Journal of Peer Learning

Volume 12 | Article 2

Spring 2019

A meta-analysis on the effect of duration, task, and training in peer-led learning

Shenghua Zha
University of South Alabama, shzha@southalabama.edu

Michele D. Estes
James Madison University, estesmd@jmu.edu

Ling Xu
James Madison University, xulx@jmu.edu

Recommended Citation
Zha, Shenghua; Estes, Michele D.; and Xu, Ling, A meta-analysis on the effect of duration, task, and training in peer-led learning, *Journal of Peer Learning*, 12, 2019, 5-28.
Available at:https://ro.uow.edu.au/ajpl/vol12/iss1/2

Journal of Peer Learning (2019) Vol 12: 5–28

A meta-analysis on the effect of duration, task, and training in peer-led learning

Shenghua Zha, Michele D. Estes, and Ling Xu

ABSTRACT

This meta-analytic study compared the effect of peer-led learning versus non peer-led learning on students' cognitive achievement in post-secondary education. Twenty-eight studies published in English from six countries between 1993 and 2017 were identified and used in the analysis. Result of the analysis on the random-effect model showed a moderate but positive effect, meaning that peer-led learning was associated with higher cognitive achievement than non-peer-led learning. Three study characteristics were examined including duration, student leaders' training, and task type. Only the task type was found significant in moderating the effect of peer-led learning. Student leaders' facilitation of problem-based learning tasks outperformed other types of tasks. Results of this study not only provided suggestions for peer-led learning designers and coordinators but also called for future research of student leaders' readiness as well as online peer-led learning.

INTRODUCTION

Postsecondary educators have used peer-led learning (PLL) extensively in academic and extracurricular settings. PLL is a type of small group learning activity where student leaders, of similar status to the learners, guide or facilitate learning within a group (Eberlein et al., 2008; Ender & Kay, 2001; Topping, 2005). Instructors may select student leaders from a current or recent course, or student leaders may emerge from within a course (Quitadamo, Brahler, & Crouch, 2009). Student leader responsibilities vary; for example, while some facilitate discussions, others may take the role of a lecturer (Cuseo, 2010).

Over the last several decades, PLL has evolved as an effective pedagogical practice. In 1973, Deanna Martin developed a PLL program called Supplemental Instruction, where senior student leaders were hired to help students through difficult courses (Martin & Arendale, 1992). In 1981, this program was endorsed by the U.S. Department of Education for its effectiveness on students' final course grades and on retention in high-risk courses (Dawson et al., 2014). In 1993, the City College of New York implemented a similar program called Peer-Led Team Learning. In this program, senior students assisted undergraduate students in the class (Smith et al., 2014). Since then, PLL programs have been funded and disseminated nationwide (Dreyfuss, 2012). In 2008, a Peer-Led Team Learning program received the James Flack Norris Award for Outstanding Achievement in Teaching Chemistry (Gosser, Kampmeier, & Varma-Nelson, 2010).

According to reviews by Dreyfuss (2012) and Dawson (2014), the increasing and international reputation of PLL resulted in exponential growth in the number of individual studies conducted. Hooker (2010) conducted a narrative review of Peer-Led Team Learning studies administered between 1995 and

2003. Dawson et al. (2014) reviewed studies of Supplemental Instruction published between 2001 and 2010. Each confirmed the positive impact of PLL on student course grades and retention. PLL programs were shown to benefit learners while student perception was also favourable, and most students reported the PLL experience as engaging and productive.

While narrative review was used in these studies, meta-analysis is considered a more rigorous method for identifying the overall effect of one treatment (Glass, 1976). Meta-analysis as a statistical approach synthesizes results from individual quantitative studies that examine a similar treatment but often reach different conclusions regarding the power of the treatment. The meta-analysis summarizes and identifies the direction and magnitude of an overall effect of a treatment across studies (Borenstein et al., 2009). This method is especially helpful when attempting to identify small, yet consistent, effects from individual studies (Kulik & Kulik, 1982).

Researchers examined K-12 student academic achievement in relation to PLL activities (Ginsburg-Block, Rohrbeck, and Fantuzzo, 2006; Rohrbeck, et al., 2003). Findings in the two meta-analytic studies were consistent with findings from the narrative reviews shared previously in this paper. For example, Ginsburg-Block, Rohrbeck, and Fantuzzo (2006) found that PLL had an overall positive effect on student academic achievement when compared with students in non-PLL activities. Specifically, the academic achievement of students in PLL activities was about 0.48 standard deviations higher than that of non-PLL students.

This medium effect could not be determined by merely reading individual studies with mixed results and is one advantage of using meta-analysis (Cohen, 1977). Another advantage is that this method enables researchers to relate study characteristics to the effect across studies (Bernard et al. 2004). For example, Cohen, Kulik, and Kulik (1982) found through meta-analysis that structured PLL programs, and shorter PLL session duration, were associated with positive and greater effect than un-structured programs and programs with longer duration.

In the past decade, just two meta-analytic studies were conducted on general PLL, and each was focused on K-12 education (Ginsburg-Block et al., 2006; Rohrbeck et al., 2003). These studies provided insights into the overall effect of PLL and the efficacy of some characteristics in the design of PLL for K-12 education. PLL in post-secondary education, on the other hand, has not been thoroughly investigated even though some well-known models such as Supplemental Instruction and Peer-Led Team Learning have been widely adopted and examined in colleges and universities since the 1990s.

In this article, we review the design framework of PLL in particular and the design of learning tasks that serve as the foundation of the three variables, or characteristics, examined in this study. The three characteristics are (1) the types of tasks required for student leaders to perform in PLL, (2) training received by student leaders, and (3) duration of an average PLL session. The authors collected and analysed 28 studies, published between 1993 and 2017 in six countries, in the multilevel analysis. Near the end of this paper, we discuss the results and implications for instructional design and future research in the PLL area.

LITERATURE REVIEW

PLL is grounded in Piaget's social learning theory, from which it gathers significance and versatility. Piaget (1932) believed that children communicate and learn better with their peers than with teachers who hold authoritative roles. Peer interaction may be used to generate cognitive conflicts that, in turn, help learners reconstruct and develop social and intellectual competence (Piaget, 1932). With regard to social identity theory, Tajfel and Turner (1979) proposed that an individual's self-image is affected by the group to which he or she is attached. If a person is in a group where a positive association is developed, he or she will develop a positive self-image and high self-esteem. As such, peer groups have been considered a beneficial support for students who are experiencing physical, emotional, and cognitive adjustment (Tarrant, 2002).

While the research of Piaget (1932) and Tajfel and Turner (1979) focused on children and adolescent development, Schmidt and Moust (1995) studied PLL among post-secondary students. Using social congruence theory, the researchers found that student leaders demonstrated both cognitive and social congruence with peers who have similar knowledge structures, learning pathways, and social roles. These students had mutual understandings and were able to communicate with other students easily. This finding is critical for student leaders and instructors who seek to identify and address difficult learning problems. Results of this study, and following studies, confirm that social and cognitive congruence have a positive yet indirect impact on student academic achievement (Lockspeiser et al., 2008; Schmidt & Moust, 1995).

Tasks

Social learning theory, social identity theory, and social congruence theory support PLL from a social learning perspective. However, these theories lack guidelines for the design of the PLL backbone: learning tasks. At what cognitive level should a PLL learning task be designed in order to engage students? Further, what tasks should student leaders perform to help students increase achievement? While it may be worthwhile to examine a variety of tasks that support higher-order thinking, this paper will focus on one popular method used in some PLL sessions called problem-based learning (PBL).

