A Spotlight on Lack of Evidence Supporting the Integration of Blended Learning in K-12 Education:
A Systematic Review

Mark Poirier, Parenting and Special Education Research Unit, Katholieke Universiteit Leuven, Leuven, Belgium

Jeremy M. Law, School of Interdisciplinary Studies, College of Social Science, University of Glasgow, Glasgow, UK

Anneli Veispak, Tallinn University, Tallinn, Estonia

https://orcid.org/0000-0001-8432-2549

ABSTRACT

In an effort to highlight the need for, and lack of, quality empirical research in K-12 blended learning environments, this systematic narrative review investigated and reported on the quantity and quality of recent empirical research in K-12 blended learning, published between 2009 and February 2017. In addition to assessing the quality and scope of these studies, the effectiveness of blended learning environments on learning outcomes and potential contributing variables were discussed. Eleven articles were identified and found to meet the inclusionary criteria and measures of quality set by this review, extending the corpus of 5 articles identified by a previous 2009 meta-analysis commissioned by the U.S. Department of Education to 16. Mixed findings regarding the benefit of blended learning in a K-12 setting were reported across the literature, thereby highlighting the need for more extensive research in this domain.

KEYWORDS

Blended Learning, Effectiveness, Hybrid Learning, K-12, Review

1. INTRODUCTION

Modernity inspires innovation. Nowhere is this more evident than in an educational context. One such innovation is distance education (Bonk & Graham, 2006). For decades, distance education has offered an opportunity for isolated or busy individuals to learn without the expense and ongoing commitment of attending permanent, brick and mortar schools. Over the years, the practice of correspondence education has evolved to become more interactively streamlined through the use of web-based technologies (Garrison & Kanuka, 2004). As accessibility and content of various web-based platforms and resources have grown, so too have their integration and blending with traditional face-to-face instruction.

The term 'blended' or 'blended learning' has been applied quite broadly throughout literature and educational practice. Graham (2006) noted that early models exploring blended learning attempted to address the question of "what is blended?" and resulted in 3 distinct conclusions: 1) a combination of online and face-to-face instruction (Young, 2002; Friesen, 2012); 2) a combination of instructional

modalities or delivery media (Thomson, 2002); 3) a combination of different instructional methods (Driscoll, 2005).

The most prominent definition adopted within recent educational research has associated the term 'blended' with both learning and learning environments in situations where technology and/ or online resources are combined in some form with traditional in-class instruction. In an effort to refine this definition and establish essential characteristics of blended learning Bernard, Borokhovski, Schmid, Tamim, and Abrami (2014) argued that no less than 25% of time should be spent in online environments with 50% or more time spent face-to-face. Others, conversely, have argued for the use of more discrete categorisation where: 1) traditional instruction was defined as having 0% of online content delivery; 2) 1%-29% of online delivery was defined as 'web facilitated' instruction; 3) 'blended' was defined as 30%-79% online; and iv) the use of 80% or greater online content delivery, defined as 'online' (Allen & Seman, 2007).

Due to the lack of location-based restrictions inherent in online learning, some blended learning environments and designs may result in the use of a flipped classroom pedagogical model where students receive and engage with lecture material at home at their own pace while completing assignments and discussions in class (Lage, Platt, & Treglia, 2000).

In the current systematic review, blended learning will be defined as the deliberate integration and combination of online, computer-based learning and instruction with face-to-face methods of learning and instruction, regardless of the application of a flipped or traditional pedagogical structure. As empirical studies examining blended learning in primary and secondary educational (K-12) settings are scarce, a wide-ranging online to face-to-face time allotment criteria is applied to increase the likelihood of inclusion similar to the meta-analysis of Means, Toyoma, Murphy and Baki (2013). Within this review, 'deliberate combination' is considered as a more important criteria than that of the amount of time spent in either online or in-class environments.

1.1. The Increasing Popularity of Blended Learning Environments

Over the past decade, online and blended learning environments have become a sweeping trend throughout the world (Werth, Werth, & Kellerer, 2013). Rudestam and Schoenholtz-Read (2010) noted that this trend has been primarily driven by the belief in its potential for greater flexibility and cost-effectiveness in accessing content and quality instruction, as well as providing new levels of interactivity, social networking, collaboration and reflection.

Many post-secondary institutions have aggressively adopted this revolution in instructional approach. For instance, the University of Ottawa is aiming to classify at least 20% of offered courses as blended by 2020 (Smith, 2014), and 57.5% of undergraduate classes taken nationwide have been reported to involve some blending by the Canadian Higher Education Strategy Association (2011). Additionally, over 4.6 million students in US higher education institutions were taking at least one online course in 2008 (Allen and Seaman, 2010).

Similarly, primary and secondary schools have begun to move towards greater integration of online learning with traditional educational environments. For instance, the United States saw an increase of K–12 students participating in some form of online/blended course instruction from roughly 45,000 in 2000, to over 4 million students in 2010 with a projected estimate of 5 million K-12 students engaged in blended learning courses by 2016 (Staker, 2011; Picciano, Seaman, Shea, & Swan 2011). Likewise, in Canada the Ontario Ministry of Education (2016) has published a commitment to make blended learning available for all Ontario students from kindergarten to grade 12 (ages 4-18).

