Author: Wang, J.
Title: Technology integration in education: policy plans, teacher practices, and student outcomes
Issue Date: 2021-07-06

The effects of mobile technology usage on cognitive, affective, and behavioral learning outcomes in primary and secondary education: A systematic review with meta-analysis

5

This chapter was submitted in an adapted form as:

Wang, J., Tigelaar, D. E., & Admiraal, W. The effects of mobile technology usage on cognitive, affective, and behavioral learning outcomes in primary and secondary education: A systematic review with meta-analysis.

Abstract

The impact of mobile technology usage on student learning in various educational stages has been continuously studied in empirical and review studies. In the course of researchers' inquiry into the extent of enhancement of learning outcomes, systematic quantitative analyses of mobile devices' effects on both cognitive and non-cognitive learning outcomes, with a particular emphasis on primary and secondary education, are lacking. This study aimed to synthesize the effects of using mobile technology on cognitive, affective, and behavioral learning outcomes in primary and secondary education. Based on our inclusion and exclusion criteria, we found 61 studies of 56 peer-reviewed papers (N = 6406) from electronic databases and major journals in educational technology and mobile learning between 2014 and 2020. We then examined 15 moderators that were expected to affect student learning outcomes. Compared with traditional technology and non-technology groups, using mobile technology produced medium positive and statistically significant effects on primary and secondary students' learning, in terms of cognitive (g = 0.547), affective (g = 0.514), and behavioral (g = 0.543) learning outcomes. Further moderator analyses revealed that student factors (i.e., students' socioeconomic status), learning process (i.e., hardware used, student-to-hardware ratio) and study quality (i.e., learning content/ topic equivalence, software/ tool equivalence, procedure of effect size extraction) were among the variables that moderated the summary effect sizes (ESs) for at least one learning outcome dimension significantly. The findings and their implications for researchers, policymakers, and practitioners are discussed.

Keywords

Mobile technologies; Learning outcomes; Primary education; Secondary education; Meta-analysis

5

5.1 Introduction

Mobile technology is characterized by wireless internet-connected devices, including smartphones, clickers, tablets, and laptops, etc. Considering the rapid growth and affordability of mobile technology, mobile learning, known as "learning across multiple contexts, through social and content interactions using personal electronic devices" (Crompton, 2013, p. 4), has become a fast-growing research field in the world (Soloway & Norris, 2018). The proliferation of mobile technology provides researchers with the opportunity to reimagine teaching and learning with mobile technology (Mayer, 2020). Recent literature has identified the exciting potential of integrating mobile technology in education. For example, instantly gathering student data from mobile devices can help teachers monitor students' learning progress and deliver differentiated instruction in class (Lee, Hao, Lee, Sim, & Huang, 2019), and support teachers plan and orchestrate through reflection on and evaluation of their teaching (Wise, 2019). Beyond the importance to teachers, most mobile technology research reported on increased learning achievement in the language (Alfadil, 2020), science (Chang et al., 2020), mathematics (Zhu & Urhahne, 2018), and social studies (Huang, Chen, & Hsu, 2019), followed by students' perceptions of motivation (Lee et al., 2019) and attitude (Sahin & Yilmaz, 2020). Benefit from the increased learning mobility, mobile learning also facilitates social interaction (Hwang, Lai, Liang, Chu, & Tsai, 2018) and knowledge co-creation (Lim, Shelley, & Heo, 2019). Researchers have pointed out the critical role of self-efficacy in shaping students' behavioral engagement in mobile learning (Xie, Heddy, & Vongkulluksn, 2019). Therefore, mobile technology affords students to learn, both individually and collectively (Koole, 2009). There are some minor concerns on mobile learning, however, regarding distractive effect (Zhai, Zhang, Li, & Zhang, 2019), misuse (Ravizza, Uitvlugt, & Fenn, 2017), self-control challenges (Troll, Friese, & Loschelder, 2020), and heavy cognitive load (Chu, 2014).

To date, because the non-cognitive domains were less reported and seemed less relevant for informal settings, the pooled effects of mobile technology on learning have mainly been limited to cognitive learning. We argue that the targeted learning goals for "21st-century skills" include cognitive goals, affective or intrapersonal goals, and behavioral or interpersonal goals, and research needs to go beyond concentrate on measuring cognitive learning gains (Pellegrino & Hilton, 2013). Recently, highly cited articles on mobile learning have focused more on the affective and behavioral dimensions (Lai, 2020). It is yet to be known what are the overall effects of mobile technology on affective and behavioral learning outcomes, which play a vital role in understanding students' learning from alternative perspectives.

Moreover, given the richness and complexities of mobile learning, it is important to develop a greater understanding of how to optimize the implementation and interpret the different learning outcomes (Rogaten, et al., 2019). Specifically, more research is needed on the best practices for using mobile technology in order to figure out when and how children should use mobile devices (Crompton & Burke, 2020). Because of the numerous, significant differences found in the available studies between primary and secondary education, and post-secondary education (Schmid et al., 2014), one could argue that mobile technologies are more effective for lower education levels. For example, a recent meta-analysis on audience response systems revealed that the effect is much more significant in experiments performed in non-university contexts than in the university context (Castillo-Manzano, Castro-Nuño, López-Valpuesta, Sanz-Díaz, & Yñiguez, 2016). When clickers are integrated into the classroom, students tend to feel more excited, engaged, and less anxious in learning (Lee et al., 2019). While half of the top 100 highly cited articles on mobile learning published from 2012 to 2016 were focused on primary and secondary education (Lai, 2020), the overall quantitative impact on learning outcomes and the factors that play a central role in promoting learning have received relatively little research attention.

To quantify the overall effects of mobile technology usage on cognitive and non-cognitive learning outcomes and close the research gap related to primary and secondary student learning, we employed a meta-analysis to compare mobile learning effects with traditional learning in primary and secondary education. The present study has two aims. First, we aimed to examine the overall effects of mobile technology usage on multidimensional learning outcomes from three aspects, i.e., cognitive, affective, and behavioral learning. Second, it quantifies and explains the amount of variability in the findings in the literature. Our results from an up-to-date meta-analytic synthesis may provide a rich overview of the current mobile-learning practices and their overall effects, which can inform researchers, policymakers and practitioners on how best to integrate mobile technology in teaching and learning.

5.1.1 Previous narrative reviews of learning with mobile technologies

Narrative reviews regarding mobile learning published over the past three years have been performed in various educational contexts (e.g., Chung, Hwang, & Lai, 2019; Diacopoulos & Crompton, 2020; Lai, 2020; Suarez, Specht, Prinsen, Kalz, & Ternier, 2018). These studies have examined various dimensions of learning outcomes such as Bloom's taxonomy of educational objectives (Chung et al., 2019), thinking skills (Diacopoulos & Crompton, 2020), engagement and collaboration (Diacopoulos & Crompton, 2020), and learners' agency (Suarez et al., 2018). As Lai (2020) stated, these previous seldom-discussed learning outcomes like learners' higher-order thinking and behaviors, are potential mobile learning research themes.

Academics also constrained narrative reviews to school-aged students. Crompton, Burke, and Gregory (2017) conducted a systematic review from 2010 to 2015, investigating the general characteristics of 113 mobile-learning studies conducted in PK-12 (students ages 2-18), such as research purposes, methodologies, and outcomes, domains, educational levels, contexts, and learning activities. In 2019, Crompton and his colleagues (Crompton, Burke, & Lin, 2019) published an up-to-date analysis of students' cognitive learning level as measured by Bloom's Taxonomy in PK-12 mobile learning research. They reviewed 101 articles from 2010 to 2016 and found that mobile devices were integrated into more subjects, e.g., multiple subjects and social studies. Similarly, Crompton and Burke (2020) applied the Substitution, Augmentation, Modification, and Redefinition (SAMR) framework to examine PK-12 studies from 2014 to 2019. They found that mobile technologies were sometimes used to replicate activities without functional changes. Besides, Burden, Kearney, Schuck, and Hall (2019) systematically reviewed 57 studies from 2010 to 2017 focused on innovative mobile learning practices in K-12 education. However, these studies were limited as papers were identified through either the top journals or database searches, which may not represent all works published on mobile learning. Also, the included studies were often published before 2015 (Crompton et al., 2017), conducted in a special education settings (Crompton et al., 2019; Crompton & Burke; 2020), or lack comparison groups (Crompton et al., 2017), which means they cannot generally reflect the current mainstream practice or makes it challenging to evaluate the interventions.

