The Effect of Collaborative Learning on Enhancing Student Achievement

A Meta-Analysis

Ravinder Kumar

A Thesis

in

The Department

of

Education

Presented in Partial Fulfilment of the Requirements

for the Degree of Master of Arts (Educational Technology) at

Concordia University

Montreal, Quebec, Canada

April 2017

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Ravinder Kumar

Entitled: The Effect of Collaborative Learning on Enhancing Student Achievement
A Meta- Analysis

and submitted in partial fulfillment of the requirements for the degree of

Master of Arts (Educational Technology)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

________________________________ Chair
Chair's name/ Dr. Saul Carliner

________________________________ Examiner
Examiner's name / Dr. Robert M. Bernard

________________________________ Examiner
Examiner's name / Dr. Claude Martel

________________________________ Supervisor
Supervisor's name / Dr. Richard F. Schmid

Approved by ____
Chair of Department or Graduate Program Director

________________ ____
Date Dean of Faculty

Abstract

The Effect of Collaborative Learning on Enhancing Student Achievement: A Meta-Analysis

Ravinder Kumar, M.A.
Concordia University, 2017

The purpose of this meta-analysis was to investigate the effect of collaborative learning on student achievement. The sample of the study consisted of 20 representative studies involving 2434 participants selected from an extensive literature search based on the use of collaborative activities in a formal education setting cutting across multiple grade levels and subject domains. Analysis of representative studies (k = 28) yielded a moderately weighted average effect size of 0.26. A mixed effects model was used for the analysis of the moderators of effect size. The research design was not significant across true and quasi- designs. Two groups – high collaboration (Experimental condition) and no or less collaboration (control condition) – were compared to measure the effect of collaborative learning on the dependent variable of student achievement. The analyses of moderator variables were not significant or suffered from a lack of statistical power (i.e., grade levels). Implications for the use of the collaborative learning are discussed, and recommendations for future researchers are suggested along with the limitation and conclusions.

Keywords: collaborative learning, meta-analysis, student achievement

Acknowledgements

As a graduate student, I experienced many memorable learning moments during the entire program. The thesis option provided me with a real task to synthesize my experiences as a capstone of all the learning. The journey was rigorous and challenging. However, the continuous support and guidance provided by some of the wonderful individuals in the department helped me complete this task successfully. I want to extend my immeasurable appreciation and deepest gratitude for the support offered by the following persons :

Foremost, **Dr. Richard Schmid**, my supervisor, chair, and professor, for his full support from the day one. He showed confidence in my abilities and helped me through his continuous advices, suggestions, feedback, comments, and support during the wrting process of the entire thesis. In addition, I am indebted to him for his official support to complete the departmental formalities for my thesis defense and submission. His commitment, patience, and sincerity were manifested during every meeting and/or communication with him which boosted my morale to stay in tune at every stage of this work. In addition, I thank him again for serving as my referee for various research grant projects and mentorship.

Dr. Robert Bernard, my professor and research committee member, and a specialist in meta-analysis, for his guidance during every stage of this project for all his advices on technical matters. I sincerely thank him for allowing to use the dataset for my study sample. In addition, he always showed his willingness to help me out whenever it was needed. His expertise has contributed immensely to keep the quality of the current work intact.

Dr. Claude Martel, my second committee member and professor, for his emotional support and time-to-time help during the program. His expertise in applied research and educational consultancy helped me devise better strategies for a successful planning of this project. I always found him a great person with his empathetic touch to uplift my morale and motivation.

Evgueni Borokhovski, another faculty member from Concordia University, for his regular support and motivation during the entire journey of this work. He has been exceptional to support me through all the phases of this work including coding, data analysis and reporting of the results. He always rendered his services to polish this work. He is such an appealing person that I befriended with him during my interactions with him.

Saul Carliner, another professor from the faculty of education technology, for offering his service to chair my defense committee. He has been a great influence on my learning, as his quality books on instructional design and resourceful peer-reviewed publications throughout his academic career.

Lastly, I would extend my thanks to all the other teaching community members for their support and teaching. In addition, I express my gratitude to all the administrative staff of the Department of Educational technology for their continuous support and care during the entire program. In particular, I would like to thank Mr. Sean Gordon, my program coordinator, for his amazing interactions and timely information during the my entire journay in this program.

Dedicated to

With the incessant grace and blessings of the **Almighty** I dedicate this work with all my love and gratitude to Mrs. **Sunita**, a departed soul and integral part of my family, for her continuous motivation and support throughout my academic career,

to my **parents and brother** *who always supported me,*

to **Anushaka and Aryan**, *my sweet buds for their affection,*

and most importantly

to **Anju, my** *beloved and everything, for her love, trust, and support throughout my life.*

Table of Contents

List of Figures

List of Tables

Chapter One : Introduction

Background of the Project

Teaching and learning in a modern classroom is no longer an act of transferring knowledge. The act of teaching has become a multidiciplinary enterprize to develop critical thinking, interaction, and collaboration among learners (Nelson ,1994). Given these multidisciplinary changes in curriculum and its relative learning objectives, the need to collaborate in order to create learning environments has gained momnetum in this decade or so. Instead of teacher-centred approaches, the focus has shifted to learner-centred and learning-centred strategies. In the current educational landscape, learners are no more the empty vessels to be filled in, rather they need to be the co-creators of knowledge; they should be willing to take ownership of their learning and contribute to the development of knowledge.

Collaborative learning, according to Dillenbourg (1999), is "a situation in which two or more people learn or attempt to learn something together," and more specifically as joint problem solving (p. 1). The use of collaborative approaches in pedagogies have gained currency in the last few decades. According to Wilczenski, Bontrager, Ventrone, & Correia, (2001), collaborative learning entails "students working together without immediate teacher supervision in groups small enough that all students can participate collectively in a task" (p. 270). Further, Roschelle and Teasley evaluate collaboration as the "mutual engagement of participants in a coordinated effort to solve a problem together" (as cited in Dillenbourg et al., 1996, p. 2). Thus collaboration can be termed as "coordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem" (Roschelle & Teasley, 1995, p. 70). A collaboratively rich environment has been strongly recommended to improve students' cognitive performance, social relationships, and metacognition (Paris & Winograd, 1990). Keeler & Anson (1995) also asserted that cooperative learning significantly improved learning performance and reduced attrition among students.

According to Lai (2011), research on collaboration has stemmed from three distinct strands: research comparing group performance to individual performance, identifying the conditions which favor or challenge collaboration as more or less effective, and research exploring the characteristics of interactions that evaluate the impact of collaboration on learning and achievement including the moderators such as use of new technologies that facilitate numerous interactions. For example, research have been conducted in this domain including designing interactive learning environments (Borokhovski, Tamim, Bernard, Abrami, & Sokolovskaya, 2012), technology integration in postsecondary education (Schmid, Bernard, Borokhovski, Tamim, Abrami, Surkes, Wade & Woods, 2014), and collaboration and its impact on student learning (Uribe, Klein, & Sullivan, 2003; Beldarrain, 2006; Williams, 2009; Tomcho, & Foels, 2012).

Statement of the Problem

In light of these issues, the purpose of this research is to undertake a meta-analysis to investigate the effect of collaborative learning on student achievement and the role of other associated moderators such as technology, subject domains, grade levels and duration to facilitate and/or support collaborative learning impacting student achievement.

