2012). In mobile device research, the more frequently studied populations have been students in higher education as compared to K–12 students (Wu et al., 2012). Among the major academic disciplines and professions, applied sciences (29%), humanities (20%), formal sciences (16%), social sciences (4%) and natural sciences (3%) were the most common subject areas targeted in mobile device studies (Wu et al., 2012). In these studies, researchers found positive effects of mobile devices on student achievement in various subject areas (Baya'a & Daher, 2009; Riconscente, 2013; Varma, 2014; Wu et al., 2012), although there are also a few studies revealing neutral (Carr, 2012; Ketamo, 2003) and even negative effects of mobile devices (Doolittle & Mariano, 2008). We address the following three components: (a) mobile devices in science achievement, (b) mobile devices in mathematics achievement and (c) mobile devices in reading achievement.

Mobile devices in science achievement

Studies on K-12 grade students investigating the effects of mobile devices on student achievement in science have focused on using different mobile learning approaches (Looi et al., 2011; Varma, 2014), implementing different programmes to scaffold the learning process (Chu, Hwang, Tsai, & Tseng, 2010; Hung, Hwang, Lin, Wu, & Su, 2013) and comparing science achievement in student groups using either mobile devices or computers (Nedungadi & Raman, 2012). When compared to other major content areas in K–12 education such as mathematics and reading, the integration of mobile devices in teaching and learning science has been more frequent, especially in informal science settings (Scanlon, Jones, & Waycott, 2005). Many experimental studies in science education show that students using mobile devices in the treatment group outperformed students in the control group (Hwang, Wu, Zhuang, & Huang, 2013; Yang, Hwang, et al., 2013b).

Mobile devices in mathematics achievement

The use of mobile devices in mathematics classrooms has been rapidly increasing (Carr, 2012; Ross, Morrison, & Lowther, 2010). Mathematics teachers and students have been using a variety of mobile devices such as smartphones, iPads, tablets, iPods and PDAs for teaching and learning mathematics (Agostinho et al., 2015; Nedungadi & Raman, 2012). The use of mobile devices in the mathematics classroom has had positive impacts on student achievement, motivation, attitudes and cognitive skills (Carr, 2012; Li & Pow, 2011). For example, Riconscente (2013) investigated the effects of mathematical games using mobile devices in learning fraction concepts in elementary school and found that student achievement significantly improved. However, intervention studies investigating the effects of mobile devices on student achievement in mathematics are limited (Carr, 2012).

© 2017 John Wiley & Sons Ltd

Mobile devices in reading achievement

Reading is a critical skill for student academic achievement (Perfetti, Landi, & Oakhill, 2005). However, the extant research notes that reading difficulties for school age children are the most common challenge (Rafdal, Mcmaster, Mcconnell, Fuchs, & Fuchs, 2011). Researchers investigated whether students' reading achievements increase when appropriate interventions are provided (Torgesen, 2002; Vaughn et al., 2012; Wanzek & Vaughn, 2008). More recently, the implementation of mobile devices has become more common among reading intervention studies. Research shows that mobile devices are effective in improving student performance through increasing their motivation and engagement in reading (Chai, Vail, & Ayres, 2015; Crowley, McLaughlin, & Kahn, 2013). For example, an experimental study by Yang, Tseng, Liao, and Liang (2013a) showed that the group of students using mobile devices to read and learn poems outperformed the control group using a traditional textbook. Furthermore, many research studies focusing on students with special needs documented that the use of mobile devices has positive impacts on student learning (McClanahan, Williams, Kennedy, & Tate, 2012). Overall, research suggests that mobile devices are useful in improving student achievement in reading.

Incorporating technological devices into learning offers a more enriched learning environment compared to traditional classrooms. Bringing these devices into classrooms has become more common. However, designing studies (i.e., experimental and quasi-experimental) using mobile devices to explore the effectiveness of these devices may be challenging because of the difficulties in the various implementation methods of the mobile devices. A meta-analysis has the potential of avoiding the limitations of a usual literature review (Caird, Willness, Steel, & Scialfa, 2008), and it provides insights about how the mobile devices should be integrated into teaching and learning environments. A meta-analysis study may guide researchers to implement a study through a statistical combination of results from many studies as well as by examining the effects of moderator variables (Rosenthal & DiMatteo, 2001). More importantly, a meta-analysis study warrants various research dimensions to be included in the statistical analysis. Conducting a meta-analysis is useful in offering researchers more insights on how to integrate mobile devices into teaching and learning. Thus, the purpose of this paper is to meta-analytically estimate the true effect of mobile devices on student achievement. This meta-analysis seeks to answer the following questions:

1. What is the overall impact of the use of mobile devices on achievement in science, mathematics and reading among K–12 students?
2. How does the relationship differ across study characteristics (i.e., device type, subject area and intervention language), student characteristics (i.e., grade level) and/or methodology characteristics (i.e., research design and implementer)?

Method

We first describe the inclusion criteria used in this meta-analysis study. Then, we give the details of the coding procedures, followed by a discussion of the effect size calculation and the publication bias. Next, we explain the statistical analysis performed in the study. We complete the method section with an explanation of the homogeneity test.

Inclusion criteria

In this meta-analysis, we used the following inclusion criteria to determine the eligibility of studies. Studies were considered eligible if (a) mobile device interventions were implemented in grades K–12, (b) interventions were provided as part of the regular school curriculum, (c) study designs were either experimental or quasi-experimental, (d) enough data were reported to be able to calculate effect sizes, (e) they were published in peer-reviewed journals to establish the validity and reliability of the manuscripts.

We used a computer search in the ERIC, ProQuest and PsycINFO databases to find targeted studies about the effects of mobile devices on student achievement in science, mathematics and reading content areas, which correspond to the major content areas of K–12 and the key subjects as indicated by The Programme for International Student Assessment. Key descriptors or root forms of mobile device terms (i.e., iPad, tablet, hand-held device*, hand-held computer*, handheld, mobile device, smartphone, mobile learning, assisted learning and PDA) were used in combination with key performance indicators (i.e., success*, achievement, comprehend*,

© 2017 John Wiley & Sons Ltd

interven*, reasoning and perform*) to search for related articles. The preliminary search returned many medical articles because of the term ‘tablet’. Therefore, we decided to exclude some medical terms (i.e., clinic*, health and medic*) to reduce the number of search results. A total of 1044 studies were obtained from the search procedure. The search results were extracted to an online document-sharing tool and became available to the research team.