5.1.2 Previous meta-analyses of effects of mobile technology usage on learning outcomes

Numerous experimental or quasi-experimental studies have been conducted to investigate the effects of mobile technology usage. The findings of these primary studies as listed in Table 5.1 have been synthesized in at least nine meta-analyses. However, most meta-analyses had a limited scope, either to synthesize a single outcome variable (Castillo-Manzano et al., 2016; Cho, Lee, Joo, & Becker, 2018; Fabian, Topping, & Barron, 2016; Tingir, Cavlazoglu, Caliskan, Koklu, & Intepe-Tingir, 2017; Yang, Sung, & Chang, 2020), or to center on specific subjects (Castillo-Manzano et al., 2016; Cho et al., 2018; Mahdi, 2018; Tingir et al., 2017), or particular mobile devices (Castillo-Manzano et al., 2016; Hunsu, Adesope, & Bayly, 2016).

We found three broader meta-analyses aimed at various mobile technology use for potential benefits of cognitive and non-cognitive learning in all grades and disciplines in the past five years. More specifically, Sung, Chang, and Liu (2016) investigated the effects of integrating mobile devices on learning that cut across all levels of learning stages, school subjects, and mobile technology types, from 1993 to 2013. They found a significant medium average effect size of $g = 0.523$ for learning achievement and $g = 0.433$ for affective outcome variables (e.g., motivation, attitude, participation, and engagement), compassing 110 journal articles and 18749 participants. The authors answered core questions

Table 5.1. Nine meta-analyses of mobile learning research over the last five years, ordered by year of publication.

Meta-analyses (Year)	Inclusive dates	K	Average ES	Education levels	Subjects	Devices	Learning outcomes
Castillo-Manzano, Castro-Nuño, López-Valpuesta, Sanz-Díaz, and Yñiguez (2016)	2008-2012	53	0.48	University, non-university	Various	Audience-response devices	Examination scores
Fabian, Topping, and Barron (2016)	2003-2012	14	0.48	Elementary education	Mathematics	Various	Student scores
Hunsu, Adesope, and Bayly (2016)	2001-2014	86 for cognitive; 25 for non-cognitive	0.05 for cognitive; 0.23 for non-cognitive	K-12, college	Various	Clicker-based technologies	Cognitive and non-cognitive learning outcomes (i.e., Attitudinal and behavioral learning outcome)
Sung, Chang, and Liu (2016)	1993-2013	108 articles for cognitive; 22 for affective	0.523 for cognitive; 0.433 for affective	Kindergarten, elementary school, middle school, high school, college, adults	Various	Various	Cognitive and affective learning outcomes
Tingir, Cavlazoglu, Caliskan, Koklu, and Intepe-Tingir (2017)	2010-2014	14	0.48	K-12	Science, mathematics and reading	Various	Achievement score
Sung, Yang, and Lee (2017)	2000-2015	163	0.516	Kindergarten, elementary school, junior high school, senior high school, college, graduate school, and adults	Various	Various	Learning achievement, learning attitude, and peer interaction for collaborative learning
Cho, Lee, Joo, and Becker (2018)	2008-2017	22	0.51	Primary, secondary, and post-secondary	Language	Various	Tests
Mahdi (2018)	2009-2015	16	0.67	Young and adult learners	Language	Various	Vocabulary tests
Yang, Sung, and Chang (2020)	2001-2017	87	0.803	Various	Various	Various	Cognitive and affective learning outcome for inquiry-based Learning

about cognitive learning outcomes. For example, do students learn academic content better with mobile technology than conventional technology (Mayer, 2020). The authors concluded that students using mobile devices in education performed better than those not. Besides, unlike other reviews, Sung and his colleagues focused on different teaching methods rather than mobile learning in general, for example, inquiry-based learning (Sung, Yang, & Lee, 2017) and collaborative learning (Yang et al., 2020).

Although the above mentioned meta-analyses have added academic understanding to the effects of mobile technology usage, they did not distinguish between affective and behavioral learning outcomes from non-cognitive outcomes, nor consider conducting moderator analyses related to these non-cognitive outcome variables. Moreover, it is hard to determine what happens to primary and secondary students and see how mobile devices boost their learning in various ways. To address these concerns, the current study took a step further by investigating the effects of mobile technology usage on different learning outcomes emphasizing primary and secondary education. In sum, this study differs from previous studies for the following reasons. First, an addition from 2014 on is necessary because of the large number of studies. Secondly, the current study is not limited to cognitive learning outcomes but also includes non-cognitive learning outcomes. We examined the effects of mobile technology usage on three dimensions of outcomes: cognitive learning, affective learning, and behavioral learning. Third, we considered a series of factors from both educational and methodological aspects, which are supposed to moderate the effectiveness of the mobile technology intervention.

5.1.3 Potential moderator variables considered

In addition to uncertainty about the overall effects of mobile technology on different types of the outcome variable, the potential influences of several moderators need further exploration, which were derived from relevant studies conducted earlier. We adopted the 3P (presage - process - product) model (Biggs, 2003) to determine the primary aspects of moderators that could reflect the full picture of teaching and learning within the mobile technology integration context. The 3P model provides us to comprehend the relationships among student and teaching context presage factors, learning process factors, and product factors (learning outcomes) within the context of mobile technology usage. Moreover, higher methodological quality studies could have provided substantially different results than less quality studies (Cheung & Slavin, 2016), thus we chose to study quality factors commonly presented in experimental studies on mobile technology intervention. Therefore, the potential moderators in this meta-analysis have been grouped into four categories: student factors, teaching context, learning process, and study quality.

The following two moderator variables have been considered as student factors: community type and student socioeconomic status (SES). Moreover, teaching context factors include education level, school type, learning environment, school subjects, teacher training on content and technology. Hardware used, student-to-hardware ratio, software used, teaching method and duration of intervention are selected as the learning process factors. Finally, we examine whether the different results between the studies could be explained by research design, instructor equivalence, degree of technology use in the control group, and the procedure of effect size extraction. Although researchers have constantly discussed the significance of the above variables (see e.g., Chauhan, 2017; Schmid et al., 2014; Sung et al., 2016; Zheng, Warschauer, Lin, & Chang, 2016), at this point, we give our special attention to learning process factors which might provide a deeper insight in the implementation and evaluation of the interventions of interest. Below we go into details on our rationale for the selected moderator variables related to the learning process.

In order to guide the decision on instructional designs and keep advancing mobile learning in all different situations, the learning process factors can typically be described by three main aspects: human resources, technological resources and intervention duration. Human resources primarily refer to teachers, especially the type of pedagogy they adopted that supports students to acquire knowledge and their interaction processes, and technological resources primarily relate to the degree of resource access and differences in resource usage that supports educational processes. Intervention duration refers to the duration between time prior intervention and time post intervention.

Regarding technological resources access, the effects of technology on learning retention and joyful learning environment are most likely when each student had access to an individual digital device (Chou, Chang, & Lin, 2017). However, Kay, Benzimra, and Li (2017) found that students were distracted more when using mobile devices on their own. Next, the most common variables with regard to the difference of resource usage are hardware and software used for learning. As an example, Sung et al (2016) found that the effect sizes differed significantly among the various hardware including handheld, laptops and mixed devices, and larger effects were reported for learning-oriented software designed for educational purposes than for general software designed for commercial purposes. More importantly, we believe that the value of mobile technology lays in how it is integrated with pedagogy and curriculum. Several meta-analyses (see e.g., Sung et al., 2017; Yang et al., 2020) have shown that different teaching methods implemented in mobile learning context produce different effects. Furthermore, we included duration of intervention as moderator variable. Empirically, short interventions might not yield effects because students need some time to familiarize with hardware and software (Sung et al., 2017). Nevertheless, if the intervention duration is too long, the effects could decline because students feel less motivated (Lee et al., 2019).