Research Questions

The current study attempts to investigate the effect of collaborative learning on student achievement while comparing the two conditions – high collaboration (experimental condition) and less or no collaboration (control condition). In addition, the study is interested in exploring the moderator variables which affect student collaboration and achievement outcomes. The study also examines if the degree of collaboration has a varying effect at different grade levels, subject domains, and duration which in turn might influence student achievement outcomes. The following three research questions guide this meta-analysis:

1) Does collaborative learning have any statistically significant effect on student achievement outcomes when compared learning without (or under lesser degree of) collaboration?
2) Do different types of technology have varying effects on student achievement when used to enhance/support /promote collaboration?
3) Do grade levels and subject domains have any moderating effects on student`s achievement?

Significance of the Study

The study is significant in that interaction and collaboration help develop and improve student performance and their achievement outcomes. Accroding to Van Boxtel, Van der Linden, and Kanselaar (2000), collaborative learning activities help learners to find explanations of their understanding that assists them elaborate and reorganize their knowledge. Various studies have found that collaboration among students has considerable impact on their achievement outcomes (Borokhovski, Bernard, Tamim, Schmid, & Sokolovskaya, 2016; Fjermestad, 2004; Schmid et al., 2014). Some meta-analyses which focussed on the use of technology in distance and/or online learning, have found a positive impact of technology on student collaboration and on learning outcomes (Abrami, Bernard, Bures, Borokhovski, & Tamim, 2011; Beldarrain, 2006). Another group of studies explored design-based cooperative /collaborative learning, and reported a significant impact of collaborative activities on student learning outcomes (Lou, Abrami, & Appolonia, 2001; Chou & Min, 2009; Lee, 2007; Puzio & Colby, 2013; Wright, Kandel-Cisco, Hodges, Metoyer, Boriack, Franco-Fuenmayor & Waxman, 2013). Other researchers who investigated the pedagogical and theoretical aspects of technology integration (Arts, Gijselaers & Segers, 2002; Jehng, 1997; Kanuka & Anderson, 1999; Pedró, 2005; Lou, Bernard, & Abrami, 2006; Mantri, Dutt, Gupta & Chitkara, 2008), found mixed effects of technology intergration on student learning outcomes.

Furthermore, Bernard, Abrami, Borokhovski, Wade, Tamim, Surkes, and Bethel (2009) were interested in the relative effectiveness of the designed and contextual interaction treatments in distance education . They found a strong association between the strength of interaction treatments and achievement for asynchronous DE courses compared to courses containing mediated synchronous or face-to-face inter (p. 1243). Borokhovski et al. (2012) found that higher levels of collaboration and cooperation among students could be achived by employing technology to enable , support and facilitate discussions (p. 321). Schmid et al. (2014) explored the impact of technology-enhanced instruction in postsecondary education and reported that learning is best supported when students are engaged in active and meaningful exercises via technological tools that provide cognitive support. (p. 285). According to Baepler, Walker, and Driessen (2014), in an active learning classroom, student faculty contact could be reduced by two-thirds and students achieved learning outcomes that were at least as good, and in one comparison significantly better than, those in a traditional classroom (p. 227).

Borokhovski et al. (2016) tried to map out the added value of planned collaborative activities versus unplanned grouping of students in the context of postsecondary education. They reported that the designed-based treatments outperformed contextual treatments on measures of achievement and they strongly suggested the value of planning and instructional designing in technology integration in post-secondary education (p. 15). These studies have contributed to the literature on collaborative learning. Though Kyndt, Raes, Lismont, Timmers, Cascallar, and Dochy (2013) have explored the effects of face-to-face cooperative learning in this regard, no study has attempted to investigate the effect (s) of collaborative learning on student achievement in both online and and face-to -face modes collectively across varied levels of formal education and subject domains. Therefore, it would be interesting to investigate how the effect of collaborative learning on student achievement varys, especially, amidst the different grade levels and subject domains in both online and face-to-face modes of instruction.The current meta-analysis will fill this gap by exploring the effect of collaborative learning on student achievement in the context of formal education settings cutting across all grades and subject domains. Indeed, the findings of this study will contribute to the pool of existing body of knowledge on collaboration. Furthermore, while the outcomes of this meta-analysis may have limited generalizability, this study would serve as a springboard for the prospective studies in the domain.

Student interaction and collaboration. Interactions among student-student, student-teacher and student-content are vital to engage learners for collaboration (Schmid et al., 2014). Since interaction is a primary condition to collaborate, three types of interactions, as mentioned above, have been effective, particularly in various theoretical frameworks on distance and online education (Anderson, 2003; Beldarrain, 2006; Moore, 1989). Interactions among students trigger the learning process. In this connection, the understanding of collaborative and cooperative learning is a must (Decuyper, Dochy, & Van den Bossche, 2010). As collaboration is an umbrella term, it encompasses interactions between and/or among student-student, teacher-students and student-content (Anderson, 2003; Beldarrain, 2006; Moore, 1989). Roschelle and Teasley define collaboration as a "mutual engagement of participants in a coordinated effort to solve a problem together" (Dillenbourg et al., 1996, p. 2). The central assumption of collaboration lies in the interaction between individuals or groups. It may take many forms – cooperative, collective, peers, group and/or team learning. According to some authors (Bruffee 1995; Eastmond, 1995; Webb & Palincsar, 1996), small-group learning strategies reside along a continuum from loosely structured (i.e. collaborative) to highly structured (i.e. cooperative).

Presumably, the aim of collaborative learning is the co-construction of knowledge through interaction. So collaborative learning refers to a variety of educational approaches that encourage students to work together, including cooperative learning; problem-based instruction; guided design; writing groups; peer teaching; workshops; discussion groups; and learning communities (Smith & McGregor, 1992). Vygotsky (1962) stressed the need for collaborative learning either between teacher-students or among students to assist students in advancing through their zone of proximal development (the gap between what can a student achieve alone and in collaboration with the group). Collaboration can be manifested in the face-to-face form (in-class group project) and online technology-mediated formats (synchronous and asynchronous media–wikis, blogs, forums and any other forms of online communication). However, Kirschner and Erkens (2013) relate:

> It has become clear that simply placing learners in a group and assigning them a task does not guarantee that they will work together … coordinate their activities … engage in effective collaborative learning processes… lead to positive learning outcomes (p. 1).

Regardless of the environments in which collaborative learning is implemented, there is an agreement that simply grouping students does not promote higher achievement or more positive relationships among students (Johnson & Johnson, 1996). Collaborative learning means specific forms of interaction that have a purposive intent and a general set of structural elements to facilitate it (Harasim, Hiltz, Teles, & Turoff, 1995).

Concerning collaborative learning, the allied concept of cooperation is also important. In cooperative learning individuals work together to achieve shared goals; they collaborate to maximize their own and each other's learning (Johnson & Johnson, 1989). According to Johnson and Johnson (2009), five elements must be present in cooperative/collaborative learning: 1) positive interdependence, 2) individual accountability, 3) face-to-face promotive interaction, 4) social and interpersonal skills, and 5) group processing. However, in the current project, the focus is on the purposive cooperative/collaborative activities which enhance students' achievement. Collaboration is understood as interaction between and/or among student-student (Bernard et al., 2009 ; Borokhovski et al., 2012), teacher-students and student-content (Anderson, 2003; Beldarrain, 2006; Moore, 1989) which enhance student learning outcomes, excluding collaboration meant for non-academic intentions.