Afterwards, all five researchers of this study individually reviewed the titles, abstracts and keywords of the first 50 studies and were assigned to make a decision of ‘yes’, ‘maybe’ or ‘no’ based on the inclusion/exclusion criteria for training purposes. Disagreements between the researchers were resolved through an online group meeting with full team participation. When titles, abstracts and keywords of articles provided insufficient information, the researchers read the full-paper version of the articles and reviewed them to ensure the articles met the criteria. We followed the same procedure for the second set of 50 studies. After reaching an adequate level of consistency among researchers, we assigned 100 articles to each set of two researchers to make the inclusion/exclusion decisions. Any disagreement between researchers was resolved after meetings for each set of studies. This procedure was repeated until all of the 1044 studies were reviewed. As a result, these studies were narrowed down to the 14 studies that met our inclusion criteria (Table 1), published between 2010 and 2014. The database search resulted in no studies published before 2010 so that we set the starting date as 2010. In addition, after reviewing the literature on the effects of mobile devices on student achievement, we identified moderator variables from the 14 selected studies as follows: (a) device type, (b) subject area, (c) intervention language, (d) grade level, (e) study design and (f) implementer.

Coding procedures

We created our coding sheet based on the elements specified in the What Works Clearinghouse Design and Implementation Assessment Device (Institute of Education Sciences, 2003). We coded four categories: (a) study information, (b) participant information, (c) methodology and (d) findings to calculate effect size. Study information was coded in four categories including instruction language (i.e., English, Taiwanese or Chinese), school type (i.e., elementary, middle or high school), device type (i.e., tablet, PDA, smartphones or mobile devices) and subject (i.e., science, mathematics or reading). Participant information was coded in six different subcategories including number of total sample size, number of experiment group, grade, race/ethnicity, disability status and socio-economic status. Methodology information was coded by research design, sampling method, measurement construction, role of person implementing intervention and duration of intervention.

Before starting actual coding, researchers received training in three steps: (a) introducing the coding categories and discussing what each code meant, (b) modeling the coding process by the expert trainer and (c) practising coding with pair coders. Interrater reliability was calculated based on the agreement among coders and ranged from 87% (findings) to 94% (study information). Finally, coders came to a complete agreement after discussing similarities and differences in their individual codes.

Effect size calculations

Using the notations from Lipsey and Wilson (2001) and Borenstein, Hedges, Higgins, and Rothstein (2009), we calculated Cohen’s effect sizes, Hedge’s unbiased effect sizes, variances and weights. Cohen’s effect sizes (*d*) were calculated using standardized mean differences from treatment and control group and standard errors from each study.

$$d = \frac{\overline{Y}^{\mathrm{T}} - \overline{Y}^{\mathrm{C}}}{S_{Y\mathrm{pooled}}} \tag{1}$$

We corrected a potential bias that may exist because of small samples by using Hedge’s unbiased effect sizes (g) and calculated it to adjust the sample sizes with a $c(m)$ correction factor.

$$g = d{*}c(m) \tag{2}$$

In equation (2), $c(m) = 1 - 3/(4m - 1)$ in which $m = n^{\mathrm{T}} + n^{\mathrm{C}} - 2$, where n^{T} stands for the number of subjects in treatment group and n^{C} stands for the number of subjects in the control group. Then, we calculated the variance of the effect size using equation (3).

$$v = \frac{n^{\mathrm{T}} + n^{\mathrm{C}}}{n^{\mathrm{T}} n^{\mathrm{C}}} + \frac{g^2}{2(n^{\mathrm{T}} + n^{\mathrm{C}})} \tag{3}$$

© 2017 John Wiley & Sons Ltd

Table 1. Details of included articles for this meta-analysis

Study	Language	Device type	Subject area	Grade level	Study design	Implementer	Effect size(s) (*g*)	95% confidence interval
Ahmed and Parsons (2013)	English	Smartphone	Science	High school	Quasi-experimental	Teacher	0.277	[−0.077, 0.629]
Billings and Mathison (2012)	English	Mobile device	Science	Elementary	Quasi-experimental	Teacher	0.491	[0.130, 0.845]
							0.520	[0.206, 0.830]
							0.782	[0.397, 1.155]
							0.624	[0.298, 0.944]
Carr (2012)	English	Tablet	Mathematics	Elementary	Quasi-experimental	Researcher	0.389	[−0.003, 0.775]
de-Marcos et al. (2010)	Non-English	Smartphone	Science	Secondary (middle)	Quasi-experimental	Teacher	0.761	[0.349, 1.160]
							0.348	[−0.054, 0.744]
							0.368	[−0.165, 0.891]
Huang et al. (2010)	Non-English	PDA	Science	Elementary	Quasi-experimental	Teacher	1.172	[0.395, 1.890]
Hwang et al. (2013)	Non-English	PDA	Science	Elementary	Quasi-experimental	Teacher	0.599	[0.054, 1.127]
Kim et al. (2011)	Non-English	Mobile device	Reading	Elementary	Quasi-experimental	Teacher	0.276	[−0.369, 0.910]
							0.125	[−0.443, 0.689]
Lin (2014)	English	Tablet	Reading	High school	Experimental	Teacher	1.448	[0.955, 1.915]
Liu and Lee (2013)	Non-English	Tablet	Mathematics	High school	Quasi-experimental	Researcher	0.257	[−0.165, 0.674]
							0.927	[0.574, 1.271]
							−1.616	[−2.083, 1.122]
Nedungadi and Raman (2012)	English	Smartphone	Science	Secondary (middle)	Quasi-experimental	Teacher	−0.190	[−0.711, 0.336]
Riconscente (2013)	English	Tablet	Mathematics	Elementary	Experimental	Teacher	0.736	[0.162, 1.288]
							0.255	[−0.306, 0.808]
Varma (2014)	English	PDA	Science	Elementary	Experimental	Teacher	0.243	[−0.464, 0.937]
							0.548	[−0.176, 1.245]
							0.703	[−0.040, 1.410]
							0.252	[−0.462, 0.954]
Yang, Hwang, et al. (2013b)	Non-English	Smartphone	Science	Elementary	Quasi-experimental	Teacher	1.071	[0.162, 1.288]
							0.559	[0.053, 1.052]
Yang, Tseng, et al. (2013a)	Eng.	PDA	Reading	Secondary (middle)	Quasi-exp.	Researcher	0.740	[0.224, 1.238]

© 2017 John Wiley & Sons Ltd

The inverse variance weight was computed for each effect size. This procedure assigned larger weights to the studies with larger samples.