In its most comprehensive form, PBL is interwoven throughout the curriculum and scaffolded as any pedagogical intervention. When using PBL in class, instructors or student leaders present a real-world problem, or problem simulation, to a learner group. Problems are designed at a particular level to engage students in higher-order thinking. Students work in the group, or as individuals, to identify and solve the given problem by applying what they know. The instructor or student leader facilitates this process.

Studies have unanimously agreed on the positive effect of PBL on student motivation and attitude (Prince, 2004). In addition, Hung, Jonassen, and Liu (2008) conducted a narrative review of PBL and reported its positive impact on student application of knowledge. However, the researchers did not find a consensus regarding the impact of PBL on student acquisition of knowledge, especially conceptual knowledge.

In a qualitative review of seven meta-analyses, Strobel and van Barneveld (2009) found that PBL benefitted student skill development. Results of their analysis on student knowledge acquisition clarified mixed results from Hung et al. (2008) and showed that PBL helped students' long-term knowledge retention but not immediate knowledge acquisition and retention.

PLL has been widely implemented in foundational undergraduate courses to decrease the historically high drop-fail-withdraw rate (Kokkelenberg, Dillon, & Christy, 2008; Lee & Choi, 2010). Conceptual knowledge is essential for success in these courses. In many PLL sessions, student leaders have been required to engage students in conceptual learning tasks such as repeating the lecture and reviewing the notes (Batchelder et al., 2010; Ogden et al., 2003).

PBL can deepen this activity, as it is associated with more complex learning tasks involving higher-order thinking. PBL has the potential to engage students in cognitive conflict and challenge, which is strongly recommended in PLL design by Topping (1996, 2005). Studies have shown that when students are engaged in higher-order cognitive activities, they revisit and correct relevant conceptual knowledge (Cracolice & Busby, 2015; Jensen et al., 2014). As a result, the engagement improves their understanding of conceptual knowledge. This conclusion is supported by results of both PBL reviews (Hung et al., 2008; Strobel & van Barneveld, 2009), especially the latter.

It is important to note that long-term knowledge retention in these studies was measured after a period of 12 weeks to two years. Some of the PLL programs were conducted in regular semesters when other relevant courses were offered. Student learning outcomes were usually measured at the end of a course, a time not immediately after a PLL activity, nor 12 weeks after a PLL session was delivered. How did a task that student leaders implemented in PLL affect student comprehensive achievement at the end of a course? Did problem-based learning work well in PLL activities? While individual studies present mixed results, a meta-analytic study could help identify an overall effect and give definitive answers to these questions.

Training

Hung et al. (2008) noted that it is a challenge to design PBL, suggesting that there is a need for PLL course instructors to train and help student leaders. Additionally, Topping (1996, 2005) pointed out that PLL puts a heavy demand on student leaders. They need to be good at detecting, diagnosing, and managing learner misconceptions, and verbalizing and communicating with students effectively. To do this, student leaders must have expertise in subject content as well as facilitation and interaction skills (Attarzadeh et al., 2011; Skalicky & Caney, 2010). Results of PLL studies on student performance and reflection confirm the need for these kinds of proficiencies (Neville, 1999; Secomb, 2008).

Salas, Nichols, and Driskell (2007) conducted a meta-analysis of non-academic small group work studies. Results indicated that group performance was better when group members received training than when they did not. In many PLL white papers, especially those with brand names such as Supplemental Instruction and Peer-Led Team Learning, student leader training was underscored and described in detail.

However, recent studies of PLL student leader training tend to focus on either leader prior experience or subject-matter expertise (Ahmed, Elseed, & El-Sheikh, 2007; Zumbach & Spraul, 2007). For example, in a study of 206 third-year dental students, sessions facilitated by experienced student leaders did not significantly outperform those facilitated by inexperienced student leaders (Park et al., 2007). In another study of two computer science classes from a pre-university secondary school, when skilled leaders helped students, the students were found to have lower cognitive load and better cognitive performance than those helped by student leaders without facilitation skills (Hsiao et al., 2012).

In a majority of PLL programs or activities, student leaders are required to attend orientations to understand the purpose and expectations, and to gain basic facilitation skills before they conduct their first PLL session or activity (Attarzadeh et al., 2011). In some PLL programs, student leaders also attend weekly meetings with instructors and pedagogy experts to improve their competencies (Mitchell, Ippolito, & Lewis, 2012). It takes institutions, departments, and course instructors much time and effort to plan and coordinate these trainings. Therefore, it is imperative to understand how theses trainings prepare student leaders and affect student achievement.

Duration

Another key factor to consider in the PLL task design is the duration of a learning session. Studies have shown that duration is closely related to student engagement. For example, in a review of literature, Wilson and Korn (2007) found that students retain most content learned in the first 10, and last 20, minutes of a lecture class. Reardon et al. (2008) surveyed 467 undergraduate business students at four universities and found that students preferred shorter and more frequent class sessions, and they believed this format resulted in better learning outcomes than longer and less frequent class sessions. These studies provide evidence that post-secondary students engage at specific points in time and that attention decreases in learning sessions with long durations.

PLL programs tend to use durations typically found in a regular lecture class, ranging from 50 to 180 minutes per session. There is little empirical evidence on how long an average PLL session should last to better engage students in learning tasks and thereafter improve their achievement. Cohen, Kulik, and Kulik (1982) in a meta-analysis of PLL in K-12 education examined the overall length of PLL programs. They found that PLL sessions with a total of zero- to four-week sessions were associated with significantly greater effect than sessions lasting more than four weeks. However, they did not examine how long an average PLL session lasted in those PLL studies or how that duration variation impacted student learning.

RESEARCH QUESTIONS

The purpose of this meta-analysis was to review studies in post-secondary education published between 1993 and 2017 and to identify the effect of PLL on students' cognitive achievement. The research questions were:

1. In general, what was the effect of PLL on student cognitive achievement? Did students have better cognitive achievement in PLL than non-PLL?

2. How were the following three characteristics associated with the direction and magnitude of the effect of PLL: duration of an average PLL session, student leaders' training, and tasks?

METHODS

Literature search

A comprehensive search was conducted in the academic databases or search engines including PsycINFO, ERIC, EBSCO, SCOPUS, ProQuest Digital Dissertations, and Google Scholar. The following keywords and the Boolean logic were used in the search process:

> *Keywords* = (peer *OR* student *OR* learner)
>
> *AND*
>
> *Keywords* = (led *OR* lead *OR* leader *OR* tutor *OR* teacher *OR* mentor *OR* facilitator *OR* facilitate *OR* facilitation *OR* moderator *OR* moderate)

In addition, we searched the systematic review and annotated bibliography articles (Arendale, 2014; Arendale, 2017). References in the resulting studies were checked so as to include those articles that did not have the keywords but did investigate the effect of PLL. We sent e-mails to the authors of studies where there was insufficient data either for the three study characteristics or to calculate effect sizes. If no responses were received and there was no other way to acquire the data needed for our analysis, those articles were dropped from the study.

Source inclusion and exclusion criteria

Studies included in this meta-analysis were selected based on the following criteria:

1. They were written in English.
2. They were available to the public between 1993 and 2017.
3. Studies were situated in postsecondary education.
4. They were quantitative studies that compared the effects of PLL (experimental group) versus non-PLL (control group). Non-PLL included traditional instructor-led learning, or expert-led learning, or traditional individual work in class. As a result, studies that compared only different design and delivery of PLL were excluded. In addition, studies that measured only pre- and post-performance in PLL were also excluded.
5. They measured students' cognitive achievement in PLL and non-PLL. For example, studies that measured only students' psychomotor performance in a lab course were removed from the search results.
6. They focused on PLL where senior students had distinct facilitation, tutoring, or coaching roles. As a result, studies of the branding programs such as Supplemental Instruction and Peer-Led Team Learning were included in this study. At the same time, studies of other types of peer-assisted learning where no students took such roles were not included.