Technology and pedagogy. Technology has been revolutionizing the learning sciences significantly since the 1990s when the internet access with personal computers became widely available for educational purposes (Kozma, 1994), and computers provided more capacity for information processing and other advanced functions. The historical debate between Clark and

Kozma over the use of technology as a medium or more than a medium opened the floodgate for experimentation on and initiation of multiple instructional strategies in and out of the classroom, promising the significance of technology to support student learning and attitudes (Dede, 1996, 2004; Kozma, 1991, 1994; Mayer, 2008).

Since then, technology has become a buzz word in the parlance of the current educational practices. In the last fifteen years, the influx of technological innovations has bombarded the field of education immensely. Technology is not merely a set of hardware or software paraphernalia, but also is, as Ross, Morrison, and Lowther (2010) define, "a broad variety of modalities, tools, and strategies for learning, [whose] effectiveness, therefore, depends on how well [they] help teachers and students achieve the desired instructional goals" (p. 19). Technology-mediated instruction has complimented education in enhancing teaching and learning both inside and outside the classroom. Uses of computer-assisted technology and cloud-based authoring technology have facilitated multiple forms of pedagogical and instructional strategies such as online learning/distance education (Bernard, Rojo de Rubalcava, & St-Pierre, 2000 ; Bernard, Abrami, Lou, Borokhovski, Wade, Wozney, 2004 ; Bernard et al., 2009), blended learning (Henrie, Halverson, & Graham, 2015 ; Schmid et al., 2014), and MOOCs (Margaryan, Bianco, & Littlejohn, 2015). In addition, technology has widened the access and scope of learning and offered options to collaborate worldwide.

The use and impact of technology on student learning, however, are two different things. This is so because the impact is directly decided by the manner and the purpose of the use of technology. Technology as a moderator variable in collaborative learning is interesting for many reasons. Given recent technologies such as cognitive tools, communication methods, search and retrieval strategies, and other presentational tools, it would be interesting to investigate the use of technology in collaborative activities that impacts student achievement. A meta-analysis by Susman (1998) found that participants in collaborative, computer-based conditions showed a greater increase in elaboration, higher-order thinking, metacognitive processes, and divergent thinking than participants in individual computer-based instruction.

Given this perspective, it would be interesting to know how the use of technology in its varied manifestations facilitate collaboration which might affect student achievement. In other words, this study explores whether and how technology-supported collaborative activities help learners interact and collaborate leading to enhanced achievement outcomes.

Learning and achievement. Dillenbourg (1999) defines "learning" as a "biological and/or cultural process" that takes place over several years; a "joint problem solving in which learning is assumed to occur as a by-product of interactions" (p.4). Learning can be more effective when students discuss their ideas, experiences, and perceptions with peers (Jonassen & Kwon, 2001; Kanuka & Anderson, 1998). In a face-to-face classroom setting, collaborative learning can affect learning outcomes, social skill development, and self-esteem positively (Johnson & Johnson, 1996; Slavin, 1990). Similarly, the use of collaborative learning strategies in computer-based instruction (CBI) is also supported by research (Cavalier & Klein, 1998; Dalton, Hannafin, & Hooper, 1989 ; Hooper 1992; Hooper & Hannafin, 1991; Klein & Doran, 1999; Kulik & Kulik, 1991; Sherman & Klein, 1995). In addition, theorists also have propounded that computer-mediated collaboration can positively enhance learning, problem-solving, and other higher-order thinking (Adelskold, Alklett, Axelsson, & Blomgren, 1999 ; Johnson & Johnson, 1996 ; Jonassen, Prevish, Christy, & Stavulaki, 1999). In the context of the current study, the analysis is restricted to the collaborative activities which affect student achievement and how different moderator variables, including technology, grade levels, subject domains, and duration influence student collaboration.

Achievement, as a dependent variable, informs the academic performance by means of standardized and/or validated measures (Schmid et al. 2014). However, in the current study, the term achievement is confined to students' accomplishment, participation, success, and attainment in the parlance of collaborative activities measured by exams, tests, and grades.

Chapter Two : Literature Review

Historically, according to Dillenbourg, Baker, Blaye, and O'Malley (1996), a great amount of research on collaborative and cooperative learning stemmed from the works of Piaget and Vygotsky. Piaget's system of developmental stages describing children's cognitive progress contributed to the development of socio-constructivists paradigm. Similarly, his concept of cognitive conflict provided a cognitive framework to understand learner's experiences. In addition, social interactions mediate cognitive conflict that allow learners to interact with peers at more advanced developmental levels.

On the contrary, regarding cognitive conflict Vygotsky stressed the value of social interaction itself for causing individual cognitive change, as opposed to being merely stimulated by it (Dillenbourg et al., 1996). Internalized social interaction causes conceptual changes in participants that help them negotiate meaning. A similar concept, the zone of proximal development, according to Vygotsky, is the distance between what a student can accomplish individually and what he/she can accomplish with the help of a more capable "other." While Piaget suggests pairing children based on different developmental stages to facilitate cognitive conflict, Vygotsky, on the other hand, recommends pairing children with adults. Unlike Piaget and Vygotsky who maintain that cognitive conflict causes conceptual change, socio-culturalists privilege collaborative learning that takes place within the zone of proximal development (Dillenbourg et al., 1996).

According to Kreijns, Kirschner, and Jochems (2003), a new strand of research regarding collaborative learning emerged in the late 1990s that focused on new technologies for mediating, observing, and recording interactions during. On the whole, four strands came into existence out of the seminal works of Piaget, Vygotsky and their shared concept of cognition and research built on them – the "effect" paradigm, the "conditions" paradigm, the "interactions" paradigm, and "computer-supported" paradigm respectively (Dillenbourg et al., 1996, pp. 8 -17). In the next paragraphs, the author discusses these paradigms.

The “effect” paradigm investigates outcomes of collaboration rather than the collaborative process itself, and compares group performance with individual performance. This paradigm maintains that a collaborative classroom culture can have powerful effects on student learning and performance. Webb (1993) found that the students who worked in groups on computational math problems scored significantly higher than equivalent-ability students who worked individually.

The “conditions” paradigm tries to determine the conditions that moderate the effectiveness of collaboration on learning, for instance, individual characteristics of group members, group heterogeneity and size, and task features. Webb’s (1991) study reported significant differences in the collaborative learning experiences of boys and girls. Boys were more likely than girls to give and receive elaborated explanations, and their explanations were more likely to be accepted by group mates than girls’ explanations (Dillenbourg et al., 1996).

To overcome the complexities of former paradigms, the “interactions” paradigm emerged to identify the intersecting systems between collaboration and learning outcomes. In a way, this paradigm tried to explain the characteristics and processes of interactions which measure the effect of collaboration on learning (Dillenbourg et al., 1996). In this regard, Webb (1991) reported that the effect of collaborative learning on student achievement depends on the quality of the interactions among them.

The fourth paradigm, the contemporary one, was developed to explore whether the theoretical benefits of collaborative learning as harvested in face-to-face settings can be repeated in a computer-mediated or computer-assisted interactions given the asynchronous, text-based interactions. For example, Curtis and Lawson (2001) found that in online media there were fewer exchanges among student during collaboration given their unfamiliarity prior to online interactions. Further, the “online medium was found effective only in planning activities and coordinating work than challenging ideas” (pp. 29-30). For the purpose of the current study, the author will rely upon the collective contributions of these approaches to investigate his research questions.