Publication bias

This meta-analysis included 27 effect sizes from 14 identified studies. Funnel plots have been used to identify publication bias (Bartolucci & Hillegass, 2010; Egger, Smith, Schneider, & Minder, 1997). Studies with larger sample sizes present towards the top of the funnel plot, whereas studies with smaller sample sizes present towards the bottom of the funnel plot (Rothstein, 2008). In our study, most of the effect sizes showed symmetry in the funnel plot.

Rosenthal's (1979) failsafe-N procedure resulted in a value of 1004 missing studies with zero effect at the target p-value of 0.05. Even though Rosenthal did not provide a specific cut-off for the maximum number of missing studies, he suggested a general guideline in which a failsafe-N that is larger than $5K + 10$, where K is the number of identified studies, would be robust to possible publication bias (Rothstein, 2008). In our study, $5K + 10$ is 80 studies, which indicates that the results of our study are robust to the threat of publication bias. We also carried out Egger's test (Egger et al., 1997), which is a linear regression method that assesses the publication bias by the funnel plot. The Egger test of the intercept was not significant, $t = -0.38$, $p = 0.70$, suggesting that the funnel plot is symmetric (Figure 1). Therefore, we concluded that publication bias was not a threat to the validity of our study.

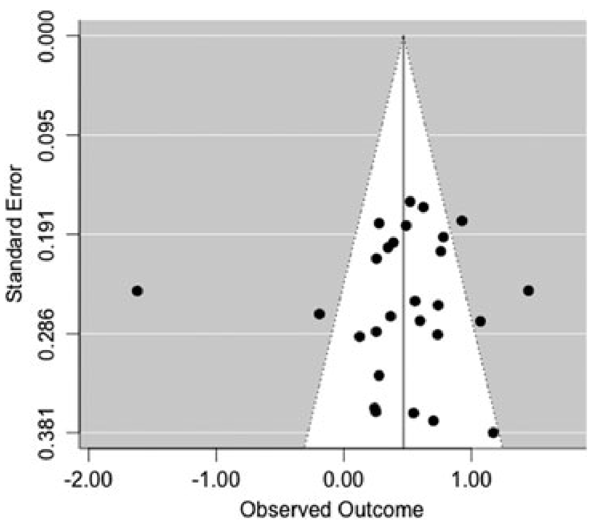

Figure 1 Funnel plot of Hedge's g and standard error

Statistical analysis

As a first step in analysis, we calculated the descriptive statistics for each moderator variable and reported them in the results section. Some of the studies had more than one outcome; thus, they had multiple effect sizes. In regular meta-analytic procedures, having multiple outcomes for each study leads to the violation of independence assumption. To take account of dependence, one could select the best effect size for each study or average the multiple outcome effects. Because some of the studies might have larger numbers of effect sizes than those of the other studies, this generally results in loss of information. Becker (2000) and Olkin and Gleser (2009) provided an approach with formulas for covariance among dependent effect size. However, following this approach is difficult because studies usually do not report the correlations of the outcomes. As a result, we followed a multilevel model approach to deal with the possible dependence problem (Berkey, Anderson, & Hoaglin, 1996; Konstantopoulos, 2011; Olkin & Gleser, 2009) in this meta-analytic study. Additionally, Bryk and Raudenbush (1992) suggested a procedure in which multilevel models can be used to test the significance of the mean of the estimates, the variance of the estimates and the association with the estimate and the moderator variables. Because the data in this study have effect sizes that were nested in the effect size IDs and effect size IDs were nested in the studies, we used a three-level multilevel modeling strategy to obtain the overall effect size and explain the variance of the study. With the notations from Bryk and Raudenbush, Konstantopoulos (2011) and Acar and Sen (2013), the level 1 model (i.e., effect size level) is demonstrated as

$$d_{ig} = \delta_{ig} + e_{ig}, \quad (4)$$

where g represents the level 3 units (i.e., study level), d_{ig} is the standardized effect size from study g, δ_{ig} is the corresponding population parameter, e_{ig} is the sampling error associated with effect size with $e_{ig} \sim N(0, V_{ig})$ and V_{ig} is assumed to be known. Then, we defined level 2 as the effect size ID level in which every effect size was assigned to a number from 1 to 27, as in total, we had 27 effect sizes in 14 studies. The level 2 model (i.e., effect size ID level) is within the study level (level 3) where the level 2 models can be used to explore whether the heterogeneity can be attributed to the effect size ID level.

$$\delta_{ig} = \beta_{0g} + r_{0g} \quad (5)$$

The fully unconditional model is

$$\beta_{0g} = \gamma_{00} + u_{og}, \quad (6)$$

where the level 3 unit means vary around a γ_{00}, which is the grand mean effect size, and u_{0g} is the error associated between studies with u_{0g} ~ N(0, τ). Then, the level 3 model (i.e., study level) is therefore

$$d_{ig} = \gamma_{00} + u_{0g} + r_{0g} + e_{ig}. \quad (7)$$

If we add the moderator variables associated with the study characteristics, equation (5) becomes

$$\delta_{ig} = \beta_{00} + \beta_{10}X_{1ig} + \cdots + \beta_{pg}X_{pig} + r_{0g}. \quad (8)$$

Equation (8) introduces p numbers of moderators to the model. We first fit a fully unconditional model to estimate the mean and variance, and then we examined the between-studies differences by fitting models with added moderator variables. The six moderator variables were device type, subject area, intervention language, grade level, study design and implementer. Because all of the moderators were categorical, we used dummy coding. Restricted maximum likelihood estimation with weights was used to estimate the model parameters. We used R (R Development Core Team, 2015) and *metafor* (Viechtbauer, 2010) to perform a weighted multilevel meta-analysis of the relationship between achievement scores and moderators. Weights were estimated by taking the inverse function of the effect size variances.

The homogeneity test

A heterogeneity test (i.e., Cochran's Q statistics) determines whether the studies all have the same effect in the population. The homogeneous model states the variation among the studies is only due to sampling error. However, if the model is heterogeneous, the variance among the studies is larger than one would expect based on sampling error alone. In addition, a significant Q test suggests that the outcomes are heterogeneous. To support the heterogeneity test results, I^2 statistics, which represents the percentage of the heterogeneity where higher values of I^2 represent higher-level of heterogeneity, was calculated.