The purpose of this study is to provide new quantitative data that are expected to deepen the knowledge base on various learning outcomes and inform evidence-based decision-making on the use of mobile technology in primary and secondary education. Following the PICO framework, the population is composed of students in primary and secondary education. The intervention is the use of mobile technology for learning. The comparison is made with a non-technology (e.g., pen and paper) or traditional technology group (e.g., desktop computers and whiteboards). The outcomes refer to measurements of cognitive (e.g., attention, memory, and understanding), affective (e.g., motivation, emotions, and attitudes), and behavioral (e.g., self-efficacy, interaction, and engagement) aspects of learning. Specifically, this meta-analysis seeks to answer the following research questions:

RQ1: When compared with traditional learning, what is the overall effectiveness of using mobile technologies in primary and secondary education on students' learning outcomes in terms of cognitive, affective, and behavioral dimensions?

RQ2: What, if any, factors based on 3P model, that is student factors, teaching context and learning process factors, moderate the relationship between mobile technology use and learning outcomes?

RQ3: For RQ1 above, what, if any, study quality characteristics explain the heterogeneity in results?

5.2 Method

5.2.1 Inclusion and exclusion criteria

Our criteria for the determination of coding studies and subsequent meta-analysis were developed based on a preliminary literature review on the use of mobile technology for educational purposes. A pre-defined criterion for identifying research samples was listed below:

(a) The study used an experimental or quasi-experimental research design.
(b) The results of the mobile technology intervention group were compared with non-technology (e.g., pen and paper) or traditional technology (e.g., desktop computers and whiteboards) groups.
(c) Learning outcomes were reported as the dependent variable, measured by either cognitive, affective, or behavioral learning outcomes.
(d) Reported original data and provided sufficient information to calculate effect sizes, such as means, standard deviations, the sample size in each group.
(e) The sample consisted of primary or secondary school students.
(f) Studies were published in peer-reviewed journals, and a full text was available.
(g) Studies were published between 2014 and 2020 and were written in English. The starting year was set in 2014 because we extended Sung et al.'s (2016) study to understand the mobile learning empirical field over recent years.

Several exclusion criteria were used. Conceptual analysis or research reviews, and qualitative research, pre-experimental studies, and editorials were excluded. Moreover, studies on gifted education, special education, or disabilities learning were excluded. Studies involving any children with special educational needs were also excluded because this may have potential impacts on the entire group's performance. In cases where studies met all the inclusion criteria but lacked sufficient descriptive statistics or inferential statistics to calculate effect sizes were excluded.

5.2.2 Literature search and data sources

Studies were identified from two different sources. First, a database search was performed on all databases available at the library of Leiden University, such as Web of Science, Elsevier, ERIC, SAGE journals. Four sets of keywords were combined: (1) population (i.e., student); (2) mobile-technology related terms (i.e., mobile technology, mobile device, personal digital assistant, handheld, iPad, laptop, tablet, smart phone, mobile phone, response system); (3) learning-related keywords (i.e., learning outcome, achievement, performance); and (4) research-design related keywords (i.e., experimental, quasi-experimental). For the search, a Boolean OR operator first linked the keywords within each set; a Boolean AND operator was used to combine keywords across the four sets. The terms of mobile technology were searched within titles, and other terms were searched within any field. 421 peer-reviewed articles were found on 25 May 2020, and twenty duplicate papers were then removed in the Mendeley. In the next step, the title and abstract of each paper were read. Based on our criteria, the first author assessed these 401 studies to determine 'yes', 'maybe', or 'no' (Liberati et al., 2009), and papers in the 'maybe' group were then assigned to other two authors for the final decision. A total of 39 eligible papers were obtained in this stage.

Moreover, we browsed the major educational technology and mobile learning journal online in June 2020, including the British Journal of Educational Technology, Computers & Education, Educational Technology Research and Development, Educational Technology & Society, Journal of Computer Assisted

Learning. After removing the 41 duplicates from the 3318 paper contained in the five journals, additional 196 studies were found after screening abstracts, resulting in 235 articles for full-text review. These articles were not found in the first stage and the main reason is that the terms of mobile technologies were searched within titles and these studies used other related terms (e.g., games, mobile learning, mobile application, online tools, and clickers).

During the final full-text screening step, at least two authors screened the articles applying the inclusion and exclusion criteria to check for eligibility. There were minor disagreements mostly related to whether mobile technologies were used, and these were discussed among the three authors until they were resolved. This step limited these studies to the 61 studies of 56 journal articles that were included in this meta-analysis. Figure 5.1 provides a flowchart describing the inclusion process and describes the reasons why studies were excluded, following the guidance of The PRISMA Group (Moher, Liberati, Tetzlaff, & Altman, 2009).

5.2.3 Coding of potential moderators

First, a coding sheet was developed mainly based on the coding variables in recent meta-analysis articles (Schmid et al., 2014; Sung et al., 2016). Evidence produced by review, however, was used to assess relationships that primary researchers never examined (Cooper & Hedges, 1994). Thus, a strategy we used to adapt the original coding sheet was to search for possible moderators by evaluating a subset of studies (Brown, Upchurch, & Acton, 2003). After the pilot testing on 22 articles, four variables (i.e., student training on technology, student training on content, learning topic/ content equivalence, and software/ tool equivalence) were added to the coding sheet. After completing the code sheet, a codebook was developed to guide the coding process for all eligible studies. All the eligible studies were independently coded by the first and the second or third author. All disagreements produced by the former procedure were addressed in several meetings, and the authors reached consensus on each coding category.

In total, we coded for 21 variables (17 from previous studies and 4 from our new data) that were supposed to be used as moderators. However, not all were included in the moderator analyses. We excluded 6 moderators either because of low variability in the outcome (i.e., school type and software used), or because very few studies reported the relevant information (i.e., student and teacher training on technology/ content). In the end, 15 variables served as moderators (see Table 5.3 for the final moderators and their categories).