7. They focused on PLL activities or programs that occurred regularly in a course period. We believed that PLL activities that happened only once in a course were not enough to generate a longitudinal effect on students' achievement in a course.

Firstly identified were 1,074 research articles published in English between 1993 and 2017. Articles were then examined to meet Criteria 3–7. As a result, 30 articles were selected, among which two were focused on online PLL. Considering the differences of organization and structure between face-to-face and online PLL, we dropped the two online PLL studies. At the end of the search and screening process, 28 individual studies from six countries were collected and kept for analysis (Table 1).

Table 1
Summary information of studies

Types of publication:	Number of studies
Dissertation/thesis	3
Conference proceeding	1
Journal	24
Years of publication:	**Number of studies**
1994–1999	1
2000–2004	4
2005–2009	6
2010–2017	17
Disciplinary areas	**Number of studies**
Mathematics	7
Chemistry	7
Biology	6
Engineering	1
Health science and medicine	5
Law	1
Psychology	1
Modes of learning	**Number of studies**
Face-to-face	28
Online	0
Course levels	**Number of studies**
Undergraduate lower-level	24
Undergraduate upper-level or graduate	4
Countries	**Number of studies**
Australia	2
Canada	1
Germany	1
Netherlands	1
UK	1
U.S. and U.S. territories	22

Coding procedure

Two raters, who were experts in qualitative coding, identified and separately coded 12 studies. An inter-rater reliability was calculated, which proved to be greater than 90%. The raters met and resolved disagreements via discussion, then developed a codebook based on the outcomes of the discussion. Thereafter, one rater coded the remaining studies using the codebook criteria and presented coding questions to the other rater. Discussions occurred until a consensus was reached. The individuals coded variables, or study characteristics.

Characteristics of individual studies

Three characteristics were examined in this meta-analysis. They were (1) types of tasks that student leaders performed in PLL activities or sessions, (2) student leader training, and (3) duration of an average PLL session (Table 2).

The types of tasks that student leaders were required or supposed to perform in PLL activities were coded based on whether the core task was to facilitate student learning through the use of PBL activities. Studies where student leaders used PBL and had learners involved in a complete PBL process were coded as 1. For example, with the aid of student leader scaffolding, students identified and solved problems and presented solutions either individually or in groups. Studies where student leaders performed only non-PBL tasks were coded as 0. These tasks included lecturing, reviewing notes, and managing role-play activities.

The variable of student leader training was coded into three values: (a) sufficient training coded as 2, (b) partial training coded as 1, and (c) no training coded as 0. Sufficient training meant that student leaders not only received formal training before their first PLL sessions, but they also met with instructors and pedagogy experts regularly to improve their facilitation skills during the semester when PLL sessions were conducted. If only one training was given to student leaders, that study was coded as partial training. If a study had a detailed description of the preparation for student leaders but did not indicate any training was offered to them, it was coded as no training.

Duration of an average PLL session referred to the number of minutes each PLL session was hosted at one time. These were quantitative values directly quoted from the authors of each study.

Outcome variable

The outcome of this study, cognitive achievement, was usually measured on examinations after the interventions were introduced in each study. However, the measurement of those examinations varied among studies. For example, some studies used course grades and some studies used national standardized tests. Therefore, effect sizes (ESs) were computed within each study to make the outcome comparable across studies.

Table 2
Characteristics from individual studies

Studies	Sample size	ES	Task type	Student leaders' training	Duration of a PLL session
Akinyele, 2010	534	0.65	1	2	120
Báez-Galib, Colón-Cruz, Resto, & Rubin, 2005	1,849	0.31	1	2	180
Batchelder et al., 2010	358	–0.07	0	1	120
Born, 2000	116	0.95	1	2	120
Dancer, Morrison, & Tarr, 2015	751	0.55	1	2	60
Fayowski & MacMillan, 2008	1,259	0.47	0	2	50
Guarcello et al., 2017	598	0.07	0	1	90
Hensen & Shelley, 2003	3,009	0.21	0	2	90
Hizer, Schultz, & Bray, 2017	404	0.43	0	1	60
Hockings, DeAngelis, & Frey, 2008	1,125	0.35	1	1	120
Hooker, 2010	83	0.41	1	2	60
Hryciw, Tangalakis, Supple, & Best, 2013	483	0.41	0	2	60
Hughes, 2011	132	0.93	1	2	60
Kalil, Jones, & Nast, 2016	110	1.02	1	2	60
Knobe et al., 2010	151	–0.01	0	0	120
Lewis & Lewis, 2008	1,740	0.13	0	1	50
Lewis, 2011	1,931	0.31	1	1	50
Loui, Robbins, Johnson, & Venkatesan, 2013	720	0.33	1	2	90
Moust & Schmidt, 1994	407	–0.12	0	1	120
Preszler, 2009	2,909	0.31	0	2	65
Quitadamo, Brahler, & Crouch, 2009	1,382	0.20	0	2	120
Reisel, Jablonski, Munson, & Hosseini, 2012	1,382	0.35	1	2	60
Reisel, Jablonski, Munson, & Hosseini, 2014	682	0.38	1	2	60
Schneiders, 2014	58	0	0	0	100
Shields et al., 2012	262	0.81	1	1	120
Snyder, Carter, & Wiles, 2015	104	0.50	1	2	60
Steele, Medder, & Turner, 2000	673	0.15	0	1	120
Tien, Roth, & Kampmeier, 2002	1,809	0.63	1	1	120

A positive ES indicated that students participating in PLL had higher cognitive achievement scores than students participating in non-PLL activities or sessions. A negative ES meant that student cognitive achievement scores after PLL were lower than scores of those who experienced non-PLL. Hedges' *g* was used in the calculation of ES to reduce the bias in Cohen's *d* due to a wide spectrum of sample sizes in individual studies (Ginsburg-Block et al., 2006; Hedges & Olkin, 1985). In studies where means and standard deviation were not reported, ES was calculated using other statistical data, such as *t* and *F* values (Rosnow & Rosenthal, 1996).

To comply with the assumption of independence in meta-analysis, we took the following actions when data from multiple groups were reported in one study (Ahn, Ames, & Myers, 2012). First, if a summary of data was already provided in the study, such as total mean and total standard deviation, the summary data were used and one ES was calculated. Second, if multiple experimental groups were compared to one control group, data from the experimental groups were merged into one large experimental group and an ES was calculated accordingly (Headrick, 2010). Third, if multiple pairs of experimental and control groups were reported on the sub-group level, such as courses, ES and variance were calculated for each pair and then merged into one ES and one variance for each study using the calculation introduced by Borenstein, Hedges, Higgins, and Rothstein (2009).

ANALYSIS

Bias assessment

Publication bias has been found in many fields, in particular, the field of social science (Peplow, 2014; Rothstein, Sutton, & Borenstein, 2005). It occurs when research with significant and positive results are more likely to be published than research with non-significant or negative results. This selective publication poses a threat to the validity of meta-analysis studies. Egger's regression was used to detect potential publication bias from the collected studies in this meta-analysis (Egger et al., 1997).