Over the past three and half decades numerous meta-analyses have been conducted to investigate the effects of collaborative and small-group instruction on student learning and achievement outcomes. Twelve meta-analyses spanning from 1981 to 2016 are discussed in the next section (See Table 1). Johnson, Maruyama, Johnson, Nelson, and Skon (1981) reported that cooperation both with and without intergroup competition is more effective than the interpersonal competition and individual efforts. Similarly, Newmann and Thompson's (1987) study which was conducted in the context of secondary education found 68% of the studies yielded positive effects in favor of the cooperative condition. Qin, Johnson, and Johnson (1995) investigated cooperative versus competitive efforts and problem solving. They found that studies with non-linguistic problems (for example the study domain of mathematics or exact sciences) showed slightly more positive effects than studies with linguistic problems. Lou, Abrami, Spence, Poulson, Chambers, and d'Apollonia, (1996) who explored the differences in achievement and attitudes at all grade levels of education, concluded that "on average, students learning in small groups within classrooms achieved significantly more than students not learning in small groups" (p. 439).

On the other hand, Johnson, Johnson, and Smith (1998) focused their research on higher education settings and adults. They reported that cooperative learning results in positive effects on achievement in comparison with competitive or individualistic learning. Springer, Stanne, and Donovan's meta-analysis (1999) investigated the effects of cooperative learning on achievement, attitudes and persistence in the context of undergraduate STEM courses. They reported that students who were learning in cooperative groups showed better achievement than students who were not learning in cooperative groups. Bowen's (2000) second meta-analysis, which focused on high school and college level chemistry students, pointed out that "on average, using aspects of cooperative learning can enhance chemistry achievement for high school and college students" (p. 119). He found that cooperative learning had a significant positive effect on student attitudes toward STEM courses.

Another meta-analysis conducted by Lou, Abrami, and D'Appolonia (2001) reported that, on average, small group learning had significantly more positive effects than individual learning on student individual achievement (mean ES = + 0.15) and on group task performance (mean ES = +0.31). Similarly, Lou et al. (2006) reported a significant correlation between student-student interaction and greater achievement success ($g+ = 0.11$, $k = 30$, $p < .05$) in the context of the undergraduate distance education courses. In the same vein, Bernard et al. (2009) were interested in three types of interaction treatments (i.e. student-student, student-teacher, and student-content). They found an explicit link between interaction and academic performance in distance education that improved student learning. The student-student interaction emerged as the most important group among the three ($g+ = 0.49$, $k = 10$, $p < .05$). In addition, they found higher achievement effect size regarding the presence of technology which appeared to have facilitated or at least improved the effectiveness of interaction among students, as reflected in achievement learning outcomes.

Furthermore, Bernard et al. (2004) found a synergistic relationship between technology and pedagogy in distance education and speculated the same relationships in the context of other learning environments where the use of technology in designing environment might promote student collaboration and knowledge construction.Tamim, Bernard, Borokhovski, Abrami, and Schmid, (2011) also provided a useful insight about the use of technology in regular formal educational contexts. The authors (Tamim et al., 2011) found that technology had a positive impact on student achievement. In the same coin, Schmid and colleagues (2014) examined the effects of technology integration on students in postsecondary classroom settings. The degree of technology use was the defining characteristic for the effect size. The authors concluded that learning was best supported when students were engaged in active and meaningful exercises using technological tools that provided them cognitive support.Technological tools themselves do not make learning happen. In this connection, Abrami and colleagues (2011) rightly remarked :

[J]ust because opportunities for interactions were offered to students does not mean that students availed themselves of them, or if they did interact, that they did so effectively. The latter case is the more likely event, so the achievement effects resulting from well-implemented interaction conditions may be underestimated in our review (p. 86).

Borokhovski et al. (2012) observed that interactively designed activities are more conducive to increase student learning than do the contextual instructional settings which are not intentionally designed to create collaborative learning environments. Recently, Borokhovski and colleagues (2016) found that designed treatments outperformed contextual treatments ($g+ = 0.52$, $k = 25$ vs. $g+ = 0.11$, $k = 20$, Q-Between = 7.91, $p < .02$) on measures of achievement and emphasized the importance of planning and instructional design in the integration of technology at postsecondary levels. Similarly, Kyndt, Raes, Lismont, Timmers, Cascallar, and Dochy (2013) reported a positive effect of cooperative learning on student achievement and attitudes. In addition, they reported that the study domain, the age group, and the students' cultures also produce variations in effect size.

Having discussed these meta-analsyes, the cooperative / collaborative learning showed a positive effect on student achivevment. Though these studies have contributed to the literature on collaborative learning, however, there is a discernible gap informed by these studies. The gap is : How do the moderators such as technology, subject domain, and grade levels collectively affect the student collaboration and in turn student achivement? The current meta-analysis will fill this gap.

Table 1. Meta-analyses addressing the effects of cooperative/collaborative learning

Review Study (Year)	***k***	**Dependent variable**	**Conditions of the Independent variable**	**Mean Effect Size**
Johnson et al. (1981)	122	Achievement	Coop./Competitive	0.78
Qin et al. (1995)	46	Achievement	Coop./Competitive	0.55
Lou et al. (1996)	51	Achievement	Within class grouping / no	0.17
Johnson et al. (1998)	168	Achievement	- Coop./Competitive	0.49
Springer et al. (1999)	37	Achievement	- Coop./Competitive	0.51
Bowen (2000)	37	Achievement	Coop./Traditional	0.51
Lou et al. (2001)	122	Achievement	Small group/ Individual	0.15
Bernard et al. (2009)	74	Achievement	Interaction treatments (ITs) with other DE instructional treatments	0.38
Schmid, R. et al. (2009)	310	Achievement	High Technology/ some or no technology	0.28
Tamim, R. M. (2009).	37	Achievement	Face-to-face/Traditional	0.30
Kyndt, Eva et al. (2013)	65	Achievement	Coop./Traditional	0.54
Borokhovski, Eugene (2016)	45	Student learning outcomes	Designed treatments /contextual treatments groups	0.52

Chapter Three : Methodology

In this section, firstly, the author defines the major terms as the working definitions associated with the research problem and the research questions to be investigated in the current study. Secondly, he informs about the proceedures used for the ethical considerations meant for this meta-analysis. Thirdly, the author unpacks the methodology which includes : 1) the study design, variables, methods and instruments ; 2) literature search ; 3) the coding study features ; and 4) process of the calculation and synthesis of the effect sizes.

Terms and Definitions

Collaboration: According to Panitz (1996), is a philosophy of interaction and personal lifestyle whereas cooperation is a structure of interaction, a classroom technique, designed to facilitate the accomplishment of an end product or goal. Cooper, Prescott, Cook, Smith, and Mueck (1990) have defined, cooperative/collaborative learning as "an instructional technique that requires students to work together in small, fixed groups on a structured learning task". The current study focuses on collaborative learning as "a situation in which two or more people learn or attempt to learn something together," and more specifically as joint problem solving (Dillenbourg, 1999, p. 1).