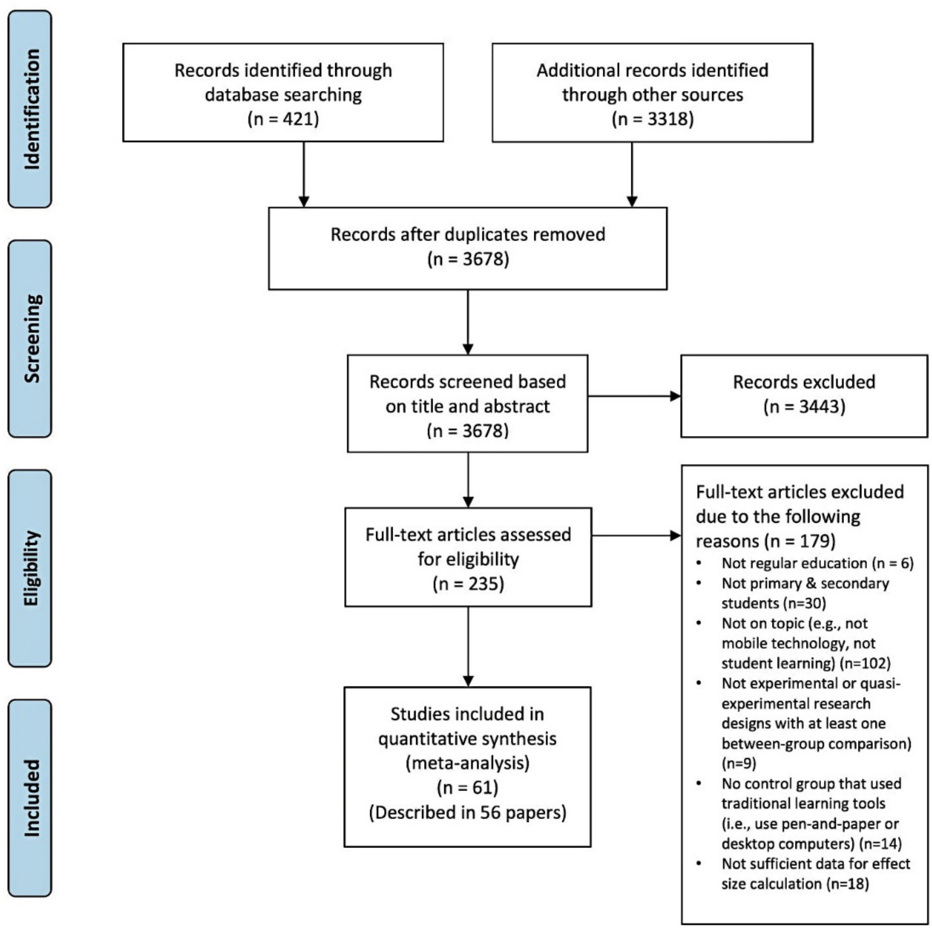

Figure 5.1. Flowchart of the study selection process following the guidelines of The PRISMA Group (Moher et al., 2009).

5.2.4 Effect size calculation

In the present meta-analysis, the standardized mean difference between the intervention and the control conditions on the posttest was the dependent variable. We chose the effect size of Hedges' *g* over Cohen' s *d* because it is more accurate for smaller samples (Borenstein, Hedges, Higgins, & Rothstein, 2009). The intervention group outperformed the control group by showing a positive effect size. Cohen (1992) indicated that the value of any pooled Hedges' *g* was viewed as following: small effect ($g = 0.2$), medium effect ($g = 0.5$), and large effect ($g = 0.8$).

Wherever applicable, the effect sizes were calculated based on the post-baseline means and standard deviations rather than scores reflecting changes from baseline to follow-up, as these are not independent (Cuijpers, Weitz, Cristea, & Twisk, 2017). If they were not available, we used other inferential statistics as long as they represent the difference between the intervention and the control condition on the posttest.

The cognitive learning outcome was the primary outcome and we also coded effect sizes based on affective and behavioral learning outcomes. When more than one appropriate outcome measure was reported in a study, we calculated effect sizes for all of those. The software Comprehensive Meta-Analysis (CMA), Version 3.3.070 was used to calculate the effect size for each contrast.

5

5.2.5 Statistical dependence of the samples

We included ten studies with multiple comparisons. Since these comparisons are not independent of each other this may yield an artificial reduction of heterogeneity which can affect the pooled effect size, we examined these possible effects by conducting sensitivity analyses in which we included only one of the comparisons per study. However, this did not result in a different result (for more details, see section 3.3). The second case of dependent data was reporting multiple outcomes or time-points per study. A study may involve different measures for the same learning outcome variable. In this case, we created a synthetic effect size for each study, which is a more conservative method for combining dependent outcomes than assuming completely independent outcomes (see Borenstein et al., 2009). When multiple time points of one dependent variable in one study could be calculated, we chose only to include the measurement that is closest to the end of the intervention that causes differences between experimental and control groups to rule out other possible explanations. Additionally, for those studies providing two or more independent experiments, and each experiment contributing independent information, we treated each experiment as a separate study, computed the effect within experiments, and then use these effects as the unit of analysis.

5.2.6 Data analysis

We conducted three meta-analyses: one on the cognitive learning outcome, one on the affective learning outcome, and one on behavioral learning outcome. Because there was a wide range of different participants, interventions and outcome measures between studies, we used the random-effects model to calculate the average effect sizes. The random-effects model allows for between-study variance beyond random error (Borenstein et al., 2009).

The first method to examine heterogeneity is to look carefully at the forest plot. Forest plots were presented to examine effect size distributions, and to assist in identifying outliers. Outliers were defined as studies in which the 95% CI was outside the 95% CI of the pooled studies and excluding outliers from a meta-analysis results in a considerable drop in the level of heterogeneity (Levy Berg, Sandell, & Sandahl, 2009). However, outlier tests are tools that help us to find certain studies that are worth examining in more detail but should not be taken as a justification of removal studies (Viechtbauer & Cheung, 2010).

Additionally, the *Q*-statistics was utilized to calculate the heterogeneity of the average effect sizes. As an indicator of heterogeneity, we calculated the I^2-statistic, which gives heterogeneity in percentages and it is assumed that a percentage of 25% indicates low heterogeneity, 50% moderate and 75% high heterogeneity (Higgins, Thompson, Deeks, & Altman, 2003).

In order to assess the effects of differences between the primary studies that might have an influence on the results we tested the effects of a priori defined variables. Moderator analyses were conducted to compare the contrasts based on categorical moderator variables in all the meta-analyses. Only categorical moderator variables that had at least four contrasts in the categories were used (Bakermans-Kranenburg, Van IJzendoorn, & Juffer, 2003). Because very few studies were found in some categories, we merged these categories. For example, we assumed that the SES of students was not low if it was not reported in the study.

Publication bias was inspected in all sets of studies. Studies with significant results are more likely to be published and thus significant findings may be overrepresented in a meta-analysis which may lead to an overestimation of the average effect size. The visual display of effect sizes against standard errors by a funnel plot is a popular way to evaluate publication bias and an asymmetrical distribution of the studies indicates the risk of missing studies (Card, 2012). In case of an asymmetrical funnel plot, we used Duval and Tweedie's trim and fill procedure to calculate the adjusted effect (Duval & Tweedie, 2000). Furthermore, Rosenthal's fail-safe N was estimated to show the number of missing studies ($5k + 10$) with zero effect to be required to generate non-significant results (Rosenthal, 1979).

5.3 Results

5.3.1 Characteristics of included studies

The final dataset consisted of 61 studies from 56 articles with a total of $N = 6406$ students. Appendix F presents the studies included in the present meta-analysis and Appendix G provides an overview of the studies. The most studied region was Taiwan (n = 26). Community types (i.e., urban, suburban and rural) were only reported in 23% of the studies. In a few studies (n = 5), students came to school with a low SES. More than half of the studies (n = 33) investigated primary school students and less than half of the studies (n = 28) investigated students from the secondary school level. For learning environment, 40 studies implemented in the formal settings. Language arts were the most studied subjects (n = 20), followed by Science (n = 18), Social studies (n =10) and Mathematics (n = 9). Handheld devices with multiple functions (including laptops, tablet PCs, and mobile phones) were the most widely studied hardware (n = 53), followed by handheld devices with one specific function (n = 5, including classroom response systems, e-book readers, PDAs, digital pen, etc.). In about half of the studies (n = 36), students owned and used a mobile device. With regard to teaching method, inquiry-oriented leaning (n = 19, including discovery and exploration, problem-solving, project-based learning, and cooperative learning) was the most frequently researched, followed by game-based learning (n = 11). The studied intervention duration were similar, that is, < 1 day (n = 18), 1 day- 4 weeks (n = 20), and > 4 weeks (n = 19). Only 9 studies utilized a true experimental design. Some studies conducted well on equivalent instructor (n = 28), equivalent learning topic/ content (n = 50), and equivalent software/ tool (n = 35). Finally, pen-and-paper conditions (n = 40) were the most often studied control groups, followed by traditional technology condition (n = 12).

5.3.2 Evaluation of publication bias

Regarding the possibility of publication bias affecting our data, funnel plots for each dependent variable were examined for asymmetry, as presented in Figures 5.2, 5.3, and 4. Duval and Tweedie's trim and fill analyses showed that no studies missing for cognitive and affective learning outcomes, and that 1 extra study for behavioral outcome variable had to be imputed to obtain a symmetric distribution of effects. The adjusted mean effect size on behavioral outcome was still positive, but showed a smaller (and significant) effect of using mobile devices for learning (g = 0.477, 95% CI [0.164, 0.789]). Finally, the fail-safe N was 6508, 607, and 262, with cognitive, affective and behavioral learning outcomes, respectively, which is much larger than the tolerable number of studies with 370, 130, and 80, respectively. Based on these analyses, we concluded that the effects of mobile technology usage on learning in primary and secondary education was reliable and robust.