Structure of multilevel meta-analysis

Data in meta-analyses are hierarchically structured by nature. PLL may be structured the same way in one study; hence, effects are nested within studies. Two-level meta-analysis, where a composite effect size is typically calculated for each study to meet the assumption of independence, has been widely used in the last 20 years (Raudenbush & Bryk, 2002; Teo, 2011). Like regression analysis, all characteristics are taken into consideration at one time in this study. It allows us to examine the impact of a set of characteristics and how those characteristics relate to each other (Sheu & Suzuki, 2001). As a result, a two-level analysis was employed using the open source *metafor* package for the R statistical environment version 3.4 in Microsoft Windows 7. This package allows the running of the multilevel random-effects model. It is considered the most flexible package running meta-analysis in R (Polanin, Hennessy, & Tanner-Smith, 2016).

The Level 1 model took the ES in each individual study as a result of an estimate of its true ES and sampling error.

Equation 1

$$g_k = \delta_{0k} + e_k, \quad e_k \sim N(0, \sigma^2)$$

where g_k meant the calculated ES in Study *k*. δ_{0k} meant a population ES of Study *k*. e_k meant the random sampling error in Study *k*. In meta-analysis, the sampling variance (σ^2) was known.

δ_{0k} in the Level 1 model became the outcome variable in the Level 2 model, the latter of which examined variation across studies and the potential effect of cross-study characteristics on the variation.

Equation 2

$$\delta_{0k} = \gamma_{00} + \gamma_{01}Task_{0k} + \gamma_{02}Training_{0k} + \gamma_{03}Time_{0k} + u_{0k}, \quad u_{0k} \sim N(0, \tau_\beta)$$

where δ_{0k} meant a population ES of Study *k*. γ_{00} meant the overall ES across studies. γ_{01} , γ_{02} , and γ_{03} were the coefficients of the variables of task types, student leaders' training, and duration of a PLL session respectively. u_{0k} was the extent to which population ES of Study *k* varied from the overall ES. τ_β was the between-study variation in ES that remains once controlling for all characteristics in the model.

A combination of Equations 1 and 2 yielded Equation 3.

Equation 3

$$g_k = \gamma_{00} + \gamma_{01}Task_{0k} + \gamma_{02}Training_{0k} + \gamma_{03}Time_{0k} + e_k + u_{0k}, \quad u_{0k} \sim N(0, \tau_\beta)$$

where g_k meant the calculated ES in Study *k*, γ_{00} meant the overall ES across studies. γ_{01} , γ_{02} , and γ_{03} were the coefficients of the variables of task types, student leaders' training, and duration of a PLL session respectively. e_k and u_{0k} referred to sampling errors at the study and cross-study levels.

To answer Research Question 1, an unconditional random-effects model was used to estimate the overall ES and the homogeneity of ES across studies before the proposed cross-study characteristics were taken into consideration. This model viewed the observed ES of each study varied around a common ES plus random effects.

Equation 4

$$g_k = \gamma_{00} + e_k + u_{0k}$$

where g_k meant the calculated ES in Study *k*. γ_{00} meant the overall ES across studies, and e_k and u_{0k} referred to sampling errors at the study and cross-study levels.

To answer Research Question 2, a conditional mixed-effects model with restricted maximum likelihood as depicted in Equation 3 was deployed to examine the association of cross-study characteristics with the ESs. The rationale of using the mixed-effects model was that firstly, one or more cross-study characteristics might account for some portion of the heterogeneity in true ESs (Hedges & Vevea, 1998). Secondly, the heterogeneity in true ES may also result from subject area and sampling differences, which was acknowledged but not an interest of this study.

RESULTS

Results showed that the intercept value of Egger's regression was positive (z = 1.57), which indicated a slight chance that studies with smaller sample sizes had higher levels of test accuracy. However, this asymmetry was not statistically significant (p = 0.11). Hence, we concluded that no significant evidence of publication bias was found from the studies collected for the analysis.

Overall effect

Overall, 28 ESs generated from 28 studies were used in the analysis (Figure 1). Results of the unconditional model showed that the average ES across studies was 0.36. In other words, the average cognitive achievement of PLL groups was 0.36 standard deviation units higher than that of non-PLL groups. According to Cohen (1977), this is a moderate effect. A calculation of I^2 showed that 89.7% variance in observed effects reflected variance in true effects rather than sampling error. The 95% confidence interval for τ ranged from 0.26 to 0.46, indicating that τ is non-zero and it was worth exploring characteristics across studies.

Impact of study characteristics

Q_M statistics was used to test whether all of the three regression coefficients were equal to zero (H_0: $\gamma_{01} = \gamma_{02} = \gamma_{03} = 0$) (Pastor & Lazowski, 2018). The result for the conditional model showed a statistical significance (Q_M= 24.12, df = 3, $p < 0.01$), meaning that this set of characteristics was a practically significant predictor of ES differences across studies. More than half (53%) of the between-study variability in true ES can be explained by this conditional model.

In this conditional model, the task type was the only variable that significantly moderated the ES after controlling for the two other variables (Table 3). Groups where student leaders facilitated PBL activities as their core tasks were 0.30 standard deviation units higher than groups where student leaders facilitated non-PBL types of tasks, such as lecturing. In other words, PLL sessions where student leaders facilitated a complete PBL process were more effective, with regard to supporting student cognitive achievement, than those sessions where student leaders did not facilitate using a complete PBL process. However, the width of the confidence interval was large, suggesting that the magnitude of the effect could be as small as 0.15 or as large as 0.45.

On average, groups where leaders participated in training were associated with a greater ES than groups where leaders did not participate in training (Table 3). Among 28 studies, only two omitted a description of any training that student leaders might have received, which very likely suggests that no training was offered. Groups where leaders received training before or during the semester when they hosted the PLL sessions were associated with higher ES than those who did not receive any training. Likewise, groups where student leaders received formal training at the beginning, as well as regular meetings to improve their facilitation skills during the semester, outperformed those where leaders received only one type of training. However, none of these mean differences showed statistical significance.