Formal educational settings: According to Coombs and Ahmed (1974), a formal education setting is structured in a hierarchical manner that spans from primary school to university levels including general academic studies and a varied of specialized programs for full-time technical and professional training.

Pedagogical uses of technology : Based on the meta-analysis of Schmid et al. (2014), the current meta-analysis distinguishes among the following pedagogical uses (major purposes) of educational technology :

1) To promote **communication** and/or facilitate the exchange of information that includes technology which enables a higher level of interaction between individuals (i.e., two-way communications among learners and between learners and the teacher);

2) To provide **cognitive support** for learners which encompasses various technologies that enable, facilitate, and support learning by providing cognitive tools (e.g., concept maps, simulations, wikis, different forms of elaborate feedback, spreadsheets, word processing);

3) To facilitate **information search and retrieval** using technology to access to additional information (e.g., web links, search engines, electronic databases); and

4) To enable or enhance **content presentation where technology** is primarily used by teachers and/or students to present, illustrate and otherwise enrich the content of instruction for example PowerPoint presentations, graphical visualizations, computer tutorials with limited interactive features.

Achievement: According to Ollendick, & Schroeder (2003), ‘‘academic achievement is the knowledge and skills that an individual learns through direct instruction’’ (p. 1). For the current meta-analysis, the author uses any objective measure of academic performance (e.g., exams or test scores) including but not limited to standardized and validated measures to estimate student achievement.

Ethical Considerations and Access to the Studies

The current meta-analysis is a part of an on-going large meta-analysis conducted by Bernard et al. (2014). The selected studies included in the current meta-analysis is a part of Bernard et al. (2014) collection of studies. Consent has been granted from Bernard and colleagues to access and use these studies before the start of this meta-analysis.

Procedures

Formulating the problem/ research questions. The current project primarily attempts to explore the differential effects of high versus low/no level of collaboration on student achievement. Of the additional interest is to decipher further the moderators which promote and /or hinder the effect of collaborative on student achievement. The author formulated three research questions to investigate the effect of collaborative learning on achievement:

1) Does collaborative learning have any statistically significant effect on student achievement outcomes when compared learning without (or under lesser degree of) collaboration?
2) Do different types of technology have varying effects on student achievement when used to enhance/support /promote collaboration?
3) Do grade levels and subject domains have any moderating effects on student`s achievement?

To answer these questions, a quantitative approach was used. More specifically, the author utilized a meta-analysis approach. Collaboration, as the independent treatment variable, used to measure its effect on student achievement. In addition, an analysis of moderator variables was conducted to examine whether moderators such as specific forms (purposes) of technology use, subject demographics, and grade levels could help or hinder collaboration that would affect student achievement outcomes.

Meta-analysis as a methodology. In general, a meta-analysis (quantitative synthesis) is a statistical analysis of the results from individual studies that address a common research question, examine systematic sources of differences in results among these studies, and leading to a quantitative summary of the results (Porta, 2008). There are two widely used approaches to a meta-analysis in the research literature. First, the Hunter and Schmidt’s (1990) psychometric meta-analysis and the second is the classic or Glass’ meta-analysis. Accordind to Bangert-Drowns and Rudner (1991*)*, Hunter and Schmidt's approach is a mix-bag of some of the best characteristics of other approaches. In this approach, all studies related to a given problem are collected without any consideration to the qualities of studies. Then, the distribution of effect sizes is corrected for any variability among studies such as sampling error, measurement error, range restriction, and other systematic artifacts. Even if any variability affect the distribution of overall effect size, then, “the effect sizes are grouped into subsets according to preselected study features, and each subset is meta-analyzed separately” (p.3). Unfortunately, this technique is not very feasible for my project, for this requires substantial information from individual studies for accurate correction of effect sizes. In reality, however, such information are not always available in all the studies.

On the other hand, Glass' approach to meta-analysis is much more of conventional. This approach starts with defining questions to be examined, then, collecting studies, coding study features and outcomes, and finally, analyzing the relations between study features and outcomes. In addition, firstly, Glassian meta-analysis applies liberal inclusion criteria and stresses not to

> disregard studies based on study quality a priori; a meta-analysis itself can determine if study quality is related to variance in reported treatment effect. Secondly, the unit of analysis is the study finding. A single study can report many comparisons between groups and subgroups on different criteria. Effect sizes are calculated for each comparison. Thirdly, meta-analysts using this approach may average effects from different dependent variables, even when these measure different constructs (Bangert-Drowns, & Rudner, 1991, p. 3).

For the current meta-analysis, I will use Glassian approach because it is quite robust for the critical re-analysis, its use of conventional statistical tests, and systematic in design.

Study design. The sample for this analysis was selected from an existing pool of studies which belonged to Bernard and colleagues' (2014) onging larger meta-analysis. The current meta-analysis included 20 studies representative of the main research question resulted for a review process involving abstracts and full study analysis of 78 potentially relevant studies. The study compared two groups – higher degree of collaboration (as the experimental condition) with low/no collaboration (the control condition). Additionally, the moderator analyses of the other associative variables such as type and major purpose of technology use, subject domains, duration, and grade levels also conducted to measure the relative influences of these factors on collaboration and for that matter on student achievement.

First, two reviewers coded five studies independently to decide whether the experimental condition of each study satisfy the inclusion criteria of been higher in collaboration than the control condition and featuring educational technology in experimental condition. In addition, twenty-four study features (methodological, substantive, and demographic) were coded for further use in the moderator variable analysis. The average pairwise agreement rate on the initial coding was 84.17% (Cohen's kappa = 0.68). Disagreements were resolved either through discussion between reviewers or by the involvement of a third party. Two studies (Terwel, Oers, Dijk, & Eden, 2009; Mastropieri, Scruggs, Norland, Berkeley, Mcduffie, Tornquist, & Connors, (2006) were excluded from the sample given the absence of technology in those studies. After establishing sufficient reliability, the author reviewed the rest of the original sample. A final sample of 20 included studies yielded 28 independent effect sizes with a total of 2434 participants.

Variables. The current study investigates the effect of collaborative learning on student achievement. As stated previously, a treatment variable collaboration is used (with high in experimental and low or no collaboration in control conditions) to measure its effect on the dependent variable of student achievement. Specifically, the meta-analysis aimed to estimate the weighted average effect size (i.e. how much better – positive effect, or worth – negative effect Experimental group compared to Control group in terms of their respective achievement outcomes). Among other variables, included in analyses as moderators, were technology type and major purpose of use, subject demographics, grade level, and treatment duration.

Searching the literature/ data sources. As mentioned earlier, the current representative sample is a part of the study collection authored by Bernard et al. (2014), comprised of numerous primary research studies identified through extensive systematic literature searches designed and conducted to identify and retrieve studies relevant to the research questions. The literature search involved more than ten electronic databases (e.g. ERIC, EdLib, Education Abstracts, Medline, ProQuest Digital Dissertations & Theses, PsycINFO, British Education Index) branching from previous relevant reviews and tables of content for major educational journals. In addition, the manual Google Internet searches and searches for various conference proceedings to form a pool of relevant studies were performed.