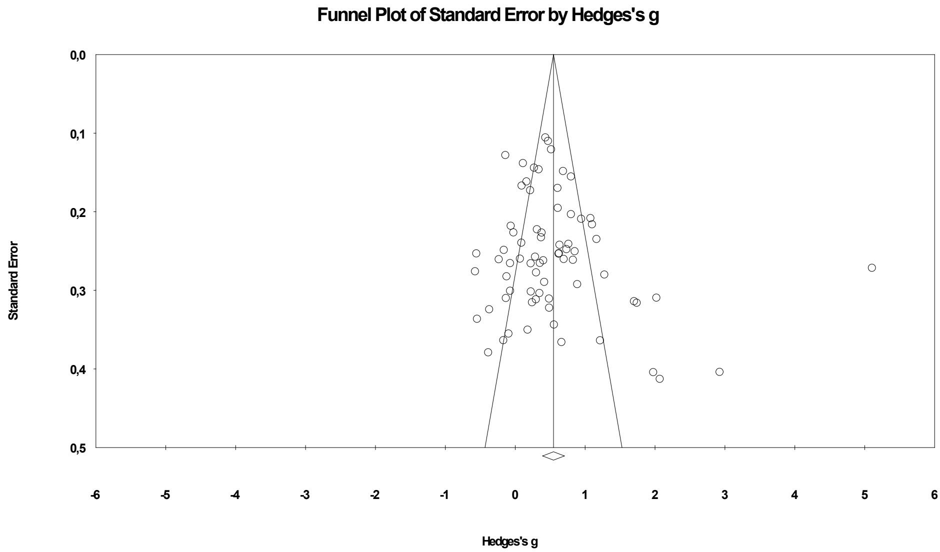

Figure 5.2. Funnel plot of the 72 effect sizes for cognitive outcomes.

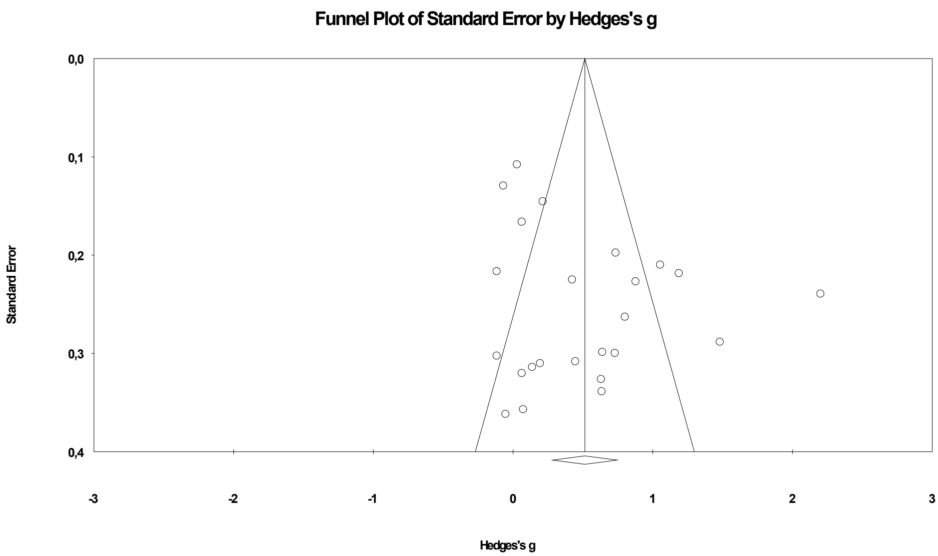

Figure 5.3. Funnel plot of the 24 effect sizes for affective outcomes.

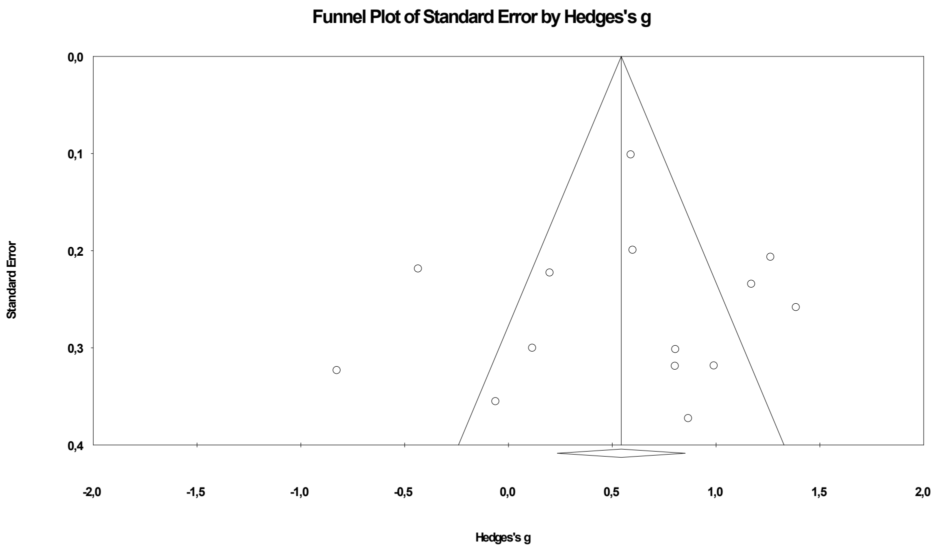

Figure 5.4. Funnel plot of the 14 effect sizes for behavioral outcomes.

5.3.3 Overall effects of mobile technology usage compared with control groups

The first research question focused on the advantages of using mobile technologies on student learning outcomes correspondingly in comparison to students learning without mobile technologies. We could compare the effects of mobile technologies with control groups on learning outcome in 72 cognitive comparisons from 59 studies, in 24 affective comparisons from 22 studies, and in 14 behavioral comparisons from 13 studies. Within each study set, effect sizes and 95% confidence intervals of each study are presented in Figures 5.5, 5.6 and 5.7.

With regard to the primary outcome variable, the overall effect shows that the use of mobile technologies had a medium positive and significant effect on cognitive learning (g = 0.547, 95% CI [0.392, 0.703]). Similar to the effects on cognitive learning, the combined effect on affective learning was medium (g = 0.514, 95% CI [0.282, 0.745]). For behavioral learning outcomes, a medium positive and significant effect size (g = 0.543, 95% CI [0.235, 0.851]) was also found. Heterogeneity is large (I^2 = 88.694 for the cognitive dimension, I^2 = 84.618 for the affective dimension, I^2 = 83.595 for the behavioral dimension)

