Received: 12 December 2020 | Revised: 30 March 2021 | Accepted: 10 April 2021
DOI: 10.1111/jcal.12557

ARTICLE

Journal of Computer Assisted Learning WILEY

A review of the meta-analysis by Tingir and colleagues (2017) on the effects of mobile devices on learning

Steve Bissonnette[1] | **Christian Boyer**[2]

[1]Department of Education, Université TÉLUQ, Québec, Québec, Canada
[2]Éditions de l'Apprentissage, SESSIONS, Montréal, Québec, Canada

Correspondence
Steve Bissonnette, Department of Education, Université TÉLUQ, 1122 Chemin Charbonneau, Lévis, Québec G7A 4T6, Canada.
Email: steve.bissonnette@teluq.ca

Abstract

Tingir et al. (2017) concluded from their meta-analysis that the subject areas taught through mobile devices had significantly higher achievement scores ($d = 0.48$) than the ones taught with traditional teaching methods. Given the relatively high positive effect of mobile devices on student achievement, we carefully analysed the selected research in this meta-analysis. We reviewed Tingir et al.'s (2017) meta-analysis based on analysis of the methodology of the selected research, while drawing on the work of Slavin (2003), Cheung and Slavin (2016), and Sung et al. (2019). Twelve of the 14 (86%) studies included in the meta-analysis done by Tingir and his team (2017) present such major methodological flaws that they should not have been included. Our analysis leads us to believe that the conclusion of Tingir et al. (2017) is not justified. It is recognized that duration of experiment is negatively correlated with effect size: the shorter the duration, the higher the effect (Burston, 2015; Slavin & Lake, 2009). Although demanding more effort, the field of education must raise the bar if it is to have knowledge of acceptable value.

KEYWORDS
effect, meta-analysis, methodological flaws, mobile devices

1 | INTRODUCTION

Tingir et al. (2017) concluded from their meta-analysis that: "[a] ccording to the results, the subject areas taught through mobile devices had significantly higher achievement scores than the ones taught with traditional teaching methods" (p. 366). This statement leads the reader to believe that the use of mobile devices is an undeniable asset in the classroom. However, several research (Abrahamsson, 2020; Baert et al., 2020; Beland & Murphy, 2016; Beneitoa & Vicente-Chirivellab, 2020) have shown positive effects of banning mobile phones in the classroom on student achievement. This effect is explained by the elimination of a major source of distraction associated with mobile phones. Although the results of Tingir et al. (2017) are consistent with two previous meta-analysis (Sung et al., 2015, 2016), it is odd that these results differ considerably from those presented in the meta-analysis conducted by Slavin et al. (2009) and Slavin (2019) on the use of technological tools in the classroom. If using new technologies in the classroom does not yield substantial positive results, how can mobile devices be effective in the same context? Likewise, the results of this type of educational organization in distance education and virtual schools, which maximize the use of mobile devices, would seem to produce serious negative effects over the long term (Barbour, 2019; Bueno, 2020). Again, how can this finding be reconciled with the Tingir et al. (2017) results of the meta-analysis on mobile devices and other meta-analysis on the subject? Is it because mobile devices under teachers' direct supervision can work, but not if this close supervision occurs in person?

The frequent methodological weaknesses of experimental research in education (Cheung & Slavin, 2013, 2016; Slavin, 2003; Slavin et al., 2009) and mobile devices have been highlighted (Burston, 2015; Sung et al., 2019). Are there research-design aspects that could explain what Tingir et al. (2017) meta-analysis presents as the effect of mobile devices?

Therefore, we decided to revise Tingir et al.'s (2017) meta-analysis based on analysis of the methodology of the selected research, while drawing on the work of Slavin (2003), Cheung and Slavin (2016), and Sung et al. (2019).

Our analysis leads us to believe that the conclusion of Tingir et al. (2017) is not justified. We follow the authors' designated sections for our review: reading, mathematics, and science.

2 | READING

The mean effect size in reading ($d = 0.666$) is the highest of the three school subjects assessed by Tingir et al. (2017). These researchers selected three studies, without differentiating between elementary ($N = 1$; sample of 160 students) and high school ($N = 2$; sample of 144 students). The total number of students is small (N total = 304 students), the number of studies low ($N = 3$), and the elementary and high school levels are mixed. It would have been much more desirable to have a sufficient amount of research to conduct an analysis by grade level (elementary, high school). Despite this very limited database, the authors still calculated an effect size of 0.666, which is considered a medium effect according to Cohen (1988; effect ≥ 0.20 = low, effect ≥ 0.50 = medium, effect ≥ 0.80 = high), which could even be described as a "medium superior" effect since it corresponds to a theoretical gain from almost 7 months of learning in the control groups. We will now briefly outline the three research studies on reading.