1.2. Existing Evidence on the Effects of Blended Learning

Previous research involving post-secondary and professional training programs has demonstrated the benefits of blended learning with regards to motivation, flexibility, student engagement and academic achievement when compared to purely online or traditional face-to-face didactics (Bonk & Graham, 2006; Means, Toyama, Murphy, Bakia & Jones, 2009).

This success has been attributed to blended learning models' ability to offer students control over time, pace, path and/or place, allowing for more student-centered learning experiences and greater flexibility for class participation and engagement. The increase in flexibility is proposed as one potential explanation for academic gains in blended learning environments when compared with traditional instructional approaches (Bernard et al., 2014; Bonk & Graham, 2006; Garrison & Kanuka, 2004). Others have argued that the identification of direct mechanisms pertaining to the success of blended learning may not be readily identifiable because they are potentially subject dependent and mediated by variables such as self-regulation (Lynch & Dembo, 2004) and student perception of learning (Ginn & Ellis, 2007).

To address questions regarding the effectiveness of online and blended learning, the U.S. Department of Education commissioned one of the most extensive reviews to date for blended and online learning, which examined the corpus of research published between 1996-2008 (Means et al., 2009). The meta-analysis revealed that instruction with blended online elements provided the greatest mean effect on learning outcomes when compared to purely face-to-face instruction or purely online instruction. Specifically, it was suggested that the provision of greater autonomy in controlling online materials triggers learner activity, reflection and a self-monitoring of understanding (metacognition), which in turn enhances learning while the inclusion of certain elements (i.e. media by way of video and online quizzes) did not influence the learning outcomes in online classes (Means et al., 2009).

One striking outcome of the Means and colleagues (2009) review was a near absence of research comparing the learning effects and outcomes of online, blended and face-to-face learning environments in the K-12 domain. Between 1996 and 2008, only nine peer-reviewed studies relating to K-12 education were identified. However, only five possessed adequate statistical data and controls to warrant their inclusion in their review. Based on these limited sources, Means and colleagues (2009) found that blended learning in a K-12 setting had mixed effects relating to learning outcomes. Three of the K–12 studies showed significant effects favouring a blended learning condition, yet the small number of studies involved limited their ability to calculate the mean effect estimate for the learning group confidently. Means and colleagues reported that the difference in effects between learning environments was determined mainly by moderating factors such as student preferences, subject content, relationship dynamics (communication), motivation or immediacy of feedback. In essence, findings from the paltry amount of scientific research in the K-12 domain provided little conclusive evidence for which the integration of blended learning could be sufficiently justified. Means et al. stressed the need for greater focus and rigour in the investigation of blended learning in K-12 settings as findings of effectiveness in higher education, professional education and adult education might not directly apply to K-12 students on a developmental level.

More recent reviews by Bernard and colleagues (2014) and Means and colleagues (2013) reported no new published research with a focus on K-12 blended learning.

1.3. Considerations in the Research of Blended Learning

When examining the effectiveness of blended learning in comparison with traditional face-to-face methods, several research design factors can affect differences. Variables to be considered and controlled for when designing a comparative study are, but not limited to: (1) individual differences of subjects (i.e., age, IQ, socio-economic status (SES), previous knowledge); (2) differences within the curriculum material (i.e. science vs language learning); (3) differences in learning time allotted within the treatment and control condition. As many studies comparing learning environments and methodologies fail in controlling for such variables, caution should be urged in the generalisation of any reports of positive learning effects. The lack of rigorous controls makes it difficult to attribute any observed advantage for blending per se as results could be equally attributed to differences in the uncontrolled variables. For instance, Means and colleagues (2009) noted that in some cases the reported advantages of blended learning could have been explained by increased time-on-task when compared to traditional face-to-face learners and not the type of learning environment.

Specific differences in the characteristics of the blended learning activities and design may additionally attribute to differences in reported effectiveness across the literature. For instance, the effectiveness of different blended learning activities could be influenced by the extent to which the activity is synchronous with the instruction and if learning is occurring in real time (physical or virtual). Means and colleagues (2013) argued that such activities offer greater spontaneity, instilling a sense of cohesion amongst learners and promoting collaboration (Hermann, Rummel, & Spada, 2001). However, Means did note that students may feel rushed to respond within these environments or encumbered by technology breakdowns. This can be contrasted with an asynchronous design where the presentation of instruction is separated by a period of time from the student responses. It has been argued that an asynchronous design offers a higher level of flexibility concerning location and time, permitting more thoughtful and reflective involvement (Veerman & Veldhuis-Diermanse, 2010; Jonassen, Lee, Yang, & Laffey, 2005).