Criteria for inclusion and exclusion and review procedure. The representative sample was selected using several qualifying criteria before including the studies into the current meta-analysis. The criteria entailed, firstly, the studies should be conducted in formal educational settings with varying grade levels from elementary and secondary to higher education. Only empirical studies with collaboration /cooperation either in face-to-face and/or virtual learning/online such as computer-based collaborative learning were included. A set of inclusion criteria, as discussed below, guided the study characteristics required to retain the studies for inclusion. Studies that did not meet the following criteria were excluded from the current meta-analysis.The inclusion of studies needed to have:

- To be published no earlier than 1996;
- To be publically available (or archived);
- Two main factors – collaboration and technology;
- Address students' achievements, as measured for example by final exams, cumulative posttest scores;
- Contain at least one between-group comparison where one group is considered the experimental condition (i.e., high level of collaboration) and the other group the control condition (i.e., lower/ no collaboration);
- Contain sufficient statistical information for effect size extraction.

Coding Study Features. Coded moderator variables (i.e., study features) used to explore between-study variability in effect sizes. The study features used were mainly based on those employed by Bernard et al., 2009 (in studying distance education) and Schmid et al., 2014 (technology integration in postsecondary education). Major study features, in addition to effect-size defining difference in degree of collaboration, include type and purpose of technology uses, presence of technology, subject demographics, grade levels, treatment duration, etc. (See Appendix A for the codebook).

Calculating Effect Sizes. The estimation of effectiveness of collaborative learning was based on extracting relevant effect sizes from included primary studies. In particular, *d*-type standardized mean difference effect size used as the common metric (i.e., Cohen's *d*). The equation for its extraction is expressed as $d = \overline{X}e - \overline{X}c/_{pooled}^{SD}$. In case of non-availability of the descriptive statistics, effect sizes were extracted from inferential statistics, such as *t*-tests, *F*-tests, or exact *p*-values, using conversion equations from Glass, McGaw, & Smith (1981) and Hedges, Shymansky, & Woodworth (1989). To correct for small sample bias, *d*s were converted to the unbiased estimates of *g* (Hedges & Olkin, 1985).

Synthesis of the Effect Size (ES). The synthesis of the data conducted using the random effects model. Model selection is justified by relative non-uniformity of treatment conditions, rather limited, i.e., non-exhaustive nature of the collection of studies, and, thus, likely heterogeneity of the distribution of effect sizes, confirmed in actual analyses (Borenstein, Hedges, Higgins, & Rothstein, 2009; Hedges & Olkin, 1985). The random effects model is used to aggregate and report average effect sizes (g +), standard errors (*SE*), confidence intervals (lower 95th and upper 95th) and z values with associated *p*-values, when systematic variation in the distribution of effect sizes is not assumed and error term (tau-squared) is randomly added to the weights of individual effects.

Furthermore, the analyses of the moderator variables was conducted to ascertain the relative effectiveness of the moderators on the dependent variable according to the so-called mixed model. In this model, the average effect sizes for categories of the moderators were calculated using the random effects model. The variance component *Q*-Between calculated across categories using a fixed effect model (Borenstein et al., 2009). All analyses, including sensitivity and publication bias analysis, were performed in Comprehensive Meta-Analysis™ 2.2.048 (Borenstein, Hedges, Higgins, & Rothstein, 2005). Eventually, a posthoc test (Bonferroni correction) was conducted at the selected levels of the moderator variables.

Chapter Four : Analysis and Results

This chapter entails various stages of the analyses and reporting of the results including an overview of the selected studies, analyses aiming at the publication bias and sensitivity for any outlier effect, and explaining heterogeneity across the included studies using moderator variable analysis of methodological, substantive, and demographic study features.

Descriptors

This section consists of the descriptive data regarding general study information, explanation of the Effect Size (ES) extraction procedures, as well as substantive, methodological and contextual/demographic features of the included studies. Microsoft Word and Excel software were used to classifiy these study features.

General study information. The general study information includes the type and the year of publication. The studies included in this meta-analysis were selected from journal publications, dissertations, and reports. The most frequent type was the journal publication consisting of 16 of the included studies representing 80% of the included sample .

Table 2. Frequency distribution of the type of publication

Type of Publication	Frequency	Relative %
Journal	16	80.00
Dissertation	2	9.09
Report	2	9.09
Total	20	100

As for the year of publication, the included studies were published between the years 1997 and 2010. The years of 2006 and 2008 offered the larger number of studies comprised four and six published studies respectively. The frequency distribution is presented in Table 3.

Table 3. Frequency distribution of year of publication

Publication year	Frequency	Relative %
1997	1	5.00
1999	2	10.00
2002	1	5.00
2004	1	5.00
2006	4	20.00
2007	2	10.00
2008	6	30.00
2009	1	5.00
2010	2	10.00
Total	20	100

Then the publication dates were grouped into four units for a broader look. In the years between 2006 and 2010, fifteen studies (75%) published. The frequency distribution within respective time frames is presented in Table 4.

Table 4. Frequency distribution of time frame of publication

Publication year	Frequency	Relative %
1997-1999	3	15.00
2000-2004	2	10.00
2006-2007	6	30.00
2008-2010	9	45.00
Total	20	100

Substantive Characteristics. The primary premise established for this meta-analysis was the investigation of the effect of collaborative learning on student achievement. The operational definition for the variable of collaboration is presented in the codebook. A higher level of collaboration was the necessary and sufficient criterion to distinguish the experimental conditions from control ones. Among the moderator variables, technology was the most important factor. The technology was further classified either as instrumental in enabling/promoting/supporting collaboration among students or as just contextually present, without any apparent influence on student collaborative work. Given the criterion of high collaboration in the experimental condition, the moderator analysis was conducted to know the role of technology in supporting, scaffolding and/or enhancing collaboration and to determine what effect it has on student achievement. In the experimental condition, technology was investigated via its types, purposes, and instrumental values when it was intentionally integrated into the collaborative activities.

Table 5 in the next section maps out the types of technology included and their purposes of use. In eight studies (40%), technology was used in a mixed manner, i.e. used for more than one purpose. In five studies representing 25% of the total, technology was used for cognitive support II (i.e., deep learning, e.g., simulations) and in three studies, technology was utilized for the cognitive support I (i.e., distributed cognition, e.g., Excel) representing 15% of the total. Interestingly, out of the final 20 included studies for analysis, 16 studies found a link between collaboration and technology forming 80% of the collection. Only four studies were found with no connection. Among the major types of technology utilized were web-based computer-driven technology including ICT and other software applications followed by multimedia and videos.