The research of Kim et al. (2011) shows a significant weakness that severely affects its results. The attrition of the experimental and control group was equal to or greater than 50% between the beginning and the end of the experiment. Yet, Dettori (2011) stated that attrition greater than 20% can seriously handicap the internal validity of research, while What Works Clearinghouse (WWC, 2014) found that an attrition rate comparable to that obtained by Kim and her colleagues leads to a potential for unacceptable bias.

Lin (2014) conducted reading research with Taiwanese ESL students in grade 10. Although the experimental group using an iPad performed better than the control group, this research has a significant methodological weakness. In fact, access to the iPad, an exclusive privilege of the experimental group, allowed students in this group to benefit from 2.55 more time reading than those in the control group. Clearly, the simple fact of reading for more time is the most likely explanation for the experimental group's superior results.

The duration of the experiment in the study by Yang, Tseng, et al. (2013) was only 180 min, which is extremely short and limited the interpretation of the results obtained. In addition, this study focused on learning, understanding, and interpreting traditional Chinese poems. The post-test evaluation consisted essentially of reinterpreting the poems studied during the experiment. This type of learning does not correspond at all to the common daily reading activities in a classroom.

In light of the above, the data on reading used in the meta-analysis by Tingir et al. (2017) do not make it possible to calculate a legitimate effect size for reading or to make a judgement about the effectiveness of mobile devices in the classroom.

3 | MATHEMATICS

The effect size in mathematics ($d = 0.160$) is the smallest of the three school subjects evaluated by Tingir and his team (2017). The researchers again selected only three studies, again not differentiating between elementary ($N = 2$; sample size 226 students) and high school ($N = 1$; sample size 316 students). Since the number of studies was still small ($N = 3$), the authors were not able to analyse the results according to grade level (elementary, high school), which would have been preferable. Despite the small pool of research, the authors calculated an effect size of 0.160, which is considered a low effect according to Cohen (1988; effect ≥ 0.20 = low, effect ≥ 0.50 = medium, effect ≥ 0.80 = high).

Riconscente's (2013) research did not demonstrate the effectiveness of the iPad in learning fractions in the experimental group, because the control group did not address fractions. In fact, the author herself stated that her study did not compare the iPad's application to traditional classroom instruction or to other computer applications, and therefore, "...the study did not enable conclusions to be drawn regarding the effectiveness [of the iPad's application] relative to other approaches to learning fractions" (Riconscente, 2013, p. 209).

Carr's (2012) research also had an important limitation: the experimental and control groups were not equivalent at baseline, with the difference benefiting the experimental group. Moreover, Carr pointed out that the statistical treatment did not succeed in completely reducing the difference to the benefit of the experimental group.

The independent variable in Liu and Lee's (2013) study included several elements, such as the use of the iPad, the use of a specific application on solids in mathematics, individualized teaching, and the use of the interactive whiteboard and collaborative work among students. This set of elements composing the independent variable makes it impossible to isolate the "handheld device" effect, yet this did not prevent Tingir and his team (2017) from including it in their meta-analysis.

With respect to mathematics, the weakness and biases of the studies selected in the meta-analysis by Tingir et al. (2017) should not have allowed them to calculate an effect size and do not provide a clear answer on the effectiveness of mobiles devices in the classroom.

4 | SCIENCE

Tingir et al. (2017) selected eight science research studies with an effect size of 0.528, again not considering the difference between elementary ($N = 5$) and high school ($N = 3$). Furthermore, the sample range of the research subjects of de-Marcos et al. (2010) varied between 14 and 21 years of age, which is not consistent with the target audience of Tingir et al. (2017), as stated in the title of their paper: "Effects of mobile devices on K-12 students' achievement: a meta-analysis." Nevertheless, Tingir et al. (2017) included college adult students aged 20 and 21 in the calculation of the overall effect size of the experimental group.

The research of Ahmed and Parsons (2013) took place over a very short period of time, which the authors did not specify, which however seems to correspond to the duration of a single teaching period on an experiment in physics (thermodynamics). It is very likely that the duration of the experiment was less than 120 min, which is too short and does not justify the generalization of the interpretation of the results obtained.

Billings and Mathison's (2012) object of learning in research is of little relevance to the field of science. This study was classified in the science section despite the fact that it is mainly related to the problem of learning English for students who do not master it. Further, this research took place at a museum and not in the typical setting of a daily science class.

Huang et al.'s (2010) research had the highest effect size ($g = 1.172$) in the science component. Despite the undeniable qualities of this study in botany, there are some caveats. The experiment lasted only 240 min and the experimental sample was composed of only 32 students equally divided between the experimental and control groups. Such a small sample with such a short experiment significantly reduces the power of the results of this research and the possibility of generalizations.