A second relevant characteristic of the blended environment is the extent of change in the students accustomed learning environment, methods and schedule. Often, concerning teacher's/class' role, blended learning environments maintain a fair element of the status quo typically witnessed in classroom learning (i.e., traditional didactics and interaction). The core differences experienced by the student often lie in the availability and means of accessing target information, while the structure, timing, physical environment and teacher/peer interactions may not be affected. However, other means of blending may employ a flipped model design, which alters the learning schedule, style, and difficulty level students are accustomed to. It is suggested that such alterations may invoke a cognitive load negatively impacting initial student achievement (Sweller, 1988).

A third important characteristic when comparing effectiveness across blended learning environments and activities is the extent of peer and teacher interaction within the learning community. In relation to blended environments, the term 'learning community' is often used to refer to the coordinated mixture of face-to-face and virtual interactions among peers which may or may not be teacher directed (Schwen & Hara, 2004). The facilitation of learning community interactions and discourse is a vital component often attributed to the success of blended learning, as the interaction between learners in a group enhances understanding and influences cognitive activity (Cohen, 1994; King, 2002).

1.4. Present Study

To justify the time and costs associated with the development of required infrastructure and curriculum, quality research and evidence supporting the benefits of blended learning environments in K-12 settings are paramount. Both policymakers and practitioners require substantial evidence to know how to most effectively utilise blended learning to arouse positive change in the learning, achievement and motivation of all students in a K-12 context.

As highlighted in past reviews, most research on the outcomes of various blended learning environments has focused on higher educational contexts. As a result, a relevant gap in the knowledge of the effectiveness of blended learning in the K-12 educational setting has become evident.

To highlight the need and lack of quality empirical research of K-12 blended learning environments, this systematic review will investigate and report on the quantity and quality of empirical research in K-12 blended learning environments published since the comprehensive review of Means and colleagues in 2009. Therefore, this review will focus on research published between 2009 and February 2017 and assess the quality and scope of these studies. The discussion will examine the results of these studies to identify if blended learning methods do indeed significantly affect learning outcomes among the K-12 student population.

More specifically within the context of this systematic review, we will identify and discuss:

1. The influence of, and differences in the educational settings, instruments of blended learning and subject/content area reported within the literature;

2. Empirical evidence pertaining to the outcomes and learning effects of blended learning;
3. Potential moderating variables related to the design of the blended learning environment of a K-12 setting.

2. METHODOLOGY

2.1. The Systematic Approach

This current systematic review has been completed according to the specific framework proposed and defined by Torgerson (2003) and Pettigrew and Roberts (2008). As such, this review will identify, appraise, and synthesise the available literature pertaining to the current research questions in the hopes that findings can be amassed in a more accessible way and aid in creating a better understanding of K-12 blending learning (Pettigrew et al., 2008).

2.1.1. Databases and Journals Searched

Searches were performed electronically in Scopus, ERIC, ProQuest Educational Journals and LIMO using combinations of inclusion and exclusion terminology in tandem with Boolean operators. References in relevant literature were also searched.

2.1.2. Inclusion/Exclusion Criteria

Inclusion and exclusionary criteria were designed to mirror those set in past work in an attempt to extend previous reviews regarding blended learning effectiveness in a K-12 context (i.e. Means et al., 2009; Means et al., 2013; Bernard et al., 2014). Therefore, for this review, inclusion criteria were as follows:

1. Interventions must compare blended learning to traditional face-to-face learning in some way;
2. For the intervention to be considered 'blended,' online portions had to be deliberately combined with traditional face-to-face instruction and materials;
3. Acceptable research designs for inclusion include random experimental design, quasi-experimental design with appropriate control, and mixed-method design;
4. The time frame for this literature search was from 2009 to February 2017;
5. The study must be within a K-12 educational context;
6. Sample size must be equal to or greater than 20 participants within the experimental group to address any concerns of statistical power of the reported results;
7. Studies must exhibit statistical control for pre-existing group differences;
8. Studies must be published in internationally ranked peer-reviewed journals (as determined by SCImago Journal Rank (SJR indicator) and/or Web of Science).

Four exclusion criteria are included:

1. Comparisons of purely online learning environments with traditional face-to-face learning environments;
2. Studies investigating education interventions solely involving populations with disabilities or gifted students;
3. Studies with a focus on professional/corporate training or adult education are excluded;
4. Mobile learning and studies on gamification are excluded.

2.2. Selection of Papers for Inclusion

Keyword database searches yielded 2,954 studies through multiple searches using permutations of combined inclusion/exclusions terminology, Boolean operators, and wildcards (see Table 1 in the Appendix for a list of the search terms used in the Boolean operators during the initial search). Of these studies, 85 were identified as initially meeting our selection criteria and were submitted for additional screening. Some research used a variation of the term 'blended' which did not meet our established definition or failed to clarify the nature of the online intervention. Seventy entries were excluded following the second phase screening process that included screening for quality and relevance using the Jadad 5-point scale (Jadad, 1996) for random control trials and the Cochrane EPOC checklist was consulted for quasi-experimental design or mixed method designs (Greenhalgh, Robert, Bate, Macfarlane & Kyriakidou, 2005). All studies failing to meet the minimum criteria set by these standards were removed. Additional reasons for exclusion were primarily due to a purely online intervention or a failure to be within the K-12 educational context.