Table 5. Shows the collaboration, use of technology and its principal purpose, and the link between collaboration & technology

Reference	Collaboration	Technology type	Technology Purpose (Major)	Link between collaboration & technology
Arts, J. A. et al. (2006)	Low	Web-based CMC & Multimedia/ Internet technology	1	YES
Barnes, L. J. (2008)		Audience Response Systems-Qwizdom & Videos	6: 1+5	YES
Cavalier, J. & Klein, J. D. (1998)		CBI: Hypercard (interactive) tutorial Computers	6: 1&5	YES
DePriter, T. N. (2008)		Computers & web-based discussion forum	6: 1&4	NO
Faro, S., & Swan, K. (2006) _1	Low	Web-based (on-line) study materials and chat &Studio/ CBI/ Online technology	6: 1&4	YES
Faro, S., & Swan, K. (2006) _2	Low	N/A	N/A	YES
Freeman, S. et al. (2007) _1		Clickers	6: 3&4	YES
Freeman, S. et al. (2007) _2		Clickers	6: 3&4	YES
Freeman, S. et al. (2007) _3		Cards	6: 3&4	YES
Freeman, S. et al. (2007) _4		Cards	6: 3&4	YES
Gersten, R. et al. (2006) _1		Video & videos	4	YES
Gersten, R. et al. (2006) _2		Video & videos	4	YES
Hernández-Ramos, P. et al. (2009)	Low	mPower software	3	YES
Hodges, T. L. (2008)		Web CT for administer the exam &Board game	2	YES
Hoon, T. S. et al. (2010)	Low	Computer/courseware	3	YES
Hosal-Akman & Simga-Mugan (2010) _1	High	PowerPoint/slides	5	YES
Hosal-Akman & Simga-Mugan (2010) _2	High	N/A	N/A	YES

Hummel, H. G. et al. (2006)		"Plea checker" computer program & Virtual Program/ emails	2	YES
Lin, J. M., Wu, C., & Liu, H. (1999) _1		SimCPU software package for computer labSimCPU (Computer software)	3	NO
Lin, J. M., Wu, C., & Liu, H. (1999) _2		SimCPU software package for computer	3	NO
Nugent, G. et al. (2008)		Mobile library, digital cameras & digital cameras	6: 3&4	YES
Olgun, O. S., & Adali, B. (2008)	Low	Internet sites & Internet search tools	4	YES
Pariseau, S. E. et al. (2007)		Computer applications (e.g., Excel) & Laptops	2	YES
Priebe, R. (1997)		Burton Comprehension Instrument (BCI) & Propositional Logic Test (PLT)	3	NO
Tsai, M. (2002) _1		Computers & electronic Bulletin Board System (BBS)	6: 1&2	NO
Tsai, M. (2002) _2		Computers & (BBS)	6: 1&2	NO
Tsai, M. (2002) _3		Computers & (BBS)	6: 1&2	NO
Tsai, M. (2002) _4		Computers & (BBS)	6: 1&2	NO
Wenk, M. et al. (2008)		Mannequin (Simulator)	3	YES
Zumbach, J. et al. (2004)		Interactive MS PowerPoint & PPT	6: 3&5	YES

Legend: Technology use: 1 = Communication/interaction, 2 = Cognitive support (distributed cognition, e.g. Excel, Word, SPSS),
3 = Cognitive support (deep understanding - e.g., simulations, knowledge creation), 4 = Informational resources
5 = Presentation, 6 = A mixture of max two (should be really two major purposes where one cannot be singled out - e.g., 6: 2&5)
N/A = 999 Missing information.
Link between collaboration and technology: Yes, and NO

The moderator analyses of the grade level, subject domain, and treatment duration were conducted to investigate their impact on student achievement resulting from instructional interventions in question. As for grade level, the included studies targeted all the grade levels (from kindergarten to post-secondary). The undergraduate level was the most frequent with ten studies forming 50% of the total collection. The second highest was high school representing four studies (20%). It is worth noting that while calculating the individual grade levels, the entire collection of 20 studies was considered. Table 6 shows the grade level distribution.

Table 6. Frequency distribution of grade level addressed in the studies

Grade level (s)	Frequency	Relative %
Kindergarten	0	0.00
Elementary (2)	3	15.00
Middle school (3)	3	15.00
High School (4)	4	20.00
Undergraduate (5)	10	50.00
Graduate (6)	0	0.00
Total	20	100

Legend: Elementary = 2, Middle school = 3, High school = 4, Undergraduate = 5, Graduate = 6

Regarding the subject matter (Table 7), two categories were formd – STEM and Non-STEM, encompassing a large variability of individual disciplines. Twelve studies related to STEM represented 60% of the population. The Non-STEM category included eight studies forming 40% of the total collection. STEM included the subjects in the domains of science, math, technology, and engineering, while Non-STEM comprised subject categories related to humanities, social sciences, languages, and arts.

Table 7. Frequency distribution of subject matter addressed in the studies

Subject domain (s)	Frequency	Relative %
STEM	12	60.00
Non-STEM	8	40.00
Total	20	100

STEM = Science, Technology, Engineering and Math,
Non-STEM = Social sciences & Humanities etc.

Effect Size Extraction and Synthesis

Procedure for calculating ES

For the ES calculation, the author used the Cohen`s *d* metric (1988), based on the division of the mean differences by the pooled standard deviations of both groups. The equations and formulas used for the calculation of ES can be found in Table 8. The information for ES calculation was gathered from means, precisely reported standard deviations, and sample sizes for the experimental and control conditions (Shymansky & Woodworth, 1989). To correct for small sample bias, *d* **was** converted to the unbiased estimator *g* (Hedges & Olkin, 1985). In the case of non-availability of the descriptive statistics, the ES was extracted from inferential statistics, such as *t*-tests, *F*-tests, or exact *p*-values, using conversion equations from Glass, McGaw & Smith (1981), and Hedges, Shymansky & Woodworth (1989).

Additionally, Cohen`s (1988) benchmark was used for the qualitative assessment of the magnitude of an ES. It states three types of magnitude of an ES: (a) $d \geq 0.20 \leq 0.50$ = small effect; (b) $d > 0.50 \leq .080$ = medium effect and (c) $d > 0.80$ is called a large effect. However, Valentine & Cooper (2003) warned against this type of fallacy saying that in the domain like education even smaller ES can be considered effective

Table 8. Study- level statistics used in meta-analysis and explanations (adapted from Bernard et al. 2014)

Equation number	Equation name	Equation	Explanation
Eq. 1	Cohen's d (standardized difference effect size)	$d = \frac{\overline{X}_E - \overline{X}_C}{SD_{Pooled}}$	Cohen's d is the basic unit of effect size in meta-analyses that compare an experimental condition with a control condition on a continuous-level dependent variable. The numerator is the ± difference between the means of the experimental condition and the control condition. The denominator is shown in Eq. 2
Eq. 2	Pooled SD	$SD_{Pooled} = \sqrt{\frac{(n_E - 1)SD^2 + (n_C - 1)SD^2}{(n_E + n_C) - 1}}$	The pooled standard deviation of the experimental and control conditions' standard deviations is the denominator of d-type effect sizes
Eq. 3	Hedges' g ($df = N - 1$) Correction for small sample size	$g = d\left(1 - \frac{3}{4df - 1}\right)$	Cohen's d is called a biased estimator because it does not correct for low sample size that tends to inflate their effect size. Hedges' g is the unbiased estimator
Eq. 4	Standard Error of g	$se_g = \sqrt{\frac{n_e + n_c}{n_e n_c} + \frac{d^2}{2(n_e + n_c)}} \cdot \left(1 - \frac{3}{4df - 1}\right)$	The unbiased standard error of g, based largely on sample size, is the estimated "standard deviation" in the population
Eq. 5	Variance of g (Within-study variance)	$v_g = se^2$	The standard error (Eq. 4) is converted to a variance by squaring it
Eq. 6	z test (test statistic)	$z_g = \frac{g}{se_g}$	The test statistic z is constructed by dividing the effect size g by the standard error of g (se_g). It tests the null hypothesis (without degrees of freedom) that $g = 0$ (does not exceed chance expectations)
Eq. 7	Two-tailed test of z_g $^+(\alpha = .05)$	Null ($g = 0$): $z_g \geq$ or $\leq \pm 1.96$ ($\alpha = .05$)	The two-tailed z value null hypothesis $g = 0$ is tested using $z = 1.96$ ($p = .025$) as the critical value
Eq. 8	95th confidence interval	$Lower\ 95th = g - (1.96 \cdot SE_g)$ $Upper\ 95th = g + (1.96 \cdot SE_g)$	The upper and lower boundaries of the 95th confidence interval define the range within which the effect size is likely to reside. Intervals that cross zero (+ and − limits, or the reverse) are judged to be not significantly different from 0. This interpretation should match the z test