Hwang et al.'s (2013) research focused on a visit to a religious temple, which is not representative of daily classroom activities. Moreover, the subject matter was social sciences, unlike the other studies in this section that focused on the natural sciences. Finally, the experiment only lasted 180 min, which considerably limits the generalizations that can be made.

Nedungadi and Raman's (2012) research compared two experimental groups using technologies, without a comparison group using traditional teaching (without technologies). This did not prevent Tingir et al. (2017) from concluding, by including this study, that the use of mobile devices "...had significantly higher achievement scores than the ones taught with traditional teaching methods" (p. 366). While it is true that, in general, the term "traditional teaching" is loosely defined in research in education, this should be pointed out more often to researchers, with a view to raising their standards in this regard.

Oddly enough, the control groups of the Varma study (2014) did not receive any teaching related to the topics studied by the experimental groups (thermodynamics). The control groups simply did something else, without any further details from the author. The results obtained by the experimental groups were superior to those of the control groups at the post-test. It goes without saying that students who are taught content learn that content compared to those who are not taught it! Therefore, there is no real comparison in this research, as in the research by Nedungadi and Raman (2012), to measure the impact of a mobile device compared to traditional teaching without such device.

Yang, Hwang, et al. (2013) did a short 140-min study. Once again, the duration of such an experiment is clearly insufficient to make any generalizations.

The studies in science selected by Tingir et al. (2017) demonstrate a set of methodological flaws that cannot provide a sound basis for determining a convincing effect of mobile device in the classroom on student achievement.

5 | CONCLUSION

Tingir and associates (2017) calculated an overall mean effect of 0.48 for mobiles devices on student achievement by combining research in reading (effect $= 0.666$; $N = 3$), mathematics (effect $= 0.16$; $N = 3$) and science (effect $= 0.528$; $N = 8$). The authors did not do a school-level analysis (elementary, high school) in their school subject groupings due to the small number of studies, particularly in reading and mathematics. The researchers concluded their article by saying, among other things, that the use of mobile devices improves the performance of K-12 students compared to classes that do not use them and that offer traditional teaching.

All of the selected research studies have characteristics that prevent them from supporting the main conclusion purported by these researchers. First of all, it must be reiterated that at least two groups of research, reading and mathematics, were based on very few studies ($N = 3$ for each subject), which makes it difficult to calculate an effect size, especially since the results from elementary and high school were combined. Further, including college adult students' performance in a meta-analysis specifically dedicated to grades K-12 is inappropriate (see de-Marcos et al., 2010). Second, the learning content of several research studies focused on topics that clearly do not represent typical or common classroom instruction (see Billings & Mathison, 2012; Hwang et al., 2013; Yang, Tseng, et al., 2013). Therefore, we seriously question the inclusion of these studies in this meta-analysis. The experiments in other studies were so short in duration that they undermine any possible generalizations related to regular classroom activities (see Ahmed & Parsons, 2013; Huang et al., 2010; Hwang et al., 2013; Yang, Hwang, et al., 2013; Yang, Tseng, et al., 2013). In this regard, Cheung and Slavin (2016) recommend selecting only research with a minimum experimental duration of 12 weeks in order to conduct rigorous, relevant, and useful meta-analysis in education. It is recognized that duration of experiment is negatively correlated with effect size: the shorter the duration, the higher the effect (Burston, 2015; Slavin & Lake, 2009). Although demanding more effort, the field of education must raise the bar if it is to have knowledge of acceptable value. A more recent meta-analysis (Chen et al., 2020) on portable devices presents results comparable to Tingir et al. (2017), but 33% of the studies were conducted over fewer than 4 weeks. Sung et al.'s two meta-analysis (2015, 2016) are based on 59% and 44%, respectively, on studies with a duration of 4 weeks or fewer.

Finally, several studies in Tingir et al. (2017) have other major critical shortcomings, such as the absence of a control group (see Nedungadi & Raman, 2012), uncontrolled non-equivalence of the experimental and control groups (see Carr, 2012), and a lack of content coverage in the control group (see Riconscente, 2013; Varma, 2014).

In sum, 12 of the 14 (86%) studies included in the meta-analysis done by Tingir and his team (2017) present such major methodological flaws that they should not have been included. Given all of the above, it is our opinion that the meta-analysis by Tingir et al. (2017) cannot be used to justify the use of mobile devices in the classroom.

Unfortunately, as of March 2021, the meta-analysis by Tingir et al. (2017) has been cited 55 times in various scientific articles and websites, including by the Education Endowment Foundation (2019).

CONFLICT OF INTEREST

We have no conflict of interest.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/jcal.12557.

DATA AVAILABILITY STATEMENT

Research data not shared.