Of the 15 remaining studies after the screening, four were later excluded from this systematic review due to: 1) failures in statistically controlling for group differences (i.e. Naidoo & Naidoo, 2007); 2) ambiguity in terminology, population, and context (i.e. Sudha & Amutha, 2015); 3) low sample size within the experimental group (i.e. Smith, 2013); 4) an uncontrolled imbalance in age and experience level of the students between groups, thus potentially confounding results (i.e. Lee, Cheung, Wong & Lee 2013).

3. RESULTS AND DISCUSSION

3.1. Scope of Research Conducted K-12 Blended Learning

3.1.1. Sample and Setting

In the 11 included studies, the sampled grades were ranging from 5th to 12th. Eight studies focused on middle school and high school with the age range of 13 to 18 years (Aidinopoulou & Sampson, 2017; Akgunduz and Akinoglu, 2016; Wendt & Rockinson-Szapkiw, 2014; Wendt & Rockinson-Szapkiw, 2015; Clark, 2015; Jia, Chen, Ding & Ruan, 2012; Yapici & Akbayin, 2012; Chang, Shu, Liang, Tseng & Hsu, 2014; Kazu & Demirkol, 2014; Pane, McCaffrey & Karam, 2014). The remaining two studies (Aidinopoulou & Sampson, 2017); Mulqueeny, Kostyuk, Baker & Ocumpaugh, 2015) investigated grade 5 students (9-10 years of age), with the latter additionally including grade 6 (11-12 years of age).

The sample sizes reported across this research was generally quite modest and a far cry from the large-scale studies that have examined adult aged subjects. Treatment and control groups were often composed of 2 to 4 classes within the same school setting, containing approximately twenty or more students per class. In one exceptional study the sample consisted of 25,500 students from over 147 schools (74 middles schools (n=6800), 73 high schools (n = 18700)) (Pane et al., 2014).

Five of the included studies investigated samples in the USA and three investigated samples in Turkey. The remaining three were from Greece, Taiwan and China (Hong Kong). Reporting of possible demographic moderators such as socio-economic status, ethnicity, gender, religion, or family structure lacked across all studies with the apparent exception of Pane et al., and 3 studies that reported a gender distribution in their study groups (Kazu et al., 2014; Clark, 2015; Akgunduz & Akinoglu, 2016; Aidinopoulou & Sampson, 2017). Such a lack of demographic controls limits any insight into how variables like socio-economic status or gender, might affect the outcomes of blended learning interventions reported within these studies.

3.1.2. Subject Areas

Subject matter across studies was relatively diverse, but all were based on concrete subject materials and quantifiable data. For example, 9 out of 11 involved science or math (1 electrical engineering, two algebra, five science, and 1 STEM). The remaining two studies investigated the effects of blended

learning on learning outcomes in history (Aidinopoulou & Sampson, 2017) and language learning (Jia et al., 2015). However, in both of these studies, either a multiple choice or a cloze test was being used, both of which are subject to a right- or a wrong answer. In the study of Aidinopoulou and Sampson (2017), an additional teacher generated task was employed that contained some open questions in the assessment of historical thinking.

In the literature, a gap often exists in the assessment of blended learning outcomes in arts and humanities when compared to sciences and math. A possible explanation may lay in a comparative study design's need for all outcome measures to be reliable and consistent. Therefore, the inclusion of subject matter which inherently permits greater subjective interpretation, such as literature or creative writing, increases variability within the scoring which is not the case of binary evaluation (right-wrong answers).

This limited breadth of subject matter included in the assessment of blended learning in K-12 has limited our ability to see the value of blended learning across both concrete and subjective aspects of education.

3.2. Empirical Evidence of Effectiveness

3.2.1. Assessment of Learning Outcomes

According to Means and colleagues (2009, 2013), blended learning is defined as an enhancement of education. It is a manifestation of additional time that can be spent on any given subject outside of the physical classroom as well as supplementary access to instructors and peers. As a result, blended learning environments do not necessarily reduce class time or the need for traditional infrastructure, yet in turn, require an initial investment of capital and time in designing blended learning tools and curricula. Therefore, it is vital to ensure this investment offers specific and measurable returns concerning student success and advances in achievement when compared with traditional instruction methods.

Nine of the eleven studies included academic achievement as a research focus. Other areas of inquiry investigate aspects which may be influential to learning in general. These aspects include motivation, attitude, cognition, engagement, communication, and community. In the following section, these factors will be discussed more explicitly.