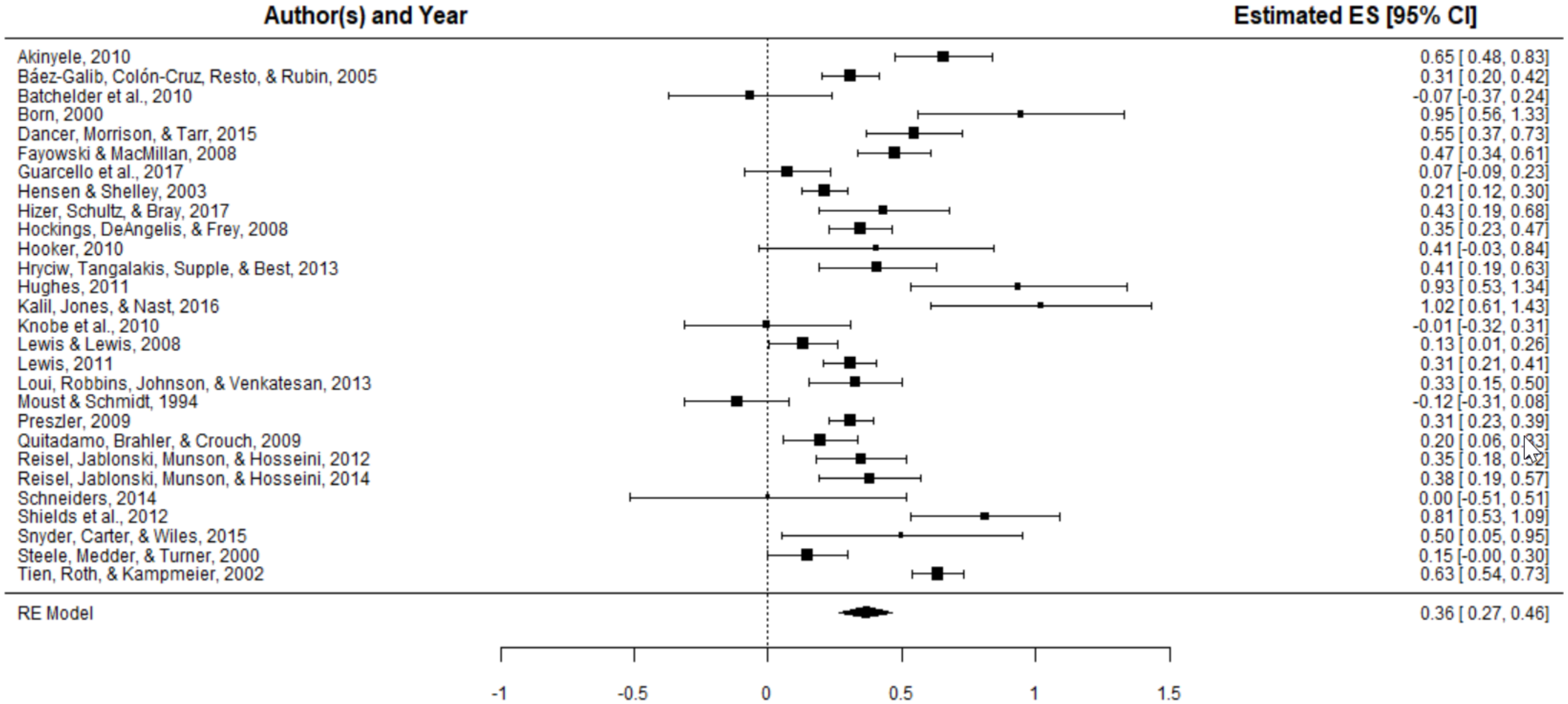

Figure 1. Forest plot for the ESs of studies

Table 3
*Coding and mean effect sizes of the study characteristics (*k *= 28)*

Characteristics	***k***	***M***	**γ**	***p***	***se***	**95% CI**
Types of tasks			0.30	0.00	0.08	(0.15, 0.45)
0: Leaders facilitated PBL as their core tasks	13	0.17				(0.05, 0.29)
1: Leaders performed a non-PBL task as their core tasks	15	0.56				(0.43, 0.71)
Student leader training			0.11	0.11	0.07	(–0.02, 0.24)
0: Leaders did not have any training before PLL starts	2	–0.003				(–0.04, 0.03)
1: Leaders went through a training before PLL started or joined additional training regularly during the semester when PLL sessions were offered.	10	0.27				(0.06, 0.48)
2: Leaders took a training before PLL started and joined additional training regularly during the semester when PLL sessions were offered.	16	0.50				(0.36, 0.64)
Duration of a PLL session			–0.001	0.31	0.001	(–0.003, 0.001)
50 minutes	3	0.30				
60 minutes	9	0.55				
65 minutes	1	0.31				
90 minutes	3	0.20				
100 minutes	1	0				
120 minutes	10	0.35				
180 minutes	1	0.31				

Note: *k* stands for the number of individual studies. *M* stands for the mean effect sizes. Γ stands for the coefficients of the regression. *Se* stands for standard error. *p* stands for the significance value of the coefficients. *95% CI* stands for 95% confidence interval.

No statistical significance was found regarding variations in the duration of an average PLL session (Table 3). Among 28 studies, a PLL session was typically hosted for 60 or 120 minutes. Groups in 60-minute sessions outperformed groups in 120-minute sessions by 0.2 standard deviation units. However, when examined across all durations, shorter sessions did not have significantly better ES than longer sessions.

In summary, this model revealed that in PLL sessions where student leaders facilitated problem-based learning tasks, student cognitive achievement was 0.30 standard deviation units higher than those sessions where leaders performed non-PBL tasks.

Equation 5

$$g_k = 0.14 + 0.30 * Task_{0k} + e_k + u_{0k}$$

Meanwhile, the remaining between-study variance was large in magnitude ($\tau^2 = 0.03$) and significantly different than zero ($Q_E = 101.25$, $df = 24$, $p < 0.001$), indicating that there were still substantial differences among studies in terms of ES that require explanation.

DISCUSSIONS

This study used a multilevel meta-analysis to unveil the general effect of PLL on student cognitive achievement. A moderate and positive effect was identified, which supported the benefit of using meta-analysis, acclaimed by Kulik and Kulik (1982). It meant that, in general, PLL was associated with improved cognitive achievement of students in post-secondary education.

Due to the fact that none of the 28 studies collected for this meta-analysis replaced traditional instructor-led learning with PLL entirely, our conclusion was that PLL was effective as a supplemental learning activity or program on student cognitive achievement. This echoed the positive findings from other narrative reviews in post-secondary education as well as prior meta-analyses in K-12 education (Cohen, Kulik, & Kulik, 1982; Dawson et al., 2014; Ginsburg-Block et al., 2006; Hooker, 2010; Rohrbeck et al., 2003).

When comparing the results from other meta-analytic PLL studies, we found that the magnitude of the ES in this study was smaller than the findings from Ginsburg-Block and Rohrbeck's studies in K-12 education (Ginsburg-Block, Rohrbeck, & Fantuzzo, 2006; Rohrbeck et al., 2003). While we acknowledge that the ES could be complicated by different measurements in individual studies, there is another possibility. Differences in K-12 and post-secondary learning environments, and learner characteristics, could have contributed to this variation. Further exploration and understanding of the reason for this difference could be helpful. For example, it may inform professionals in post-secondary education who want to design quality PLL activities or programs, particularly for first-year students experiencing the shift from high school to a college or university.

Types of tasks that student leaders facilitate or guide are the essence of a PLL design. In this study, we focused on student leader facilitation of problem-based learning (PBL) tasks and compared its effect with that of other task types. Findings showed that PBL tasks significantly moderated the ES of PLL on student cognitive achievement. Student leader facilitation of PBL worked much better than other types of tasks.

Conceptual knowledge is generally the focus of cognitive assessments at the undergraduate, introductory course level, and 86% of the papers analysed in this study examined activities in these types of courses (Table 1). Results from our meta-analysis of undergraduate postsecondary literature support some prior studies that showed the effectiveness of PBL in student knowledge retention. This could be due to students reading related chapters, or attending class lectures, before actually joining a PLL session. Another reason for this effectiveness may be that the PBL process develops learners to think at a more complex cognitive level than non-PBL tasks where learners are only expected to recall and understand concepts. Students engaging in PBL had to actively test and use their conceptual knowledge to identify and solve problems in the PLL sessions. In those sessions using non-PBL tasks, such as having students review class notes and content, student leaders focused on recall and understanding of conceptual knowledge, which could easily generate another form of recitation or lecture session.

In particular, two studies where student leaders were assigned teaching tasks had negative ES. When assigned such a role, they typically gave formal lectures with students passively listening to their explanation and receiving the information. Our findings indicated that peer teaching was less effective than teacher/expert teaching regarding their impact on students' cognitive learning. A possible reason for the negative ESs was that this type of direct teaching required student leaders to have a mastery of content. In addition, to perform well in the teaching task, student leaders need to have adequate teaching skills. This high demand of subject-matter expertise and teaching skills were not what student leaders were able to perform by the time when they facilitated PLL sessions.

Many white papers reported that student leader readiness to implement PLL tasks affected student achievement. Our results showed that there were only two studies that did not offer student leaders either formal or ongoing training. Student leaders in the other 26 studies had received either initial training before their first PLL sessions, ongoing training during the time they hosted PLL sessions, or both. We did not find that ES was significantly altered because of student leader training.