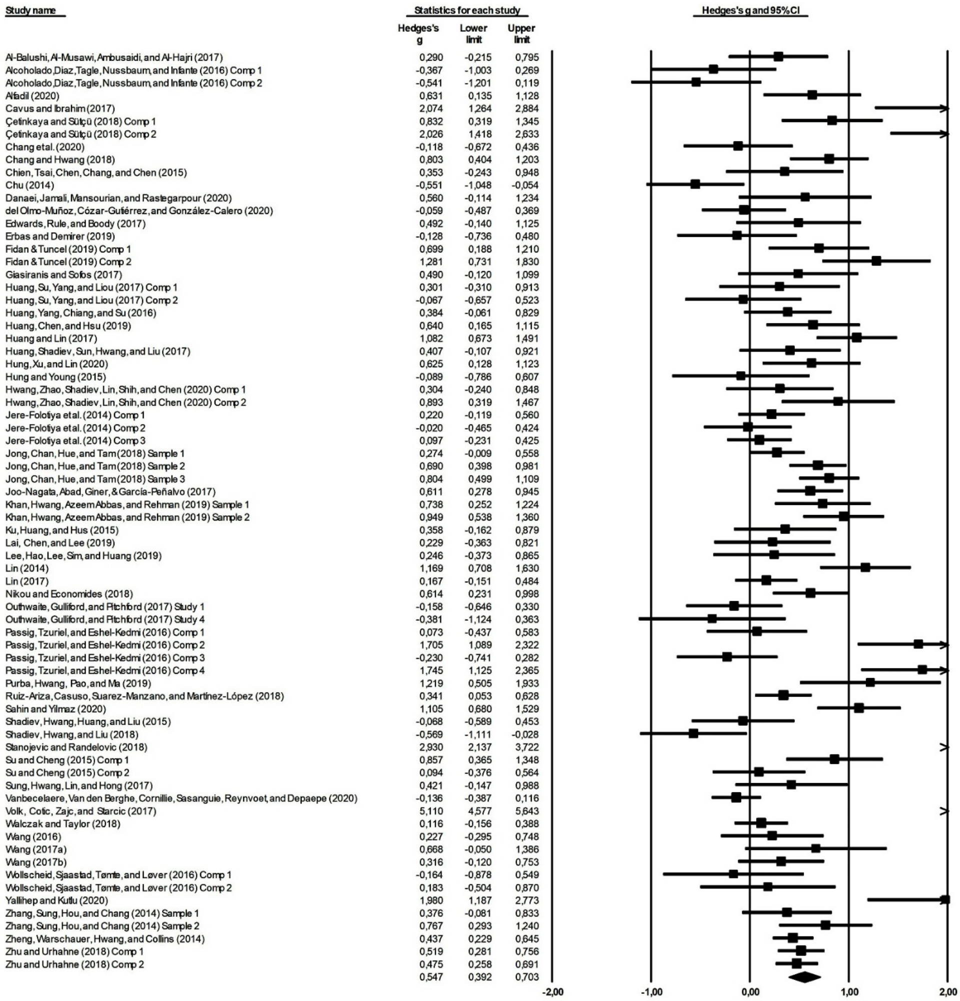

Figure 5.5. Forest plot of the 72 effect sizes for cognitive outcomes. Within one article, when multiple sample or studies were presented, the figure reports the result of each sample (sample 1, sample 2, etc.) or study (study 1, study 2, etc.) separately. Similarly, when studies used multiple comparisons, the figure reports the result of each comparison (comp 1, comp 2, etc.) separately.

5

Study name	Hedges's g	Lower limit	Upper limit
Chang and Hwang (2018)	2,203	1,733	2,672
del Olmo-Muñoz, Cózar-Gutiérrez, and González-Calero (2020)	-0,115	-0,540	0,310
Erbas and Demirer (2019)	0,139	-0,477	0,754
Fidan and Tuncel (2019) Comp 1	0,803	0,287	1,318
Fidan and Tuncel (2019) Comp 2	1,483	0,917	2,048
Furió, Juan, Seguí, and Vivó (2015)	0,632	-0,008	1,272
Giasiranis and Sofos (2017)	0,448	-0,156	1,052
Huang, Su, Yang, and Liou (2017) Comp 1	0,195	-0,413	0,804
Huang, Su, Yang, and Liou (2017) Comp 2	-0,114	-0,708	0,479
Huang, Yang, Chiang, and Su (2016)	0,425	-0,017	0,866
Huang and Lin (2017)	1,056	0,644	1,467
Hung and Young (2015)	0,074	-0,626	0,774
Joo-Nagata, Abad, Giner, and García-Peñalvo (2017)	0,065	-0,261	0,391
Lai, Chen, and Lee (2019)	0,731	0,143	1,319
Lee, Hao, Lee, Sim, and Huang (2019)	0,065	-0,563	0,693
Lin (2014)	0,879	0,434	1,325
Nikou and Economides (2018)	0,736	0,349	1,124
Quintas, Bustamante, Pradas, and Castellar (2020)	0,031	-0,181	0,242
Ruiz-Ariza, Casuso, Suarez-Manzano, and Martínez-López (2018)	0,215	-0,070	0,501
Sahin and Yilmaz (2020)	1,189	0,760	1,618
Sung, Hwang, Lin, and Hong (2017)	0,641	0,055	1,227
Vanbecelaere, Van den Berghe, Cornillie, Sasanguie, Reynvoet, and Depaepe (2020)	-0,068	-0,322	0,186
Wang (2017a)	-0,052	-0,761	0,658
Yallihep and Kutlu (2020)	0,637	-0,028	1,301
	0,514	0,282	0,745

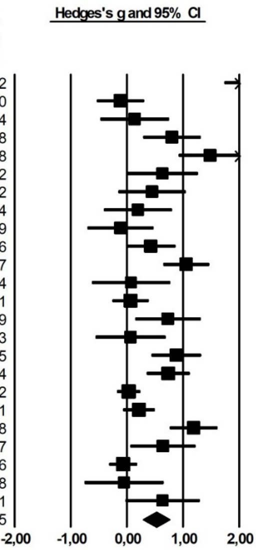

Figure 5.6. Forest plot of the 24 effect sizes for affective outcomes.

Study name	Hedges's g	Lower limit	Upper limit
Chang and Hwang (2018)	1,264	0,858	1,669
Chien, Tsai, Chen, Chang, and Chen (2015)	0,991	0,366	1,615
del Olmo-Muñoz, Cózar-Gutiérrez, and González-Calero (2020)	-0,434	-0,863	-0,005
Erbas and Demirer (2019)	-0,825	-1,459	-0,192
Huang, Su, Yang, and Liou (2017) Comp 1	0,804	0,178	1,429
Huang, Su, Yang, and Liou (2017) Comp 2	0,116	-0,472	0,705
Huang, Yang, Chiang, and Su (2016)	0,200	-0,237	0,637
Huang and Lin (2017)	0,600	0,209	0,991
Hung and Young (2015)	0,867	0,137	1,598
Lai, Chen, and Lee (2019)	0,805	0,214	1,397
Lin (2014)	1,171	0,712	1,631
Quintas, Bustamante, Pradas, and Castellar (2020)	0,591	0,392	0,789
Wang (2017a)	-0,061	-0,757	0,636
Zhang, Sung, Hou, and Chang (2014) Sample 2	1,386	0,880	1,893
	0,543	0,235	0,851

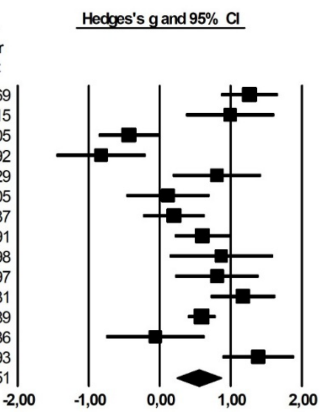

Figure 5.7. Forest plot of the 14 effect sizes for behavioral outcomes.

for the effects on all three learning outcome dimensions and highly significant ($p < 0.001$) in these analyses.

Ten studies were special since they included multiple comparisons. We examined the possible effects of this by conducting analyses with only one effect size (either the largest or the smallest effect size) per study. As Table 5.2 reveals, the resulting effect sizes were roughly the same as in the overall analyses. Heterogeneity test was not significant for cognitive ($I^2 = 9.894$, $p = 0.344$), affective ($I^2 = 0$, $p = 0.826$), and behavioral ($I^2 = 0$, $p = 0.972$) learning outcome, indicating the observed differences might not be important.

5.3.4 Moderator analyses

To answer RQ2 and RQ3, we performed moderator analyses. We calculated effect sizes and 95% CI for each level with at least four studies of all potential moderators. Results for cognitive learning outcomes are presented in Table 5.3, affective and behavioral learning outcome are presented in Table C.1 and Table C.2 respectively in Appendix H, along with all between group heterogeneity tests.