Of the included studies, mixed results were reported concerning blended learning's effect on academic achievement in the K-12 domain. Four studies demonstrated a significantly higher level of academic achievement in blended learning contexts than was found in traditional face-to-face learning (Aidinopoulou & Sampson, 2017; Kazu et al., 2014; Mulqueeny et al., 2015; Yapici et al., 2012). However, it is worth noting that the study of Aidinopoulou and Sampson (2017) although the experimental group was found to have significantly larger gains relative to controls on measures of critical historical knowledge, such as the ability to interrupt the significance of a specific event in history or to understanding the forces, choices, and circumstances that brought about a historical event, both groups were found to perform equally on measures of historical facts; such as recall of dates and places. Chang and colleagues (2014) conversely, observed a greater achievement outcome in the experimental group, but these differences were not found to be statistically significant. In the study performed by Jia and colleagues (2012), the experimental group was only found to outperform the controls on measures of vocabulary significantly. Pane et al. (2014) conducted a two-year study in 74 middle schools and 73 high schools across the U.S. to investigate the use of an online cognitive algebra program as remediation for consistently sub-standard test scores in math. Analysis of post-test outcomes found no effects of the program on learning achievements based on the first year of implementation. However, a gain was measurable and significant by the end of the second year. The estimated effect of the blended learning intervention on learning outcomes was found to be statistically significant for high school students but not for middle school. Yet, it should be noted that a follow up study of Pane's sample found that teachers implemented the blended curriculum with low fidelity as

blended-curriculum teachers in the second year reverted to more traditional approach to instruction (Karam et al., 2017), thus potentially influencing the results reported by Pane and colleagues (2014).

Clark (2015) did not find a significant difference in post-test scores following the intervention. Similarly, Wendt and Rockinson-Szapki (2014) found that learners who participated in collaborative activities in the traditional classroom had fewer science misconceptions than those who participated in the blended learning experimental group.

In summary, four studies found a statistically significant increase in academic performance. One study observed mixed results. Three studies reported higher, yet not significant, achievement scores in the intervention groups relative to controls. One observed a significant gain in academic achievement in the control group rather than in the blended learning group, and two studies did not measure academic achievement at all (Wendt, 2015; Akgunduz & Akinoglu, 2016).

Concerning the subject disciplines, no clear pattern regarding the effectiveness of blended learning for different subject matter emerged across studies. Significant learning outcome effects were reported for two studies involving biology, one focusing on mathematics and a single study involving history instruction.

Seven other outcomes were identified in the reviewed articles. Akgunduz and Akinoglu (2016) measured the difference in attitudes towards science classes and self-directed learning skills. Results indicated that students in the blended learning group had significantly larger gains in their science attitude and self-directed learning skills than the control group. Chang and colleagues (2014) investigated self-assessed perceptions of skill and observed an overall larger effect in the experimental group regarding these self-assessed aspects.

The level of student engagement was assessed by 2 of the 11 studies. Of these 2, Clark (2015) noted that in the blended learning group, students were 4% more engaged relative to controls. Similarly, Mulqueeny and colleagues (2015) demonstrated, in the blended learning group, a significant increase in engagement (i.e., more time spent independently on task, more time proactively seeking remediation on task, less time in off-task conversation and less time off task when compared to the control group); while observing less boredom, less confusion and less delight (i.e., possible distractions from learning) than their peers in face-to-face learning settings. Wendt and colleagues (2015) investigated the sense of community in blended and traditional learning environments and found that students in traditional learning environments possess a significantly greater sense of community.

3.3. Variables to be Considered in Blended Learning Environments

Broad interpretation and generalisation of the reported effects from these studies is conditioned as practice, and condition variables pertaining to the different blended environments reviewed differed across many aspects; each potentially contributing to some of the observed variances in outcomes. Of the studies reviewed, many did not directly report or provide information concerning features within the blended environments, all of which could have been responsible for a proportion of variance in the measured outcomes—an obstacle highlighted in previous reviews and meta-analyses (see Means et al., 2013). Many of the reviewed studies, for example, did not indicate measures of individual time-on-tasks, the extent to which specific media features were used, and implementation fidelity by the teacher.

In the studies, which reported greater achievement gains within blended learning environments when compared with traditional practice, discussion of potentially influential variables was not explicitly offered. Additionally, reasons for the academic success were not offered or clearly defined.

Of the 11 included studies, 4 reported no inclusion of an instrument to facilitate the peer-to-peer exchange or the promotion of a learning community (Clarke, 2015; Jia et al., 2012; Pane et al., 2014). Of the four studies that demonstrated a significantly higher level of academic achievement in blended learning contexts, 3 included means of facilitating peer-to-peer interaction (Aidinopoulou & Sampson, 2017; Kazu et al., 2014; Mulqueeny et al., 2015). Each study reported the use of various means of encouraging and facilitating peer-to-peer and learner-teacher engagement (i.e., instant

message platforms, blogs or online forums), with Yapici and colleagues (2012) being the exception, as they did not elaborate on the forums use between users (i.e., student-teacher interaction/ peer-to-peer interaction). Unfortunately, no study reported measures of the level of peer-to-peer engagement or use of these tools during the treatment period, which limits the chance to draw any conclusion of the potential effect these tools may exert on learning outcomes. Wendt and colleagues (2014) argued that this connection within a learning community (interaction) is essential and should be carried out in a synchronous manner with learning. The intervention of Wendt et al. (2014) was one of the four studies to have employed some form of asynchronous instruction within the blended design. Wendt and colleagues hypothesised that the observed insignificant learning achievement gains of their intervention group might primarily be a result of issues with the delay in asynchronous communication. Wendt noted that the asynchronous delay in responses from the teacher reinforced misconceptions in the experimental group and that the immediacy of correction in the traditional face-to-face class may have facilitated change in students' understanding, resulting in significantly higher post-test scores of the control group.