However, it was too early to conclude that trainings were ineffective given several factors. First, our review of literature has already underscored the need for student leader training in PLL. Second, only two studies did not include student leader training. This number was extremely small, making it difficult to reach statistical significance. Another factor was that the impact of training on student cognitive achievement was indirect, as the association between student leader training and student achievement was mediated by student leader tasks and student performance variables.

Experimental studies with large total and sub-group samples would help to identify the indirect impact of student leader training on student cognitive achievement. Given the current findings, helpful future studies may also focus on the transfer of learning from student leader training to the PLL context. It is important to understand how training effect is moderated by the strategies used to aid student leader transfer of learning in training. This would help to identify the optimal strategy for student leaders to learn and apply specific training content.

We examined the duration of a PLL session on the effect of student cognitive achievement yet failed to identify a significant relationship. Among the wide range of duration from 50 to 180 minutes, about 68% of PLL sessions were organized for 60- and 120-minute durations. Between them, groups in 60-minute sessions performed at 0.20 standard deviation units higher than groups in 120-minutes as evidenced by the effect on student cognitive achievement. This echoed the research findings in lecture-based classes that shorter sessions work better than longer sessions. However, this effect occurred only between those two duration types. Three studies organized each PLL session in 50 minutes, but their average ES was lower than that of 120-minute session. Therefore, we failed to support findings from other duration studies that shorter sessions worked better than longer sessions.

We also found that the remaining heterogeneity was still large after the three characteristics were taken into consideration. This suggests that other contextual or design factors need to be considered to fully explain the variation. For example, different measurements were used to determine cognitive achievement in each study. Some used course grades while others used national standardized tests. Although ES was calculated to make the achievement comparable across studies, the difference caused by those inconsistent measurements still existed. Future studies with a large sample size should examine this, and other factors, to try to explain the heterogeneity found in this study.

LIMITATIONS

It was noteworthy that all the individual studies used in this meta-analysis were conducted in face-to-face classrooms (Table 1). Although we found two articles about PLL in online and blended learning, we excluded them because it would complicate the results as different delivery modes could affect the ES. Since our search focused on experimental studies that compared the outcome of PLL with non-PLL learning sessions, results of our search echoed the claim made by other researchers that there is a lack of quasi-experimental or experimental PLL studies in online or blended leaning environments (Attarzadeh et al., 2011; Evans & Moore, 2013). More experimental research is needed in the online delivery mode.

The second limitation of this study was that our search strategy was restricted by the keywords and the availability and scope of online databases. We made efforts to overcome this issue by using all possible and similar words related to PLL and searching through references in the articles. Even so, it is possible that we missed articles not in those databases nor cited in collected articles. It is also possible that we missed those without any of the listed keywords that still met the selection criteria.

Another limitation is that when searching and analysing articles, some articles were excluded from the review. For example, studies with missing ESs or data to calculate ESs were dropped in the search process. Then studies with any missing variables of interest were excluded from the analysis due to the regression-type of analysis in this study. This shrinkage of sample size limited the power of this study to provide greater generalizability on the effect of PLL.

CONCLUSION

We collected 28 empirical PLL studies in post-secondary education. These were publicly available and published between 1993 and 2017. A meta-analysis was deployed using R, with two assumptions taken into consideration. First, the calculated effect of PLL when compared to non-PLL in individual studies was a combined result of its true effect size (ES) and sampling error. Second, the ES difference between studies was caused by the divergent design and delivery of PLL tasks. We selected three characteristics or variables that were critical in PLL design and delivery from our review of literature. Those characteristics were (1) types of tasks that leaders performed, (2) student leader training, and (3) the duration of an average PLL session.

Results of our study confirmed and extended findings from prior systematic reviews. We found that PLL had a moderate but positive effect on student cognitive achievement. In particular, student leader facilitation of problem-based learning during PLL sessions was associated with a significant effect after controlling for the session duration and training characteristics. These and other findings provide optimal pedagogical strategies for practitioners and reveal a need for additional experimental studies on student leader training and PLL in online environments.

REFERENCES

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

Ahmed, I. A. M., Elseed, M. A. M., & El-Sheikh, M. A. (2007). Expert and non-expert tutors: Role in problem-based learning. *Medical Science Educator, 17*(1), 9–13.

Ahn, S., Ames, A. J., & Myers, N. D. (2012). A review of meta-analyses in education: Methodological strengths and weakness. *Review of Educational Research, 82*(4), 436–476. doi:10.3102/0034654312458162

*Akinyele, A. F. (2010). Peer-Led Team Learning and improved performance in an allied health chemistry course. *The Chemical Educator, 15*, 353–360. doi:10.1007/s00897102300a

Arendale, D. R. (2014). Annotated bibliography of postsecondary peer cooperative learning programs available online. *Journal of Peer Learning, 7*, iv–vi.

Arendale, D. R. (2017). *Annotated bibliography of postsecondary peer cooperative learning programs*. Unpublished manuscript. Department of Postsecondary

Teaching and Learning, University of Minnesota. Minneapolis, MN. Retrieved from http://z.umn.edu/peerbib

Attarzadeh, F., Gurkan, D., Moges, M. A., Ramos, M. A., Gallardo, V. J., & Mehrubeoglu, M. (2011). *Challenges of implementing a peer mentoring program to support STEM learning.* Paper presented at the 118th ASEE Annual Conference and Exposition, Vancouver, BC; Canada.

*Báez-Galib, R., Colón-Cruz, H., Resto, W., & Rubin, M. R. (2005). Chem-2-Chem: A one-to-one supportive learning environment for chemistry. *Journal of Chemical Education, 82*(12), 1859. doi:10.1021/ed082p1859

*Batchelder, A. J., Rodrigues, C. M., Lin, L. Y., Hickey, P. M., Johnson, C., & Elias, J. E. (2010). The role of students as teachers: four years' experience of a large-scale, peer-led programme. *Medical Teacher, 32*(7), 547–551. doi:10.3109/0142159X.2010.490861

Bernard, R. M., Abrami, P. C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., … Huang, B. (2004). How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. *Review of Educational Research, 74*(3), 379–439. doi:10.3102/00346543074003379

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* Chichester, U.K.: John Wiley & Sons, 2009.

Borenstein, M., Hedges, L., & Rothstein, D. (2007). Meta-analysis: Fixed effect vs. random effects. Retrieved from http://www.metaanalysis.com/downloads /Meta-analysis_fixed_effect_vs_random_effects_sv.pdf

*Born, W. K. (2000). *The effect of workshop groups on achievement goals and performance in biology: An outcome evaluation.* (Doctor of Philosophy), Northwestern University, US.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences.* New York: Academic Press.

Cohen, P. A., Kulik, J. A., & Kulik, C. L. C. (1982). Educational Outcomes of Tutoring: A Meta-analysis of Findings. *American Educational Research Journal, 19*(2), 237–248. doi:10.3102/00028312019002237

Cracolice, M. S., & Busby, B. D. (2015). Preparation for College General Chemistry: More than just a matter of content knowledge acquisition. *Journal of Chemical Education, 92*(11), 1790–1797. doi:10.1021/acs.jchemed.5b00146

Cuseo, J. (2010). Peer leadership: Definition, description, and classification. *E-Source for College Transitions, 7*(5), 3–5.

*Dancer, D., Morrison, K., & Tarr, G. (2015). Measuring the effects of peer learning on students' academic achievement in first-year business statistics. *Studies in Higher Education, 40*(10), 1808–1828. doi:10.1080/03075079.2014.916671

Dawson, P., van der Meer, J., Skalicky, J., & Cowley, K. (2014). On the effectiveness of Supplemental Instruction: A systematic review of Supplemental Instruction and Peer-Assisted Study Sessions literature between 2001 and 2010. *Review of Educational Research, 84*(4), 609–639. doi:10.3102/0034654314540007

Dreyfuss, A. E. (2012). *A history of Peer-Led Team Learning 1990–2012.* Paper presented at the Inaugural Conference of the Peer-Led Team Learning International Society Brooklyn, NY.