Results

Overview of primary studies

Descriptive statistics (Table 2) showed that half of the studies were implemented in English, while half were implemented in different languages such as Chinese, Spanish and Taiwanese. Of the 14 studies, four studies reported the device type as tablets, four studies as PDAs and four studies as smartphones. In two studies, researchers reported the device type as simply 'mobile devices'. More than half of the identified studies were in science, whereas the others were in mathematics and reading. Most interventions occurred at the elementary level while a few studies were conducted at the middle school level. In addition, researchers most commonly used a quasi-experimental study design. Finally, a large number of researchers asked teachers to implement their interventions, whereas a small number of researchers implemented theirs by themselves.

Mean effect sizes were calculated for the categorical moderators. However, mean effect sizes did not differ statistically within the categories of moderators except the significant difference between reading and mathematics in the *subject* category. For instance, the analysis suggested that the effect of mobile devices on achievement among elementary, middle and high school grade levels was not significant. Although the results did not

Table 2. Descriptive statistics of included studies

Moderator variable	Identified categories	Counts (%)
Device type	Tablet	4 (28.6)
	PDA	4 (28.6)
	Smartphone	4 (28.6)
	Mobile device	2 (14.3)
Subject area	Mathematics	3 (21.4)
	Science	8 (57.1)
	Reading	3 (21.4)
Study language	English	7 (50)
	Non-English	7 (50)
Grade level	Elementary	8 (57.1)
	Middle	2 (14.3)
	High	4 (28.6)
Study design	Quasi-experimental	11 (78.6)
	Experimental	3 (21.4)
Implementer	Teacher	11 (78.6)
	Researcher	3 (21.4)

© 2017 John Wiley & Sons Ltd

suggest a significant change in effect sizes across the *grade level* category, mean effect sizes tended to decrease from elementary school to high school (Table 3). Similarly, mean effect sizes from the largest to the smallest in *device* were PDA, mobile device, smartphone and tablet. For the *language* moderator, the studies in which the instruction language was English had a larger effect size than the studies in which instruction was delivered in a different language. In the *design* category, experimental studies had a larger effect size as compared to quasi-experimental studies. Lastly, the studies where the *implementers* were the teachers had a larger effect size as compared to the studies that were carried out by the researchers. As mentioned previously, a statistically significant mean effect size difference was found between reading and mathematics in the *subject* category, and this will be discussed in more detail in the moderator analysis section.

Overall analysis

The heterogeneity test yielded a significant result, $Q(26) = 123.505$, $p < 0.001$. Also, the I^2 value was calculated as 80.67%. Based on these statistics, the effect sizes from the studies showed heterogeneity. Then, we ran a random-effect model and found the overall effect size estimate as 0.48, with a confidence interval of [0.26, 0.71], which indicated a moderately significant effect in the unconditional model. Table 4 provides the results of the unconditional model. The level 2 and level 3 variance components were estimated as 0.232 and 0.000, respectively, and were found to be significant.

Table 4. Results of unconditional model

	Estimates	Standard error	95% confidence interval
Fixed effects			
Intercept	0.483*	0.113	[0.261, 0.706]
Variance components			
Second level	0.232*	0.064	[0.106, 0.475]
Third level	0.000*	0.097	[0.000, 0.191]

*$p < 0.05$.

Moderator analysis

We ran the random-effects model including the moderators into the unconditional model to examine whether moderator variables had any effect on the outcomes of the studies. Table 5 provides the model parameters with the moderators included. The test for heterogeneity of moderators was not significant, $Q_{\text{between}}(10) = 5.517$, $p > 0.05$. However, the test for residual heterogeneity was significant, $Q_{\text{within}}(16) = 90.963$, $p < 0.05$. In addition, the overall effect size estimate was 0.17, with a confidence interval of [−0.45, 0.79], which indicated a small effect in the model where all of the categorical moderators were coded as 0. The level 2 and level 3 variance components were estimated as 0.268 and 0.000, respectively, and still found to be significant. The variance

Table 3. Description of effect size statistics for categorical moderators

		Mean effect size	Standard error	95% confidence interval
Device type	Tablet	0.343	0.348	[−0.338, 1.026]
	PDA	0.613	0.123	[0.371, 0.855]
	Smartphone	0.452	0.136	[0.184, 0.720]
	Mobile device	0.539	0.080	[0.381, 0.696]
Subject area	Mathematics	0.160	0.361	[−0.547, 0.868]
	Science	0.528	0.067	[0.395, 0.661]
	Reading	0.666	0.309	[0.059, 1.272]
Study language	English	0.517	0.094	[0.333, 0.702]
	Non-English	0.423	0.192	[0.046, 0.800]
Grade level	Elementary	0.554	0.059	[0.437, 0.670]
	Middle	0.427	0.206	[0.023, 0.832]
	High	0.280	0.385	[−0.476, 1.036]
Study design	Experimental	0.628	0.191	[0.253, 1.002]
	Quasi-experimental	0.422	0.116	[0.193, 0.651]
Implementer	Teacher	0.546	0.073	[0.402, 0.690]
	Researcher	0.145	0.420	[−0.680, 0.970]

© 2017 John Wiley & Sons Ltd

Table 5. Conditional model with moderators

	Estimates	Standard error	*p*-value	95% confidence interval
Fixed effects				
Intercept	0.171	0.317	0.588	[−0.450, 0.793]
Device – PDA	−1.552	0.964	0.108	[−3.442, 0.338]
Device – smartphone	−1.592	1.013	0.116	[−3.577, 0.393]
Device – mobile device	−1.629	0.982	0.097	[−3.553, 0.295]
Subject – science	1.876	1.063	0.078	[−0.208, 3.960]
Subject – reading	1.595*	0.814	0.050	[0.000, 3.190]
Language – English	−0.162	0.378	0.669	[−0.903, 0.580]
School – elementary	0.427	0.441	0.334	[−0.438, 1.292]
School – secondary	0.176	0.455	0.699	[−0.716, 1.069]
Design – experimental	−0.028	0.672	0.967	[−1.346, 1.290]
Implementer – teacher	−0.129	0.759	0.865	[−1.616, 1.359]
Variance components				
Second level	0.268*	0.080		[0.110, 0.672]
Third level	0.000*	0.649		[0.000, 1.272]

Note. Language = a dummy variable coded 1 to identify English language and 0 for non-English language. Device = a dummy variable coded 1 to identify PDA, smartphone or mobile device and 0 for tablet. Subject = a dummy variable coded 1 to identify reading or science and 0 for mathematics. School = a dummy variable coded 1 to identify elementary or middle school and 0 for high school. Design = a dummy variable coded 1 to identify experimental design and 0 for quasi-experimental design. Implementer = a dummy variable coded 1 to identify general education teacher and 0 for others.
*$p < 0.05$.

component of the model was similar to the unconditional model. Overall, most of the random variation was at the second level (i.e., effect size ID level).