An additional feature shared across the studies that reported greater achievement effects of blended environments when compared to purely face-to-face instruction was the minimal disruption the intervention exerted on the students established routine, and learning environment. The included studies that did not achieve significant academic results often enacted a more considerable amount of procedural change. Aidinopoulou and Sampson (2017), Clark (2015) and Kaza (2014), for example, introduced a flipped model design which altered the learning schedule students were accustomed to, requiring greater participation outside of traditional class time. In interviews with participants, Clark recorded statements regarding the difficulties students had in adjusting to the new schedule and content of the lessons. The increased cognitive load might have had an inhibiting impact on student achievement (Sweller, 1988). Given the academic test results in this particular experiment were roughly equivalent to those of the control group, would suggest the instructional success of the intervention. However, it does not fill the requirement of enhancement that a blended learning environment should possess to be a justified educational innovation.

Conversely, the study of Aidinopoulou & Sampson (2017) demonstrated that when given enough time allowing students to acclimatise to the changes, positive and significantly higher learning outcome gains can be found. Suggesting that for students to become accustomed to changes in the traditional face-to-face methods might take more time, than the 5 to 12 weeks provided by most of the studies within this review. Future work, which would control the implementation of varied exposure time is needed to accurately determine the role this adjustment period may have on short-term learning outcomes.

4. CONCLUSION

To highlight the need for, and lack of quality empirical research in K-12 blended learning environments, the current review was conducted to investigate, and report on the quantity and quality of recent empirical research of K-12 blended learning. In addition to an assessment of the quality and scope of these studies, the effectiveness of the blended learning environments on learning outcomes and potential contributing variables were also discussed. Eleven articles published since 2009 were found to meet the inclusionary criteria and quality standards set for the current review, extending the list of articles published between 1996 and 2017 to 16.

The lack of good quality extensive research focusing on blended learning in K-12 environments can potentially be explained by the inherent disruption such a study would inevitably introduce to the daily routines of large samples of students, school administrations, teachers and parents for the period required to conduct a well-controlled randomised study in a K-12 environment adequately. Additionally, the costs related to software, online platform design and hardware may pose funding barriers to conducting studies on blended learning.

The majority of the research that has investigated blended learning to this point has not been empirical, but rather exploratory, often focusing on definitions, models, and the potential of blended learning. Future work should strive to reflect more authentic learning situations longitudinally while controlling for characteristics in the context and sample, which might play a role in confounding the results. The strict application of an experimental design within blended learning research would help to increase the validity and repeatability of the work in different populations and settings, thus extending the generalizability of findings beyond one specific research context.

Based on the results of this systematic review, the existing evidence does not justify the current rapid adoption of blended learning practices and the outlay of associated expenses, which have been witnessed across North America over the past decade. More research is strongly urged and required before the effects of blended learning on learning outcomes in K-12 settings are genuinely understood.

REFERENCES

Aidinopoulou, V., & Sampson, D. G. (2017). An Action Research Study from Implementing the Flipped Classroom Model in Primary School History Teaching and Learning. *Journal of Educational Technology & Society*, *20*(1), 237–247.

Akgunduz, D., & Akinoglu, O. (2016). The effect of blended learning and social media-supported learning on the students' attitude and self-directed learning skills in science education. TOJET: The Turkish Online Journal of Educational Technology, 15(2).

Allen, M., Bourhis, J., Burrell, N., & Mabry, E. (2002). Comparing student satisfaction with distance education to traditional classrooms in higher education: A meta-analysis. *American Journal of Distance Education*, *16*(2), 83–97. doi:10.1207/S15389286AJDE1602_3

Allen, M., Bourhis, J., Mabry, E., Burrell, N. A., & Timmerman, C. E. (2006). Comparing distance education to face-to-face methods of education. In Classroom communication and instructional processes: Advances through meta-analysis (pp. 229-244).

Allen, M., Mabry, E., Mattrey, M., Bourhis, J., Titsworth, S., & Burrell, N. (2004). Evaluating the effectiveness of distance learning: A comparison using meta-analysis. *Journal of Communication*, *54*(3), 402–420. doi:10.1111/j.1460-2466.2004.tb02636.x

Bernard, R. M., Borokhovski, E., Schmid, R. F., Tamim, R. M., & Abrami, P. C. (2014). A meta-analysis of blended learning and technology use in higher education: From the general to the applied. *Journal of Computing in Higher Education*, *26*(1), 87–122. doi:10.1007/s12528-013-9077-3

Bonk, C. J., & Graham, C. R. (2006). *The Handbook of Blended Learning: Global Perspectives, Local Designs*. John Wiley & Sons.

Cavanaugh, C., Gillan, K. J., Kromrey, J., Hess, M., & Blomeyer, R. (2004). *The effects of distance education on K-12 student outcomes: A meta-analysis. Learning Point Associates/North Central Regional Educational Laboratory*.