For cognitive learning outcomes, as can be seen in Table 5.3, of all 15 variables tested, 5 moderators were found. We found indications that low SES students had lower ESs than others ($p = 0.001$), that students using handheld device with multiple functions were significantly more effective than using device with one single function ($p = 0.031$), that equivalent learning topic/ content between comparison groups resulted in a higher ESs ($p < 0.001$), that each student having one mobile device was significantly associated with the higher ESs ($p = 0.01$), and that the ESs differed significantly between the two effect size extraction procedures ($p = 0.041$).

In the moderator analyses for affective learning outcomes (see Table H.1 in Appendix H), we only found studies in which the equivalent learning topic/ content resulted in a higher differential effect size than studies in which non-equivalent learning content/ topic were applied ($p = 0.017$). In the series of moderator analyses regarding behavioral learning effects, results in Table H.2 in Appendix H showed that the effects size was only significantly associated with software/ tool equivalence ($p = 0.020$).

Table 5.2. Overall effect sizes of mobile technology usage.

Dependent variable	Effect size and 95% confidence interval				Heterogeneity			
	N	*g*	SE	95% CI	Q(p)	df(Q)	τ^2 (SE)	I^2 (%)
Cognitive learning outcome								
All studies	72	0.547	0.080	[0.392, 0.703]	628.013 (< 0.001)	71	0.386 (0.090)	88.694
Possible outliers removed	53	0.446	0.045	[0.358, 0.533]	107.356 (< 0.001)	52	0.049 (0.020)	51.563
One effect size per study (largest)	59	0.613	0.092	[0.432, 0.794]	570.021 (< 0.001)	58	0.434 (0.110)	89.825
One effect size per study (smallest)	59	0.520	0.090	[0.343,0.697]	548.781 (< 0.001)	58	0.412 (0.105)	89.431
Affective learning outcome								
All studies	24	0.514	0.118	[0.282, 0.745]	149.523 (< 0.001)	23	0.268 (0.113)	84.618
Possible outliers removed	19	0.405	0.089	[0.230, 0.579]	41.112 (0.001)	18	0.079 (0.050)	56.218
One effect size per study (largest)	22	0.527	0.125	[0.282, 0.772]	144.294 (< 0.001)	21	0.277 (0.122)	85.446
One effect size per study (smallest)	22	0.485	0.120	[0.249, 0.721]	134.777 (< 0.001)	21	0.254 (0.112)	84.419
Behavioral learning outcome								
All studies	14	0.543	0.157	[0.235, 0.851]	79.244 (< 0.001)	13	0.274 (0.158)	83.595
Possible outliers removed	10	0.610	0.111	[0.393, 0.827]	18.099 (0.034)	9	0.055(0.056)	50.273
One effect size per study (largest)	13	0.574	0.165	[0.251, 0.897]	76.815 (< 0.001)	12	0.282 (0.167)	84.378
One effect size per study (smallest)	13	0.524	0.166	[0.199, 0.850]	78.713 (< 0.001)	12	0.289 (0.170)	84.755

5.4 Discussion

5.4.1 Overall effects on learning outcomes

We conducted a systematic review with a meta-analysis of experimental and quasi-experimental studies comparing the effects of learning with and without mobile technology. Compared with traditional technology and non-technology groups, mobile technology produced medium positive and statistically significant effects on primary and secondary students' learning in terms of cognitive, affective, and behavioral learning outcomes. The current meta-analysis provides the converging 'best evidence' for the overall beneficial effects of using mobile technology in education.

5.4.2 Moderator variables

The main effects of mobile technology mentioned above are not the same for all student groups and learning contexts. Therefore, moderator analyses have been performed with characteristics of the students and learning contexts as moderators. The results from a series of moderator analyses supported the importance of some variables from three categories, i.e., student factors, learning process, and study quality, that explained differences in learning outcomes between mobile learning and traditional learning. From an educational perspective - as indicated in the 3P model -, effect sizes varied significantly for cognitive learning outcomes according to SES, hardware used, ratio. The mobile technology interventions were more beneficial for students using handheld devices with multiple functions, and using mobile devices on their own, except for students with low socioeconomic status (SES) backgrounds. Moreover, the effect of community type was on the edge of significance for cognitive learning outcomes, p = 0.055, favoring urban communities. The effect of teaching method was on the edge of significance for affective learning outcomes, p = 0.052, favoring inquiry-oriented learning. Nevertheless, because the number of included studies was small, these effects must be interpreted with caution. Furthermore, the four factors in the teaching context category (education level, school type, learning environment, and school subjects) were not significant moderators for all learning outcomes. From the methodology perspective, the

5

Table 5.3. Moderator analyses and weighted mean effect sizes for cognitive outcome variables.

Moderator category	Moderator variables	N	*g*	SE	95% CI	Q_B	*p*
Student factor	**Community type**						
	Urban	11	0.647	0.150	[0.352, 0.942]	3.695	0.055
	Not urban	6	0.271	0.125	[0.027, 0.516]		
	SES						
	Low	6	0.031	0.152	[-0.267, 0.329]	10.638	0.001
	Not low	66	0.601	0.086	[0.432, 0.770]		
Teaching context	**Education level**						
	Primary school	41	0.483	0.125	[0.238, 0.728]	0.771	0.380
	Secondary school	31	0.618	0.090	[0.442, 0.794]		
	Learning environment						
	Formal settings	48	0.495	0.114	[0.272, 0.719]	1.656	0.437
	Informal settings	17	0.560	0.066	[0.431, 0.689]		
	Unrestricted	7	0.921	0.312	[0.310, 1.531]		
	School subject						
	Language arts	23	0.466	0.112	[0.247, 0.685]	2.829	0.587
	Social studies	10	0.420	0.120	[0.185, 0.656]		
	Mathematics	14	0.623	0.330	[-0.023, 1.270]		
	Science	19	0.579	0.088	[0.407, 0.752]		
	Professional subjects	4	1.305	0.674	[-0.015, 2.625]		
Learning process	**Hardware used in intervention group**						
	Handheld devices with multiple functions	61	0.582	0.094	[0.399, 0.765]	4.675	0.031
	Handheld devices with one specific function	8	0.247	0.124	[0.005, 0.489]		
	Student-to-hardware ratio						
	Own	43	0.645	0.121	[0.407, 0.883]	6.613	0.010
	Shared	5	0.144	0.152	[-0.154, 0.442]		
	Teaching method						
	Inquiry-oriented learning	22	0.537	0.084	[0.372, 0.701]	0.653	0.884
	Game-based learning	11	0.439	0.136	[0.173, 0.704]		
	Self-directed learning	9	0.655	0.271	[0.125, 1.186]		
	Computer-assisted testing/ assessment	14	0.508	0.173	[0.170, 0.846]		
	Duration of the intervention						
	< 1 day	21	0.475	0.107	[0.264, 0.685]	0.582	0.747
	1 day-4 weeks	23	0.596	0.124	[0.353, 0.839]		
	> 4 weeks	25	0.561	0.173	[0.221, 0.900]		

Study quality	**Research design**						
	Quasi-experimental	61	0.534	0.087	[0.363, 0.705]	0.161	0.688
	Experimental	11	0.624	0.206	[0.220, 1.027]		
	Instructor equivalence						
	Same	35	0.577	0.118	[0.347, 0.808]	0.032	0.857
	Different	27	0.545	0.136	[0.279, 0.811]		
	Learning topic/ content equivalence						
	Same	57	0.652	0.099	[0.458, 0.846]	14.124	< 0.001
	Different	7	0.188	0.073	[0.045, 0.332]		
	Software/ tool equivalence						
	Same	38	0.603	0.138	[0.333, 0.873]	0.068	0.795
	Different	16	0.551	0.143	[0.272, 0.831]		
	Degree of technology use in the control group						
	Pen-and-paper	47	0.579	0.119	[0.347, 0.812]	0.255	0.614
	Traditional technology	13	0.495	0.118	[0.264, 0.726]		
	Procedure of effect size extraction						
	Calculated from exact descriptive	63	0.580	0.092	[0.399, 0.761]	4.196	0.041
	Calculated from inferential statistics	8	0.321	0.007	[0.152, 0.490]		

results on cognitive learning outcomes identified two moderators (i.e., learning topic/ content equivalence, and procedure of effect size extraction), on affective learning outcomes identified one moderator (i.e., learning topic/ content equivalence), and on behavioral learning outcomes identified one moderator (i.e., tool/ software equivalence).