Results of the moderator analysis suggested that one moderator variable, subject, was a significant predictor of the effect size and accounted for the level 2 and level 3 variance. The difference in mean effect size between reading and mathematics was significant, $\gamma_{\text{reading}} = 1.595$, $p < 0.05$ and 95% confidence interval [0.000, 3.190]. Among subjects, reading achievement was statistically more effective than mathematics achievement. However, language of the study, device type, school type, study design and implementer were not significant predictors of the mean effect size and did not explain the level 2 and level 3 variance.

Discussion

In this meta-analysis, we investigated the effects of mobile devices on student achievement in science, mathematics and reading in grades K–12. Based on our inclusion criteria, we identified 14 studies published in peer-reviewed journals between 2010 and 2014 and identified six moderator variables that were assumed to be influential in student achievement.

Regarding the first question in the study, the results suggested a moderately significant overall mean effect size. The achievement difference existed for all content areas favouring the use of mobile devices compared to the traditional methods. In their meta-analysis studies about the effect of technology on student achievement, Cheung and Slavin (2012, 2013) found smaller effect size (ES = 0.15 and ES = 0.16) than the effect size in our meta-analysis study (ES = 0.48). In another meta-analysis about the effect of computer technologies on student achievement, Li and Ma (2010) reported higher effect size (ES = 0.28) than Cheung and Slavin's studies, but again it was lower than the effect size that we found in this study. The larger effect size in our study may imply advantages of mobile devices over other educational technologies on student achievement.

Results of moderator analysis for the second research question showed that there was a significant difference between the subject predictor variables of reading and mathematics. We found that the effect of using mobile devices on achievement in reading was statistically higher than that in mathematics. In a previous study on the effect of using tablets on student achievement, Piper and Kwayumba (2014) found similar results: Students' reading achievement scores were statistically higher than mathematics achievement scores, but there were no statistical differences either between reading and science or between science and mathematics. In addition, Cheung and Slavin (2011) examined the effectiveness of various educational technology on reading achievement in a meta-analysis study and pointed out that students'

© 2017 John Wiley & Sons Ltd

reading achievement scores were higher than mathematics achievement scores when subjects were taught with integrated teaching approach via educational technology. As a result, they suggested that using the integrated teaching approach via educational technology would be more effective in reading than that in mathematics.

Furthermore, other factors may explain the significant difference between reading and mathematics. First, in recent years, there has been a growing interest in the development of interactive e-books that provide enriched opportunities for students to improve their reading proficiencies. Additionally, as indicated by Slough, Cavlazoglu, Erdogan, Wakefield, and Akgun (2012), the integration of interactive elements (e.g., videos, audios and high-quality pictures) into these e-books could make reading more appealing for students, keep them more engaged and increase their achievement. Clarke and Svanaes's (2014) study on the use of mobile devices as e-books in reading noted that these devices supported students with literacy difficulties. In mathematics, however, most of the mobile device applications were mathematical games (Riconscente, 2013). Mathematical games may enhance student interest and motivation as well as some skills necessary in learning mathematics, but their effect on academic achievement could be limited. Second, use of active learning strategies in teaching mathematics with mobile devices is still insufficient (Carr, 2012) as the teaching in mathematics frequently occurred via traditional lecturing (Weber, 2004). Lastly, learning mathematics requires other skills (e.g., computational skills, abstract thinking, problem solving and spatial thinking; National Council of Teachers of Mathematics, 2016) in addition to basic reading skills for understanding the phenomena in mathematical problems. This could be another reason for mobile devices having a smaller effect in mathematics achievement as compared to reading. Learning mathematics with mobile devices, therefore, is likely to be more challenging because of the additional skills that students need to have.

In this study, we found no significant differences in effect sizes for the other remaining moderator variables, device type and subject, including language, grade level, design of the study and intervention implementer. As indicated by Belland, Walker, Olsen, and Leary (2015), lack of significant differences is critical and needs to be discussed in the meta-analyses.

Non-significant difference in language moderator analysis suggests that mobile devices work similarly in English and non-English languages. This result indicates that language difference is not an issue that impacts how mobile devices affect students' achievement in grades K–12. The latest improvements in digital devices' language interfaces allow precise translations, and this result may suggest that using mobile devices in classrooms in different languages is not a significant factor in students' achievement.

Non-significant differences for device type suggest that device effect on student achievement does not exist. Some studies referring to mobile devices as PDAs (e.g., Huang et al., 2010, 2014) implied that PDAs are more personal devices that have additional features such as cellular phone, wireless connection and handwriting with special pens. However, with the recent technological developments, almost all types of mobile devices can include many additional features based on the purpose of their usage. Additionally, many smartphone companies have begun producing more capable devices that come with wide screens as well as same features available in large mobile devices. Thus, smartphones can now function not only as cellular phones but also as mobile devices. As a result, calling all types of handheld devices (e.g., tablet, PDA and smartphone) as mobile devices can be more appropriate because use of different terms results in no significant difference.

It is worthwhile to note that there was no significant difference among grade levels (i.e., elementary, middle and high school). The recent national survey study on students' use of mobile devices by Poll (2014) revealed that elementary school students are eager to use mobile devices more often than middle and high school students were (elementary school students = 71%, middle school students = 67%, high school students = 56%), and their attitudes towards using mobile devices to become successful in their schools were found to be to be higher than middle and high school students (elementary school students = 85%, middle school students = 82%, high school students = 73%). Similar to Poll's result, we found elementary school students' achievements to be higher than those of other grade-level students (elementary school students = 0.554, middle school students = 0.427, high school students = 0.280). However, the results of our study showed that the differences were not statistically significant.

Results of non-significant differences between quasi-experimental and experimental studies indicated that random selection of participants, which is the distinguishing

© 2017 John Wiley & Sons Ltd

factor between quasi-experimental and experimental designs, was not likely to be a distinctive process impacting the effect of mobile devices on students' achievement. Both designs have the assumption of equality between experimental and control groups before the intervention, and even a random selection may not result in ideal equality because the subjects are human beings in the social science research.

Finally, there was no significant difference among implementers. This result showed that intervention of mobile devices either with teachers or researchers does not impact students' achievement. One may argue that because teachers are more familiar with the classroom and their students, they may be more effective in their intervention. However, if researchers can spend some time with students and find opportunities to know them before interventions, the researchers may perform similar to or be more effective than teachers. Hence, we conclude that implementers as teachers or researchers in mobile device implementations do not statistically influence students' achievement in science, mathematics and reading in grades K–12.