Cavanaugh, C. S. (2001). The Effectiveness of Interactive Distance Education Technologies in K-12 Learning: A Meta-Analysis. *International Journal of Educational Telecommunications*, *7*(1), 73–73.

Chang, C. C., Shu, K. M., Liang, C., Tseng, J. S., & Hsu, Y. S. (2014). Is blended e-learning as measured by an achievement test and self-assessment better than traditional classroom learning for vocational high school students? *The International Review of Research in Open and Distributed Learning*, *15*(2). doi:10.19173/irrodl.v15i2.1708

Clark, K. R. (2015). The effects of the flipped model of instruction on student engagement and performance in the secondary mathematics classroom. *Journal of Educators Online*, *12*(1), 91–115. doi:10.9743/JEO.2015.1.5

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, *64*(1), 1–35. doi:10.3102/00346543064001001

Driscoll, M., & Carliner, S. (2005). *Advanced web-based training strategies: Unlocking instructionally sound online learning*. John Wiley & Sons.

Friesen, N. (2012). Report: Defining blended learning. Retrieved from http://learningspaces.org/papers/Defining_Blended_Learning_NF.pdf

Garrison, D. R., & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. *The Internet and Higher Education*, *7*(2), 95–105. doi:10.1016/j.iheduc.2004.02.001

Graham, C. R. (2006). Blended learning systems. In The handbook of blended learning (pp. 3-21).

Greenhalgh, T., Robert, G., Bate, P., Kyriakidou, O., & MacFarlane, F. (2005). Diffusion of innovations in health service organizations: A systematic literature review.

Jahng, N., Krug, D., & Zhang, Z. (2007). Student achievement in online distance education compared to face-to-face education. European Journal of Open, Distance and E-Learning, 10(1).

Jia, J., Chen, Y., Ding, Z., & Ruan, M. (2012). Effects of a vocabulary acquisition and assessment system on students' performance in a blended learning class for English subject. *Computers & Education*, *58*(1), 63–76. doi:10.1016/j.compedu.2011.08.002

Joanna Briggs Institute. (2014). JBI Data extraction form for experimental/observational studies [form]. Retrieved from http://joannabriggs.org/assets/docs/jbc/operations/dataExtractionForms/JBC_Form_DataE_ExpObs.pdf

Karam, R., Pane, J. F., Griffin, B. A., Robyn, A., Phillips, A., & Daugherty, L. (2016). Examining the implementation of technology-based blended algebra I curriculum at scale. *Educational Technology Research and Development*, 1–27.

Kazu, I. Y., & Demirkol, M. (2014). Effect of Blended Learning Environment Model on High School Students' Academic Achievement. TOJET: The Turkish Online Journal of Educational Technology, 13(1).

King, A. (2002). Structuring peer interaction to promote high-level cognitive processing. *Theory into Practice*, *41*(1), 33–39. doi:10.1207/s15430421tip4101_6

Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The Journal of Economic Education*, *31*(1), 30–43. doi:10.1080/00220480009596759

Lee, C., Cheung, W., Wong, K., & Lee, F. (2013). Immediate web-based essay critiquing system feedback and teacher follow-up feedback on young second language learners' writings: An experimental study in a Hong Kong secondary school. *Computer Assisted Language Learning*, *26*(1), 39–60. doi:10.1080/09588221.2011.630672

Lou, Y., Abrami, P. C., & d'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research*, *71*(3), 449–521. doi:10.3102/00346543071003449

Machtmes, K., & Asher, J. W. (2000). A meta-analysis of the effectiveness of telecourses in distance education. *American Journal of Distance Education*, *14*(1), 27–46. doi:10.1080/08923640009527043

Means, B., Toyama, Y., Murphy, R., & Baki, M. (2013). The effectiveness of online and blended learning: A meta-analysis of the empirical literature. *Teachers College Record*, *115*(3), 1–47.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies*. US Department of Education.

Mulqueeny, K., Kostyuk, V., Baker, R. S., & Ocumpaugh, J. (2015). Incorporating effective e-learning principles to improve student engagement in middle-school mathematics. *International Journal of STEM Education*, *2*(1), 1. doi:10.1186/s40594-015-0028-6

Naidoo, N., & Naidoo, R. (2011). Using blended learning to facilitate the mathematical thought processes of primary school learners in a computer laboratory: A case study in calculating simple areas. *Journal of College Teaching and Learning*, *4*(7). doi:10.19030/tlc.v4i7.1568

Ontario Ministry of Education. (2016). E-Learning Ontario: Blended Learning. Ottawa, ON: Queen's Printer for Ontario. Retrieved from http://www.edu.gov.on.ca/elearning/blend.html

Pane, J. F., Griffin, B. A., McCaffrey, D. F., & Karam, R. (2014). Effectiveness of cognitive tutor algebra I at scale. *Educational Evaluation and Policy Analysis*, *36*(2), 127–144. doi:10.3102/0162373713507480

Petticrew, M., & Roberts, H. (2008). *Systematic reviews in the social sciences: A practical guide*. John Wiley & Sons.