Eberlein, T., Kampmeier, J., Minderhout, V., Moog, R. S., Platt, T., Varma-Nelson, P., & White, H. B. (2008). Pedagogies of engagement in science: A comparison of PBL, POGIL, and PLTL. *Biochemistry and Molecular Biology Education, 36*(4), 262–273. doi:10.1002/bmb.20204

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*(7109), 629–634. doi:10.1136/bmj.315.7109.629

Ender, S. C., & Kay, K. (2001). Peer leadership programs: A rationale and review of the literature In S. L. Hamid (Ed.), *Peer Leadership: A Primer on Program Essentials* (pp. 1–11). Columbia, SC: National Resource Center for the First-Year Experience & Students in Transition, University of South Carolina.

Evans, M. J., & Moore, J. S. (2013). Peer tutoring with the aid of the Internet. *British Journal of Educational Technology, 44*(1), 144–155. doi:10.1111/j.1467-8535.2011.01280.x

*Fayowski, V., & MacMillan, P. D. (2008). An evaluation of the Supplemental Instruction programme in a first year calculus course. *International Journal of Mathematical Education in Science and Technology, 39*(7), 843–855. doi:10.1080/00207390802054433

Ginsburg-Block, M. D., Rohrbeck, C. A., & Fantuzzo, J. W. (2006). A meta-analytic review of social, self-concept, and behavioral outcomes of peer-assisted learning. *Journal of Educational Psychology, 98*(4), 732–749. doi:10.1037/0022-0663.98.4.732

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*(10), 3–8. doi:10.3102/0013189X005010003

Gosser, D. K., Kampmeier, J. A., & Varma-Nelson, P. (2010). Peer-Led Team Learning: 2008 James Flack Norris Award Address. *Journal of Chemical Education, 87*(4), 374–380. doi:10.1021/ed800132w

*Guarcello, M. A., Levine, R. A., Beemer, J., Frazee, J. P., Laumakis, M. A., & Schellenberg, S. A. (2017). Balancing student success: Assessing Supplemental Instruction through coarsened exact matching. *Technology, Knowledge, and Learning, 22*(3), 335–352.

Headrick, T. C. (2010). *Statistical simulation: Power method polynomials and other transformations*. Boca Raton, FL: CRC Press.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods, 3*(4), 486–504.

*Hensen, K. A., & Shelley, M. C. (2003). The impact of Supplemental Instruction: Results from a large, public, midwestern university. *Journal of College Student Development, 44*(2), 250–259. doi:10.1353/csd.2003.0015

*Hizer, S. E., Schultz, P. W., & Bray, R. (2017). Supplemental Instruction online: As effective as the traditional face-to-face model? *Journal of Science Education and Technology, 26*(1), 100–115.

*Hockings, S. C., DeAngelis, K. J., & Frey, R. F. (2008). Peer-Led Team Learning in General Chemistry: Implementation and evaluation. *Journal of Chemical Education, 85*(7), 990–996. doi:10.1021/ed085p990

*Hooker, D. D. T. (2010). *A study of the effects of the implementation of small peer led collaborative group learning on students in developmental mathematics courses at a Tribal Community College.* (Doctoral dissertation), Montana State University, Bozeman, Montana.

*Hryciw, D. H., Tangalakis, K., Supple, B., & Best, G. (2013). Evaluation of a peer mentoring program for a mature cohort of first-year undergraduate paramedic students. *Advances in physiology education, 37*(1), 80–84. doi:10.1152/advan.00129.2012

Hsiao, Y. P., Brouns, F., van Bruggen, J., & Sloep, P. B. (2012). Effects of peer-tutor competences on learner cognitive load and learning performance during knowledge sharing. In D. G. Sampson, J. M. Spector, D. Ifenthaler, & P. Isaias (Eds.), *Proceedings of IADIS International Conference Cognition and Exploratory Learning in Digital Age (CELDA)* (pp. 265–268). Madrid, Spain: IADIS.

*Hughes, K. S. (2011). Peer-assisted learning strategies in human anatomy & physiology. *American Biology Teacher, 73*(3), 144–147. doi:10.1525/abt.2011.73.3.5

Hung, W., Jonassen, D. H., & Liu, R. (2008). Problem-based learning. In J. M. Spector, M. D. Merrill, J. V. Merrienboer, & M. P. Driscoll (Eds.), *Handbook of Research on Educational Communications and Technology* (3rd ed., pp. 485–506). New York, NY: Routledge.

Jensen, J. L., McDaniel, M. A., Woodard, S. M., & Kummer, T. A. (2014). Teaching to the test…or testing to teach: Exams requiring higher order thinking skills encourage greater conceptual understanding. *Educational Psychology Review, 26*(2), 307–329. doi:10.1007/s10648-013-9248-9

*Kalil, A., Jones, C., & Nast, P. (2016). Application of Supplemental Instruction in an undergraduate anatomy and physiology course for allied health students. *Supplemental Instruction Journal, 2*(1), 33–52.

*Knobe, M., Munker, R., Sellei, R. M., Holschen, M., Mooij, S. C., Schmidt-Rohlfing, B., . . . Pape, H. C. (2010). Peer teaching: a randomised controlled trial using student-teachers to teach musculoskeletal ultrasound. *Medical Education, 44*(2), 148–155. doi:10.1111/j.1365-2923.2009.03557.x.

Kokkelenberg, E. C., Dillon, M., & Christy, S. M. (2008). The effects of class size on student grades at a public university. *Economics of Education Review, 27*(2), 221–233. doi:10.1016/j.econedurev.2006.09.011

Kulik, C. L. C., & Kulik, J. A. (1982). Effects of ability grouping on secondary school students: A meta-analysis of evaluation findings. *American Educational Research Journal, 19*(3), 415–428. doi:10.3102/00028312019003415

Lee, Y., & Choi, J. (2010). A review of online course dropout research: implications for practice and future research. *Educational Technology Research and Development, 59*(5), 593–618. doi:10.1007/s11423-010-9177-y

*Lewis, S. E. (2011). Retention and reform: An evaluation of Peer-Led Team Learning. *Journal of Chemical Education, 88*(6), 703–707. doi:10.1021/ed100689m

*Lewis, S. E., & Lewis, J. E. (2008). Seeking effectiveness and equity in a large college chemistry course: an HLM investigation of Peer-Led Guided Inquiry. *Journal of Research in Science Teaching, 45*(7), 794–811. doi:10.1002/tea.20254

Lockspeiser, T. M., O'Sullivan, P., Teherani, A., & Muller, J. (2008). Understanding the experience of being taught by peers: The value of social and cognitive congruence. *Advances in health sciences education: Theory and practice, 13*(3), 361–372. doi:10.1007/s10459-006-9049-8

*Loui, M. C., Robbins, B. A., Johnson, E. C., & Venkatesan, N. (2013). Assessment of Peer-Led Team Learning in an engineering course for freshmen. *International Journal of Engineering Education, 29*(6), 1440–1455.

Martin, D. C., & Arendale, D. (1992). Supplemental Instruction: Improving first-year student success in high risk courses. *The Freshman Year Experience: Monograph Series Number 7*, 19–26.

Miller, D., & Schraeder, M. (2015). Research on group learning and cognitive science: A study of motivation, knowledge, and self-regulation in a larger lecture college Algebra class. *The Mathematics Educator, 24*(2), 27–55.