Implications and recommendations

To decide whether mobile devices should be integrated into classrooms is challenging. Decisions as made individually by teachers, or more widely by decision makers, on whether to use mobile devices in class should include a close examination of conditions that make integration of mobile devices successful. In other words, one should thoroughly consider the extent to which other important factors may influence the effectiveness of mobile devices. Researchers should interpret the results of research studies comparing mobile device use in treatment and control groups cautiously because of the wide variability across studies (Tamim, Bernard, Borokhovski, Abrami, & Schmid, 2011). Other factors not identified in a study may account for the success or failure of mobile device interventions. To name a few, instructional design decisions, goals of instruction (Ross et al., 2010), fidelity of an intervention (Tamim et al., 2011), pedagogical approach, implementer's disposition towards the use of mobile device and other factors will be crucial. Therefore, we suggest that the results of studies that investigate effects of mobile devices be interpreted with caution because they might not equally apply to various settings.

This study showed that mobile devices had a moderate effect on student achievement. However, the effect could be potentially larger if studies were designed and reported rigorously. An unpleasant issue we encountered in this meta-analysis study was the lack of precision when research reports described variables and terms that were assumed to be crucial. As suggested by Lan, Lo, and Hsu (2014), we recommend researchers to describe the terms they use thoroughly (e.g., which characteristics of a device warrants it to be called a mobile device) and provide as much detail as possible, especially about the teaching strategy used during the mobile device intervention, specifics of assessment, duration and frequency of implementation, description of research participants and in what ways mobile devices are integrated into classroom (e.g., how it is warranted that mobile devices are not simply a tool for delivering the content; Tamim et al., 2011). As recommended by Bartolucci and Hillegass (2010), providing a comprehensive description of a study is helpful for both research and practice. A comprehensive description of a research study provides opportunities for other researchers to include the study into meta-analyses with many pertinent variables. Providing a thorough description is also helpful for practitioners because it enables them to examine what other factors should be considered before making a decision in a mobile device intervention. Finally, because the results of this meta-analysis demonstrated many non-significant effect sizes of various moderators, we agree with the suggestion of Tamim et al. (2011) that more studies are needed that go beyond mere comparison of technology versus non-technology interventions; it is essential to conduct more empirical studies that investigate the impact of variables when mobile devices are used to increase students' achievement.

Conclusion

Introduction of diverse technologies has led us to be more engaged in classroom practices through a variety of technological devices. Of these technologies, mobile devices have gained popularity and become more practical and functional to meet our demands in this technology-driven era. In the relevant literature, a growing body of research has investigated the effects of mobile devices on student achievement and revealed that mobile devices have significant effects on student achievement in different subject areas in grades K–12.

© 2017 John Wiley & Sons Ltd

Additionally, a number of review studies about mobile devices (i.e., Hung & Zhang, 2012; Hwang & Tsai, 2011; Wu et al., 2012) provided substantial information about the effects of mobile devices on student achievement. With all these borne in mind, this meta-analysis reports what has been done about the effects of mobile devices by combining findings from related studies in the literature. Specifically, examining the relevant experimental and quasi-experimental research in the field, this study documented how the achievement scores in science, mathematics and reading subject areas differ depending on the use of mobile devices or traditional teaching methods in the classrooms. According to the results, the subject areas taught through mobile devices had significantly higher achievement scores than the ones taught with traditional teaching methods. Regarding the analysis of moderator variables, use of mobile devices in reading subject area yields more effective achievement scores than that in mathematics.

Limitations

The interpretation of the results entails limitations. During the moderator variable identification process, we targeted several moderator variables to be included in this meta-analytic study. However, we were only able to use six of these variables as some of the included studies did not explicitly report some of their variables. This prevented us from incorporating more moderator variables that may have been significant in our meta-regression model. In this regard, Bartolucci and Hillegass (2010) indicated the ambiguous description of the methodology and inconsistent use of variables across studies as the potential limitations of meta-analyses. In this study, we faced these limitations. For example, we considered the duration of intervention with mobile devices as a critical moderator variable influencing student achievement but found that most of the researchers did not explicitly report their intervention duration. Students' socio-economic status, gender, ethnicity, parents' education level, measure of outcome and students' attitudes towards using mobile devices were other moderator variables we initially considered adding into our analysis but could not, owing to the unavailability of these data across the identified studies. As a result, we had to exclude some of the potential variables in data analysis. Therefore, we suggest that researchers explicitly report the data that may explain the significant effect of mobile device on student achievement and document detailed explanations of their implementations.

Funding

This study was not funded by any foundation.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

Note: References with asterisk (*) show the studies used for the meta-analysis.

Acar, S., & Sen, S. (2013). A multilevel meta-analysis of the relationship between creativity and schizotypy. *Psychology of Aesthetics, Creativity, and the Arts*, *7*(3), 214.

Agostinho, S., Tindall-Ford, S., Ginns, P., Howard, S. J., Leahy, W., & Paas, F. (2015). Giving learning a helping hand: Finger tracing of temperature graphs on an iPad. *Educational Psychology Review*, *27*, 427–443.

Ahmed, S., & Parsons, D. (2013). Abductive science inquiry using mobile devices in the classroom. *Computers & Education*, *63*, 62–72.*

Sahin, A., Cavlazoglu, B., & Zeytuncu, Y. (2015). Flipping a college calculus course: A case study. *Journal of Educational Technology & Society*, *18*, 142–152.

Slough, S., Cavlazoglu, B., Erdogan, N., Wakefield, J., & Akgun, O. (2012). Designers discussing design: An investigation into the design of an innovative science textbook designed to take advantage of multi-touch tablet technology and the cloud. In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2012* (pp. 1227–1232). Chesapeake, VA: AACE.

Bartolucci, A. A., & Hillegass, W. B. (2010). Overview, strengths, and limitations of systematic reviews and meta-analyses. In *Evidence-based practice: Toward optimizing clinical outcomes* (pp. 17–33). Berlin Heidelberg: Springer.

Baya'a, N., & Daher, W. (2009). Learning mathematics in an authentic mobile environment: The perceptions of students. *International Journal of Interactive Mobile Technologies*, *3*, 6–14.

Becker, B. J. (2000). Multivariate meta-analysis. In *Handbook of applied multivariate statistics and mathematical modeling* (pp. 499–525). San Diego: Academic Press.