Picciano, A. G., Seaman, J., Shea, P., & Swan, K.Sloan Foundation. (2012). Examining the extent and nature of online learning in American K-12 education: The research initiatives of the Alfred P. Sloan Foundation. *The internet and higher education*, *15*(2), 127–135. doi:10.1016/j.iheduc.2011.07.004

Rockman, S. (2007). Ed pace final report. San Francisco, CA: Rockman. Retrieved from www.rockman.com/projects/146.ies.edpace/finalreport

Schmid, R. F., Bernard, R. M., Borokhovski, E., Tamim, R. M., Abrami, P. C., Surkes, M. A., & Woods, J. et al. (2014). The effects of technology use in postsecondary education: A meta-analysis of classroom applications. *Computers & Education*, *72*, 271–291. doi:10.1016/j.compedu.2013.11.002

Shachar, M., & Neumann, Y. (2003). Differences between traditional and distance education academic performances: A meta-analytic approach. *The International Review of Research in Open and Distributed Learning*, *4*(2). doi:10.19173/irrodl.v4i2.153

Sitzmann, T., Kraiger, K., Stewart, D., & Wisher, R. (2006). The comparative effectiveness of web-based and classroom instruction: A meta-analysis. *Personnel Psychology*, *59*(3), 623–664. doi:10.1111/j.1744-6570.2006.00049.x

Smith, N. V. (2013). Face-to-face vs. Blended Learning: Effects on Secondary Students 'Perceptions and Performance. *Procedia: Social and Behavioral Sciences*, *89*, 79–83. doi:10.1016/j.sbspro.2013.08.813

Smith, V. (2014, November 5). What is blended learning? Retrieved from http://www.universityaffairs.ca/features/feature-article/blended-learning/

Staker, H. (2011). *The Rise of K-12 Blended Learning: Profiles of Emerging Models*. Innosight Institute.

Sudha, A., & Amutha, S. (2015). Higher secondary learners' effectiveness towards web-based instruction (WBI) on chemistry. *Universal Journal of Educational Research*, *3*(7), 463–466. doi:10.13189/ujer.2015.030706

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285. doi:10.1207/s15516709cog1202_4

Torgerson, C. (2003). *Systematic reviews*. Bloomsbury Publishing.

Ungerleider, C., & Burns, T. C. (2002, April). Information and communication technologies in elementary and secondary education: A state of the art review. In Actes du Colloque 2002 du Programme pancanadien de recherche en éducation (PPRE): La technologie de l'information et l'apprentissage (pp. 2-28).

Wendt, J. L., & Rockinson-Szapkiw, A. (2014). The effect of online collaboration on middle school student science misconceptions as an aspect of science literacy. *Journal of Research in Science Teaching*, *51*(9), 1103–1118. doi:10.1002/tea.21169

Wendt, J. L., & Rockinson-Szapkiw, A. J. (2015). The effect of online collaboration on adolescent sense of community in eighth-grade physical science. *Journal of Science Education and Technology*, *24*(5), 671–683. doi:10.1007/s10956-015-9556-6

Werth, E., Werth, L., & Kellerer, E. (2013). Transforming K-12 Rural Education through Blended Learning: Barriers and Promising Practices. International Association for K-12 Online Learning.

Williams, S. L. (2006). The effectiveness of distance education in allied health science programs: A meta-analysis of outcomes. *American Journal of Distance Education*, *20*(3), 127–141. doi:10.1207/s15389286ajde2003_2

Yapici, İ. Ü., & Akbayin, H. (2012). The effect of blended learning model on high school students' biology achievement and on their attitudes towards the internet. TOJET: The Turkish Online Journal of Educational Technology, 11(2).

Zhao, Y., Lei, J., Yan, B., Lai, C., & Tan, H. S. (2005). What makes the difference? A practical analysis of research on the effectiveness of distance education. *Teachers College Record*, *107*(8), 1836–1884. doi:10.1111/j.1467-9620.2005.00544.x

APPENDIX

Table 1. Positive and negative search terms included in the literature search

Positive Search Terms*	Negative Search Terms*
"blended learning" "blended class" "blended environment" "blended program" "blended instruction" "flipped" "flipp*" "hybrid learning" "hybrid program" "hybrid class" "K-12" "kindergarten" "primary education" "secondary education" "primary school" "secondary school"	"higher learning" "distance learning" "adult learners" "higher education" "undergraduate*" "distance learning" "occupation*" "professional"

* Included and excluded terms were combined with Boolean operators (i.e., AND, OR, NOT, AND NOT) and wildcards (*) in various permutations.

Mark Poirier is a primary school educator who has been working in education for 14 years. He has studied and been awarded degrees from Trent University, the University of Ottawa and KU Leuven.

Jeremy Law is a lecturer of Education at the School of Interdisciplinary Studies at the University of Glasgow.

Anneli Veispak is currently a freelance writer, supervisor, editor and a lecturer collaborating with University of Leuven (KU Leuven), University of Glasgow and Tallinn University.