ORCID

Steve Bissonnette https://orcid.org/0000-0002-6340-6062

REFERENCES

Abrahamsson, S. (2020). *Distraction or teaching tool: Do smartphone bans in schools help students?*. Department of Economics, Norwegian School of Economics. https://sites.google.com/view/saraabrahamsson/research?authuser=0

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

Baert, S., Vujić, S., Amez, S., Claeskens, M., Daman, T., Maeckelberghe, A., Omey, E., & De Marez, L. (2020). Smartphone use and academic performance: Correlation or causal relationship? *Kyklos*, *73*, 22–46. https://doi.org/10.1111/kykl.12214

Barbour, M. K. (2019). What virtual and blended education research reveals. Section II. In A. Molnar, G. Miron, N. Elgeberi, M. K. Barbour, L. Huerta, S. R. Shafer, & J. K. Rice (Eds.), *Virtual schools in the U.S. 2019* (pp. 41–83). National Education Policy Center.

Beland, L. P., & Murphy, R. (2016). Ill communication: Technology, distraction & student performance. *Labour Economics*, *41*, 61–76.

Beneitoa, P., & Vicente-Chirivellab, O. (2020). *Banning mobile phones at schools: Effects on bullying and academic performance*. University of Valencia and ERICES. https://www.erices.es/upload/workingpaper/99_99_0420.pdf.

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.

Bueno, C. (2020). *Bricks and mortar vs. computers and modems: The impacts of enrollment in K-12 virtual schools* (EdWorkingPaper: 20-250). Annenberg Institute at Brown University. https://doi.org/10.26300/kahb-5v62

Burston, J. (2015). Twenty years of MALL project implementation: A meta-analysis of learning outcomes. *ReCALL: The Journal of EUROCALL*, *27*(1), 4.

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.

Chen, Z., Chen, W., Jia, J., & An, H. (2020). The effects of using mobile devices on language learning: A meta-analysis. *Educational Technology Research and Development*, *68*(4), 1769–1789.

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.

Cheung, A. C., & Slavin, R. E. (2016). How methodological features affect effect sizes in education. *Educational Researcher*, *45*(5), 283–292.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd ed.)*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

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

Dettori, J. R. (2011). Loss to follow-up. *Evidence-Based Spine-Care Journal*, *2*(1), 7–10. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3427970/

Education Endowment Foundation. (2019). *Evidence for Learning Teaching & Learning Toolkit: Digital technology*. https://educationendowmentfoundation.org.uk/pdf/generate/?u=https://educationendowmentfoundation.org.uk/pdf/toolkit/?id=134&t=Teaching%20and%20Learning%20Toolkit&e=134&s=

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 and Education*, *54*, 47–58.

Hwang, G. J., Wu, P. W., 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*(4), 338–354.

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.

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

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

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

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

Slavin, R. (2003). A reader's guide to scientifically based research. *Educational Leadership*, *60*(5), 12–16.

Slavin, R. (2019, November 14). A powerful hunger for evidence-proven technology [Blog post]. https://robertslavinsblog.wordpress.com/2019/11/14/a-powerful-hungerfor-evidence-proven-technology/

Slavin, R. E., & Lake, C. (2009). Effective programs in elementary mathematics: A best evidence synthesis. *Review of Educational Research*, *78*(3), 427–455.

Slavin, R., Lake, C., & Davis, S. (2009). *Meta-findings from the best evidence encyclopedia*. Center for Data-Driven Reform in Education, Johns Hopkins University. http://www.bestevidence.org/resources/resources.htm

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

Sung, Y. T., Chang, K. E., & Yang, J. M. (2015). How effective are mobile devices for language learning? A meta-analysis. *Educational Research Review*, *16*, 68–84.

Sung, Y. T., Lee, H. Y., Yang, J. M., & Chang, K. E. (2019). The quality of experimental designs in mobile learning research: A systemic review and self-improvement tool. *Educational Research Review*, *28*, 100279.

Tingir, S., Cavlazoglu, B., Caliskan, O., Koklu, O., & Intepe-Tingir, S. (2017). Effects of mobile devices on K-12 students' achievement: A meta-analysis. *Journal of Computer Assisted Learning*, *33*(4), 355–369.

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

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

Yang, C. C., Tseng, S. S., Liao, A. Y. H., & Liang, T. (2013). Situated poetry learning using multimedia resource sharing approach. *Educational Technology & Society*, *16*(2), 282–295.

What Works Clearinghouse. (2014). *Procedures and standards handbook (version 3.0)*. U.S. Department of Education, Institute of Education Sciences. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_v3_0_standards_handbook.pdf

How to cite this article: Bissonnette S, Boyer C. A review of the meta-analysis by Tingir and colleagues (2017) on the effects of mobile devices on learning. *J Comput Assist Learn*. 2021;1–5. https://doi.org/10.1111/jcal.12557