Belland, B. R., Walker, A. E., Olsen, M. W., & Leary, H. (2015). A pilot meta-analysis of computer-based scaffolding in STEM education. *Educational Technology & Society*, *18*, 183–197.

© 2017 John Wiley & Sons Ltd

Berkey, C. S., Anderson, J. J., & Hoaglin, D. C. (1996). Multiple-outcome meta-analysis of clinical trials. *Statistics in Medicine*, *15*, 537–557.

Billings, E. S., & Mathison, C. (2012). I get to use an iPod in school? Using technology-based advance organizers to support the academic success of English learners. *Journal of Science Education and Technology*, *21*, 494–503.*

Borenstein, M., Hedges, L. V., & Higgins, J. P. T. (2009). Introduction to Meta-Analysis (2). Hoboken, GB: Wiley.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park: Sage Publications.

Caird, J. K., Willness, C. R., Steel, P., & Scialfa, C. (2008). A meta-analysis of the effects of cell phones on driver performance. *Accident Analysis & Prevention*, *40*, 1282–1293.

Carr, J. M. (2012). Does math achievement h'APP'en when iPads and game-based learning are incorporated into fifth-grade mathematics instruction? *Journal of Information Technology Education*, *11*, 269–286.*

Chai, Z., Vail, C. O., & Ayres, K. M. (2015). Using an iPad application to promote early literacy development in young children with disabilities. *The Journal of Special Education*, *48*, 268–278.

Cheung, A. C., & Slavin, R. E. (2011). *Effectiveness of educational technology applications for enhancing reading achievement in K–12 classrooms: A meta-analysis*. Baltimore, MD: Center for Research and Reform in Education, Johns Hopkins University.

Cheung, A. C., & Slavin, R. E. (2012). How features of educational technology applications affect student reading outcomes: A meta-analysis. *Educational Research Review*, *7*, 198–215.

Cheung, A. C., & 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.

Chu, H. C., Hwang, G. J., Tsai, C. C., & Tseng, J. C. (2010). A two-tier test approach to developing location-aware mobile learning systems for natural science courses. *Computers & Education*, *55*, 1618–1627.

Clarke, B., & Svanaes, S. (2014). An updated literature review on the use of tablets in education. Tablets for Schools. UK: Family Kids & Youth.

Crowley, K., McLaughlin, T., & Kahn, R. (2013). Using direct instruction flashcards and reading racetracks to improve sight word recognition of two elementary students with autism. *Journal of Developmental and Physical Disabilities*, *25*, 297–311.

de-Marcos, L., Hilera, J. R., Barchino, R., Jiménez, L., Martínez, J. J., Gutiérrez, J. A., … Otón, S. (2010). An experiment for improving students performance in secondary and tertiary education by means of m-learning auto-assessment. *Computers & Education*, *55*, 1069–1079.*

Doolittle, P. E., & Mariano, G. J. (2008). Working memory capacity and mobile multimedia learning environments: Individual differences in learning while mobile. *Journal of Educational Multimedia and Hypermedia*, *17*, 511–530.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ: British Medical Journal*, *315*, 629–634.

Huang, K. L., Chen, K. H., & Ho, C. H. (2014). Promoting in-depth reading experience and acceptance: Design and assessment of tablet reading interfaces. *Behavior & Information Technology*, *33*, 606–618.

Huang, Y. M., Lin, Y. T., & Cheng, S. C. (2010). Effectiveness of a mobile plant learning system in a science curriculum in Taiwanese elementary education. *Computers & Education*, *54*(1), 47–58.*

Hung, P. H., Hwang, G. J., Lin, Y. F., Wu, T. H., & Su, I. H. (2013). Seamless connection between learning and assessment-applying progressive learning tasks in mobile ecology inquiry. *Journal of Educational Technology & Society*, *16*, 194–205.

Hung, J. L., & Zhang, K. (2012). Examining mobile learning trends 2003–2008: A categorical meta-trend analysis using text mining techniques. *Journal of Computing in Higher Education*, *24*(1), 1–17.

Hwang, G. J., & Tsai, C. C. (2011). Research trends in mobile and ubiquitous learning: A review of publications in selected journals from 2001 to 2010. *British Journal of Educational Technology*, *42*(4), E65–E70.

Hwang, G. J., Wu, P. H., Zhuang, Y. Y., & Huang, Y. M. (2013). Effects of the inquiry-based mobile learning model on the cognitive load and learning achievement of students. *Interactive Learning Environments*, *21*, 338–354.*

Institute of Education Sciences. (2003). What Works Clearinghouse study review standards. Retrieved 10 January 2005, from What Works Clearinghouse Web site: http://www.whatworks.ed.gov/reviewprocess/study_standards_final.pdf

Ketamo, H. (2003). xTask – An adaptable learning environment. *Journal of Computer Assisted Learning*, *19*(3), 360–370.

Kim, P., Hagashi, T., Carillo, L., Gonzales, I., Makany, T., Lee, B., & Gàrate, A. (2011). Socioeconomic strata, mobile technology, and education: A comparative analysis. *Educational Technology Research and Development*, *59*, 465–486.*

Konstantopoulos, S. (2011). Fixed effects and variance components estimation in three-level meta-analysis. *Research Synthesis Methods*, *2*(1), 61–76.

Lan, Y. C., Lo, Y. L., & Hsu, Y. S. (2014). The effects of meta-cognitive instruction on students' reading comprehension in computerized reading contexts: A quantitative meta-analysis. *Educational Technology & Society*, *17*(4), 186–202.

Li, Q., & Ma, X. (2010). A meta-analysis of the effects of computer technology on school students' mathematics learning. *Educational Psychology Review*, *22*(3), 215–243.

© 2017 John Wiley & Sons Ltd

Li, S. C., & Pow, J. C. (2011). Affordance of deep infusion of one-to-one tablet-PCs into and beyond classroom. *International Journal of Instructional Media*, *38*(4), 319–326.

Lin, C. C. (2014). Learning English reading in a mobile-assisted extensive reading program. *Computers & Education*, *78*, 48–59.*

Lin, M. F., Fulford, C. P., Ho, C. P., Iyoda, R., & Ackerman, L. K. (2012). Possibilities and challenges in mobile learning for K-12 teachers: a pilot retrospective survey study. In *Wireless, Mobile and Ubiquitous Technology in Education (WMUTE), 2012 IEEE Seventh International Conference on* (pp. 132–136). IEEE.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis* (Vol. *49*). Thousand Oaks, CA: Sage.

Liu, T. Y. (2009). A context-aware ubiquitous learning environment for language listening and speaking. *Journal of Computer Assisted Learning*, *25*, 515–527.