Mitchell, Y. D., Ippolito, J., & Lewis, S. E. (2012). Evaluating Peer-Led Team Learning across the two semester General Chemistry sequence. *Chemistry Education Research and Practice, 13*(3), 378–383. doi:10.1039/c2rp20028g

*Moust, J. C., & Schmidt, H. G. (1994). Effects of staff and student tutors on student achievement. *Higher Education, 28*, 471–482.

Neville, A. J. (1999). The problem-based learning tutor: Teacher? Facilitator? Evaluator? *Medical Teacher, 21*(4), 393–401.

Ogden, P., Thompson, D., Russell, A., & Simons, C. (2003). Supplemental Instruction: Short- and long-term impact. *Journal of Developmental Education, 26*(3), 2–8.

Park, S. E., Susarla, S. M., Cox, C. K., Silva, J. D., & Howell, T. H. (2007). Do tutor expertise and experience influence student performance in a problem-based curriculum? *Journal of Dental Education, 71*(6), 819–824.

Pastor, D. A., & Lazowski, R. A. (2018). On the multilevel nature of meta-analysis: A tutorial, comparison of software programs, and discussion of analytic choices. *Multivariate Behavioral Research, 53*(1), 74–89. doi:10.1080/00273171.2017.1365684

Peplow, M. (2014). Social sciences suffer from severe publication bias. *Nature*. Retrieved from https://www.nature.com/news/social-sciences-suffer-from-severe-publication-bias-1.15787.

Piaget, J. (1932). *The moral judgment of the child*. Glencoe, Ill: Free Press.

Polanin, J. R., Hennessy, E. A., & Tanner-Smith, E. E. (2016). A Review of meta-analysis packages in R. *Journal of Educational and Behavioral Statistics, 42*(2), 206–242. doi:10.3102/1076998616674315

*Preszler, R. W. (2009). Replacing lecture with peer-led workshops improves student learning. *CBE Life Science Education, 8*(3), 182–192. doi:10.1187/cbe.09-01-0002

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education, 93*(3), 223–231. doi:10.1002/j.2168-9830.2004.tb00809.x

*Quitadamo, I. J., Brahler, C. J., & Crouch, G. J. (2009). Peer-Led Team Learning: A prospective method for increasing critical thinking in undergraduate science courses. *Science Educator, 18*(1), 29–39.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd. ed. Vol. 1). Thousand Oaks, CA: SAGE Publications.

Reardon, J., Payan, J., Miller, C., & Alexander, J. (2008). Optimal class length in marketing undergraduate classes: An examination of preference, instructor evaluations, and student performance. *Journal of Marketing Education, 30*(1), 12–20. doi:10.1177/0273475307312193

*Reisel, J. R., Jablonski, M., Munson, E. V., & Hosseini, H. (2012). *Analysis of the impact of formal peer-led study groups on first-year student math performance.* Paper presented at the 119th ASEE Annual Conference and Exposition, San Antonio, TX.

*Reisel, J. R., Jablonski, M. R., Munson, E., & Hosseini, H. (2014). Peer-Led Team Learning in mathematics courses for freshmen engineering and computer science students. *Journal of STEM Education: Innovations and Research, 15*(2), 7–15.

Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology, 95*(2), 240–257. doi:10.1037/0022-0663.95.2.240

Rosnow, R. L., & Rosenthal, R. (1996). Computing contrasts, effect sizes, and counternulls on other people's published data: General procedures for research consumers. *Psychological Methods, 1*(4), 331–340.

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005). *Publication bias in meta-analysis: Prevention, assessment and adjustments*. West Sussex, UK: Wiley.

Salas, E., Nichols, D. R., & Driskell, J. E. (2016). Testing three team training strategies in intact teams. *Small Group Research, 38*(4), 471–488. doi:10.1177/1046496407304332

Schmidt, H. G., & Moust, J. H. C. (1995). What makes a tutor effective? A structural-equations modeling approach to learning in problem-based curricula. *Academic Medicine, 70*(8), 708–714.

*Schneiders, B. B. (2014). *The effectiveness of calculus workgroup on student performance in calculus: A mixed-methods approach.* (Ph.D. dissertation), University of Colorado at Boulder, Boulder, CO.

Secomb, J. (2008). A systematic review of peer teaching and learning in clinical education. *Journal of Clinical Nursing, 17*(6), 703–716.

Sheu, C. F., & Suzuki, S. (2001). Meta-analysis using linear mixed models. *Behavior Research Methods, Instruments, & Computers, 33*(2), 102–107. doi:10.3758/BF03195354

*Shields, S. P., Hogrebe, M. C., Spees, W. M., Handlin, L. B., Noelken, G. P., Riley, J. M., & Frey, R. F. (2012). A Transition Program for Underprepared Students in General Chemistry: Diagnosis, Implementation, and Evaluation. *Journal of Chemical Education, 89*(8), 995–1000. doi:10.1021/ed100410j

Skalicky, J., & Caney, A. (2010). PASS student leader and mentor roles: A tertiary leadership pathway. *Journal of Peer Learning, 3*(1), 24–37.

Smith, J., Wilson, S. B., Banks, J., Zhu, L., & Varma-Nelson, P. (2014). Replicating peer-led team learning in cyberspace: Research, opportunities, and challenges. *Journal of Research in Science Teaching, 51*(6), 714–740.

*Snyder, J. J., Carter, B. E., & Wiles, J. R. (2015). Implementation of the Peer-Led Team Learning instructional model as a stopgap measure improves student achievement for students opting out of laboratory. *CBE-Life Education, 14*(1). doi:10.1187/cbe.13-08-0168

*Steele, D. J., Medder, J. D., & Turner, P. (2000). A comparison of learning outcomes and attitudes in student- versus faculty-led problem-based learning: An experimental study. *Medical education, 34*(1), 23–29. doi:10.1046/j.1365-2923.2000.00460.x

Strobel, J., & van Barneveld, A. (2009). When is PBL more effective? A meta-synthesis of meta-analyses comparing PBL to conventional classrooms. *Interdisciplinary Journal of Problem-Based Learning, 3*(1). doi:10.7771/1541-5015.1046

Tajfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worshel (Eds.), *The Social Psychology of Intergroup Relations*. Monterey, CA: Brooks/Cole Pub. Co.

Tarrant, M. (2002). Adolescent peer groups and social identity. *Social Development, 11*(1), 110–123.

Teo, T. (2011). Basic concepts of Hierarchical Linear Modelling (HLM) for educational technology research. *British Journal of Educational Technology, 42*(3), E51-E53. doi:10.1111/j.1467-8535.2011.01171.x

*Tien, L. T., Roth, V., & Kampmeier, J. A. (2002). Implementation of a Peer-Led Team Learning Instructional Approach in an Undergraduate Organic Chemistry Course. *Journal of Research in Science Teaching, 39*(7), 606–632. doi:10.1002/tea.10038

Topping, K. J. (1996). The effectiveness of peer tutoring in further and higher education: A typology and review of the literature. *Higher Education, 32*(3), 321–345.

Topping, K. J. (2005). Trends in peer learning. *Educational Psychology, 25*(6), 631–645. doi:10.1080/01443410500345172

Wilson, K., & Korn, J. H. (2007). Attention During Lectures: Beyond Ten Minutes. *Teaching of Psychology, 34*(2), 85–89. doi:10.1080/00986280701291291

Zumbach, J., & Spraul, P. (2007). The role of expert and novice tutors in computer mediated and face-to-face problem-based learning. *Research and Practice in Technology Enhanced Learning*, 2(2), 161–187.