Although previous research has indicated the influence of socioeconomic status on education equality among children (Li & Ranieri, 2013), previous meta-analyses of mobile technology interventions (see e.g., Tingir et al., 2017) failed to examine this moderator effect due to lacking relevant information. The finding that students with low SES benefited less than their peers is of particular importance in understanding the new digital divide and offering a valuable direction to explore differences amongst subgroups such as ethnicity, migration status, and community types. Furthermore, in line with previous meta-analysis (Sung et al., 2016), handheld devices with multiple functions often induced better cognitive learning outcomes. Handheld devices with diverse functions such as instant-feedback, speech recognition, and peer-assessment enrich learning opportunities and meet students' demands, prompting higher learning achievement. Besides, in contrast to the assumption of Haßler, Major, and Hennessy (2016), the current meta-analysis proved the higher learning gains in a student-device ratio of one-to-one environment than the shared-device learning environment. A possible explanation is that individual student mobile device supported student-centered and individualized learning (Zheng et al., 2016) and enabled teachers or computer systems to provide immediate feedback to individual students (Castillo-Manzano et al., 2016). No significant effects were found in variables in the teaching context category. An important implication of these findings is that mobile technology interventions can have an equally powerful effect on students' learning across teaching contexts. With regard to the research methodology category, the finding that instructor equivalence was not found to be a significant moderator is in accordance with previous meta-analysis on college students' learning outcomes in technology-enabled active learning environments (Shi, Yang, MacLeod, Zhang, & Yang, 2020). The influence of other features of the study quality, such as learning topic/ content equivalence, tool/ software equivalence, and procedure of effect size extraction, have not been investigated as potential moderators in past meta-analyses. However, in this study, these features served as a significant moderator variable for at least one learning outcome dimension. In sum, this calls for future research to consider the features of study quality to explore whether the moderator effects exist and might contribute to the observed differences.

5.4.3 Limitations and future research

Many studies were not included in this meta-analysis because the necessary information was not reported. Out of 235 potentially relevant journal articles found in the databases and journal websites, only 61 studies could finally be used for the analyses. Studies were excluded not only because they lacked statistical data but also because of other missing information that is important for meta-analyses. As stated by Sung, Li, Yang, and Chang (2019), mobile-learning research has suffered from methodological shortcomings that might hinder the ability of mobile-learning research to obtain reliable evidence for sustaining innovative practices and creating valid theories. To this end, Sung, Li, Yang, and Chang (2019) suggest mobile-learning researchers should utilize valid designs for their research tools, procedures, and statistical methods and focus on presenting their research results more clearly by applying the checklist for the Rigor of Education-Experiment Designs (CREED). Owing to the limited number of empirical mobile-learning studies, the quality of experimental research was not used as a criterion for the inclusion or exclusion of research samples, except that these studies were peer-reviewed; instead features of study quality were analyzed as potential moderators. Furthermore, we had few studies examining differential effects on affective and behavioral learning outcomes. We recommend that outcomes beyond cognitive learning outcomes are given more attention in research designs to fully explore the complex array of student outcomes in a learning situation. Other factors, such as training of teachers and students on technology/ content, software used, and school type, could provide more practical and theoretical insights into the effects of using mobile technologies on school students' learning. These variables were not included in the moderator analyses of the present study due to low variability in categories or missing information in the studies. Lastly, because all included studies were written in English, we suggest that future meta-analyses could consider adding more articles written in different languages to yield more robust findings than using an English single language.

5.4.4 Implications for policymakers and practitioners

The findings above may provide insight into the optimal arrangement of mobile learning regarding the presage (e.g., SES), process (e.g., student-to-hardware ratio, hardware used), study quality (e.g., learning content/ topic equivalence, software/ tool equivalence), and product (e.g., cognitive, affective and behavioral learning) variables, which are the central concerns of mobile learning policymakers, practitioners, and parents.

5

First, the study is timely given the current debates by policymakers and politicians, about the use of mobile devices in schooling. There is a focus in the media and much professional commentary on the adverse effects of school-aged students' use of mobile devices, including health problems like eyesight (China), potentially ethical issues (Indonesia), cyber-safety (Japan), classroom management concerns (Malaysia), and technology addiction (South Korea) (Churchill, Pegrum, & Churchill, 2018). The current meta-analysis provides a clear indication for policymakers on the effectiveness of mobile technology usage and evidence-based guidance on the use of mobile devices in schooling that provides a counterpoint to some of the current concerns. For example, some people believed that the use of mobile devices is not good for students' eyes, but in fact, the individual device helps students with poor eyesight see the learning content more clearly compared with look up at the backboards or whiteboards, especially those sitting in the back rows in a large classroom. For children, a mobile device is fast becoming a must-have not a nice-to-have, and it extends learning time and space (Norris & Soloway, 2015) and may sometimes serve as an unavoidable alternative for online learning (Dhawan, 2020). We recognize that hardware alone does not fulfill its potential in education and change teaching and learning fundamentally. However, different from traditional classroom learning and supported by mobile technologies' innovative features and their educational affordances, student-centered and active learning will become the new norm in tomorrow's education systems. More importantly, while the academic success of students historically determines the quality of school learning, the quality of the "learning process" has increased in importance and extends the understanding of learning outcomes (OECD, 2019). Therefore, policymakers who hesitate to scale up the use of mobile devices in education are encouraged to take actions either for improving educational quality or for bridging the digital divide. And before approving all actions under a given policy, there is an urgent need to articulate strategic intentions supplemented by established decision-making mechanisms and support.

Second, educational practitioners and parents may need to be convinced of the value of mobile learning to better prepare and support student learning. Long-term educational technology integration with appropriate supporting logistics may increase teachers' readiness to use digital technology (Christensen & Knezek, 2017) and the level of commitment to integrating their teaching with the students' learning (Khlaif, 2018). For example, if there is a lack of targeted teacher training in the preparation stage, and insufficient technical and pedagogical support during the phases of implementation, teachers might not be able to provide innovative teaching methods, and they might even reduce the time available for students to use mobile devices. Moreover, these conditions should include removing the negative effects, such as distraction, increased cognitive load, and mobile phone addiction. One way to solve these problems is to strengthen learners' self-regulation skills, as they are especially important for informal learning like homework performance (Nikou & Economides, 2018). Besides, the role of parents is important, as researchers pointed out that students' view of parental support is not only related to their learning motivation but also to their actual behaviors in self-regulating their learning (Sha, Looi, Chen, Seow, & Wong, 2012).

5.4.5 Conclusions

As interest in the tendencies of mobile learning and the affordances of mobile technologies, it is not only crucial of reimagining teaching and learning with mobile technology in primary and secondary education, but also valuable of reassessing the effectiveness of mobile technology usage on different learning outcomes as well as how to use mobile technologies for learning effectively, enjoyably, and engagingly. This study using the best evidence from experimental or quasi-experimental studies aimed to answer whether school students learn better with mobile technology and which factors explain the differences in results. Results of our meta-analyses of 72 cognitive comparisons from 59 studies, 24 affective comparisons from 22 studies, and 14 behavioral comparisons from 13 studies, indicated that mobile technology usage was positively and significantly associated with cognitive, affective, and behavioral learning outcomes. From both educational and methodological perspectives, the impacts of mobile technology usage were moderated by multiple factors, especially the student factors, learning process, and study quality factors. In the near future, researchers need to optimize the quality of experimental studies, and educational stakeholders need to take responsibility and get ready to adopt and support mobile technology usage in educational practices.