Liu, N., & Lee, G. (2013). iPad infuse creativity in solid geometry teaching. *The Turkish Online Journal of Educational Technology*, *12*, 177–192.*

Looi, C. K., Zhang, B., Chen, W., Seow, P., Chia, G., Norris, C., & Soloway, E. (2011). 1:1 mobile inquiry learning experience for primary science students: A study of learning effectiveness. *Journal of Computer Assisted Learning*, *27*, 269–287.

McClanahan, B., Williams, K., Kennedy, E., & Tate, S. (2012). A breakthrough for josh: How use of an iPad facilitated reading improvement. *TechTrends: Linking Research and Practice to Improve Learning*, *56*, 20–28.

National Council of Teachers of Mathematics. (2016). Common core state standards for mathematics. Retrieved from http://www.nctm.org/uploadedFiles/Standards_and_Positions/Common_Core_State_Standards/Math_Standards.pdf

National Educational Technology Standards. (2007) Profiles for technology (ICT) literate students. Retrieved from https://www.iste.org/docs/pdfs/nets-s-2007-student-profiles-en.pdf?sfvrsn=4https://www.iste.org/docs/pdfs/nets-s-2007-student-profiles-en.pdf?sfvrsn=4

Nedungadi, P., & Raman, R. (2012). A new approach to personalization: Integrating e-learning and m-learning. *Educational Technology Research and Development*, *60*, 659–678.*

Olkin, I., & Gleser, L. (2009). Stochastically dependent effect sizes. In *The handbook of research synthesis and meta-analysis* (p. 357). New York: Russell Sage Foundation.

Partnership for 21st Century Skills. (2011). Framework for 21st century learning. Retrieved from http://www.p21.org/storage/documents/1.__p21_framework_2-pager.pdfhttp://www.p21.org/storage/documents/1.__p21_framework_2-pager.pdf

Perfetti, C., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 227–247). Malden, MA: Blackwell Publishing.

Piper, B., & Kwayumba, D. (2014). USAID/Kenya Primary Math and Reading (PRIMR) initiative: Kisumu information and communication technology (ICT) intervention. Endline report. USA: Research Triangle Institute (RTI). Retrieved from https://www.eddataglobal.org/countries/index.cfm?fuseaction=pubDetail&ID=664

Poll, H. (2014). Pearson student mobile device survey 2014. Retrieved from http://www.pearsoned.com/wp-content/uploads/Pearson-K12-Student-Mobile-Device-Survey-050914-PUBLIC-Report.pdf

R Development Core Team. (2015). R: A Language and Environment for Statistical Computing. Vienna, Austria: the R Foundation for Statistical Computing. ISBN: 3-900051-07-0 Available online at http://www.R-project.org/.

Rafdal, B. H., Mcmaster, K. L., Mcconnell, S. R., Fuchs, D., & Fuchs, L. S. (2011). The effectiveness of kindergarten peer-assisted learning strategies for students with disabilities. *Exceptional Children*, *77*, 299–316.

Riconscente, M. M. (2013). Results from a controlled study of the iPad fractions game Motion Math. *Games and Culture*, *8*, 186–214.*

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*, 638–641. doi:10.1037//0033-2909.86.3.638.

Rosenthal, R., & DiMatteo, M. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology*, *52*(1), 59–82. doi:10.1146/annurev.psych.52.1.59.

Ross, S. M., Morrison, G. R., & Lowther, D. L. (2010). Educational technology research past and present: Balancing rigor and relevance to impact school learning. *Contemporary Educational Technology*, *1*(1), 17–35.

Rothstein, H. R. (2008). Publication bias as a threat to the validity of meta-analytic results. *Journal of Experimental Criminology*, *4*(1), 61–81. doi:10.1007/s11292-007-9046-9.

Scanlon, E., Jones, A. C., & Waycott, J. (2005). Mobile technologies: Prospects for their use in learning in informal science settings. Retrieved from http://jime.ubiquitypress.com/articles/10.5334/2005-25/

Tamim, R. M., Bernard, R. M., Borokhovski, E., Abrami, P. C., & Schmid, R. F. (2011). What forty years of research says about the impact of technology on learning: A second-order meta-analysis and validation study. *Review of Educational Research*, *81*(1), 4–28. doi:10.3102/0034654310393361.

Torgesen, J. K. (2002). Lessons learned from intervention research in reading: A way to go before we rest. R. Stainthorp and P. Tomlinson (Edn.). Learning and teaching reading. British Journal of educational psychology, monograph series II: Psychological aspects of education – Current trends, (1), 89–104.

© 2017 John Wiley & Sons Ltd

Varma, K. (2014). Supporting scientific experimentation and reasoning in young elementary school students. *Journal of Science Education and Technology*, *23*, 381–397.*

Vaughn, S., Wexler, J., Leroux, A., Roberts, G., Denton, C., Barth, A., & Fletcher, J. (2012). Effects of intensive reading intervention for eighth-grade students with persistently inadequate response to intervention. *Journal of Learning Disabilities*, *45*, 515–525.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48 .http://www.jstatsoft.org/v36/i03/

Viswanathan, P. (2015). What is a mobile device? Retrieved from http://mobiledevices.about.com/od/glossary/g/What-Is-A-Mobile-Device.htmhttp://mobiledevices.about.com/od/glossary/g/What-Is-A-Mobile-Device.htm

Wanzek, J., & Vaughn, S. (2008). Response to varying amounts of time in reading intervention for students with low response to intervention. *Journal of Learning Disabilities*, *41*, 126–142.

Weber, K. (2004). Traditional instruction in advanced mathematics courses: A case study of one professor's lectures and proofs in an introductory real analysis course. *The Journal of Mathematical Behavior*, *23*, 115–133.

Wu, W. H., Wu, Y. C. J., Chen, C. Y., Kao, H. Y., Lin, C. H., & Huang, S. H. (2012). Review of trends from mobile learning studies: A meta-analysis. *Computers & Education*, *59*(2), 817–827.

Yang, C. C., Tseng, S. S., Liao, A. Y., & Liang, T. (2013a). Situated poetry learning using multimedia resource sharing approach. *Journal of Educational Technology & Society*, *16*, 282–295.*

Yang, C. C., Hwang, G. J., Hung, C. M., & Tseng, S. S. (2013b). An evaluation of the learning effectiveness of concept map-based science book reading via mobile devices. *Journal of Educational Technology & Society*, *16*, 167–178.*

© 2017 John Wiley & Sons Ltd