Synthesis of ES. An analytical approach of the random effects model (Borenstein et al., 2009; Hedges & Olkin, 1985) was chosen for this meta-analysis. In the random effects model, effect sizes are weighted by the inverse of the sum of their within-study variance (*Vi)* and average between-study variance (tau-squared). This resulted in no between-study variance left unaccounted for after the analysis is performed. The random effects model was used to interpret and report average effect sizes (*g*+), standard errors (SE), confidence intervals (lower 95th and upper 95th) and *z* values with relative *p*-values. In addition, a fixed effect model was used to estimate total between-study variability (Q-Total) and test for heterogeneity. I^2 (i.e., the percentage of heterogeneity in effect sizes exceeding chance sampling expectations, e.g. Higgins, Thompson, Deeks, & Altman, 2003). A total of 28 effect sizes were extracted from the twenty studies. Four studies, namely, Faro, & Swan (2006) ; Freeman, O'connor, Parks, Cunningham, Hurley & Haak et al. (2007) ; Gersten, Baker, Smith-Johnson, Dimino, and Peterson (2006) ; Lin, Wu & Liu (1999), and Tsai (2002) produced more than one independent effect size.

Results

Sample size. The current meta-analysis is a part of the ongoing larger meta-analysis conducting by Bernard et al. (2016). A collection of 78 studies was considered for the current project. Only empirical studies that address collaboration/cooperation either in face-to-face, real-life classroom and/or virtual learning/online such as computer-based collaborative learning are included. A set of inclusion criteria below presents the study characteristics used to retain the studies for the meta-analysis. The included studies needed to be published no earlier than 1996 and be publically available (or archived), and necessarily feature some form of student collaborative work as the major instructional variable. Moreover, they must have contained at least one between-group comparison where one group is considered the experimental condition (i.e., higher level of collaboration) and the other group the control condition (i.e., lower/ no collaboration) and contains sufficient statistical information for effect size extraction.

The primary abstract reviews resulted in twenty-two eligible studies according to the set criterion (see criterion for inclusion/exclusion in the methods section). After this, reviews of the full-text was conducted of the selected studies to ensure compliance with all of the project inclusion criteria. A close analysis of the studies reported two studies (Mastropieri et al., (2006) and Terwel et al., (2009) as ineligible given missing data, and they were therefore removed from the collection of the final analysis. A total of twenty-eight effect sizes were extracted from twenty studies involving 2434 participants. Two groups – collaborative learning environment with high collaboration versus traditional instruction setting with less or no collaboration – were compared for the relative effect on student achievement. The twenty studies entailed a variety of collaborative activities to measure their impact on student learning outcomes. Table 9 presents the list of included studies in this meta-analysis with the title for each.

Table 9. Studies included in the meta-analysis with titles

Reference	Title (s)
Arts, J. A. et al. (2006)	Enhancing problem-solving expertise by means of an authentic, collaborative, computer supported and problem-based course
Barnes, L. J. (2008)	Lecture-Free High School Biology Using an AUDIENCE RESPONSE SYSTEM
Cavalier, J. & Klein, J. D. (1998)	Using Cooperative Learning and Objectives with Computer-Based Instruction
DePriter, T. N. (2008)	Individual or collaborative learning: An investigation of teaching strategy in the distance learning mathematics classroom
Faro, S., & Swan, K. (2006)	An investigation into the efficacy of the studio model at the high school level
Freeman, S. et al. (2007)	Prescribed Active Learning Increases Performance in Introductory Biology
Gersten, R. et al. (2006)	Eyes on the Prize: Teaching Complex Historical Content to Middle School Students with Learning Disabilities
Hernández-Ramos, P. et al. (2009)	Learning History in Middle School by Designing Multimedia in a Project-Based Learning Experience
Hodges, T. L. (2008)	Examination of gaming in nursing education and the effects on learning and retention
Hoon, T. S. et al. (2010)	Effect of an Interactive Courseware in the Learning of Matrices
Hosal-Akman & Simga-Mugan (2010)	An assessment of the effects of teaching methods on academic performance of students' in accounting courses
Hummel, H. G. et al. (2006)	Effects of cueing and collaboration on the acquisition of complex legal skills
Lin, J. M., Wu, C., & Liu, H. (1999)	Using SimCPU in cooperative learning laboratories
Nugent, G. et al. (2008)	The Impact of a Field-Based, Inquiry-Focused Model of Instruction on Preservice Teachers' Science Learning and Attitudes
Olgun, O. S., & Adali, B. (2008)	Teaching Grade 5 life Science with a Case Study Approach

Pariseau, S. E. et al. (2007)	The Effect of Using Case Studies in Business Statistics
Priebe, R. (1997)	The Effects of Cooperative Learning in a Second-Semester University Computer Science Course.
Tsai, M. (2002)	Do male students often perform better than female students when learning computers? A study of Taiwanese eighth graders' computer education through strategic and cooperative learning
Wenk, M. et al. (2008)	Simulation-based medical education is no better than problem-based discussions and induces misjudgment in self-assessment
Zumbach, J. et al. (2004)	Using multimedia to enhance problem-based learning in elementary school

Inter-rater reliability. Two trained raters were involved in the reviewing and coding of studies throughout the entire process of this meta-analysis. Following is the agreement rates regarding each stage:

- Study features coding – 81% ($k = 0.62$)
- Effect size calculation (for accuracy of data extraction, selection and application of equations) – 96.06%, ($k = 0.92$).
- Effect size comparison decisions (for defining the treatment and control conditions and number of ES and data sources to use) – 91.66% ($k = 0.83$)

Achievement Outcomes. In response to the research question, "Does collaborative learning have any statistically significant effect on student achievements outcomes?" a statistically significant random weighted effect size of $g+ = 0.266$ was found. It is a low-to-moderate positive effect size according to Cohen`s standards (For detail see Cohen, 1988). Comprehensive Meta-Analysis™ software (Borenstein et al. 2005) was used to carry out analyses and derive the outcomes. Main results are presented in Table 10. The summary statistics are based on $k = 28$. It shows an average effect size with both fixed and random effects models. The table depicts the lower limit, and upper limits of the CI (0.09 and 0.44 respectively) and the z-value along with the two-tailed probability is 2.93. The fixed weighted average effect of $g+ =$

0.281 which is also low to moderately low average effect size according to Cohen`s standards. Overall the weighted average effect sizes for both random and fixed models are significant with $p < 0.001$. The heterogeneity statistics as shown for the fixed effect model, *Q*-Total = 99.67, $p < .001$ and $I^2 = 72.91\%$ which reports that almost 73% variability is due to real differences in the effect size and only 27% can be attributed to the sampling error. The *Q*- value statistics validates that the distribution is significantly heterogeneous. This magnitude of between-study variability, according to Higgins, Thompson, Deeks, and Altman (2003) is moderately high.

Table 10. Overall weighted average random effects and fixed effects sizes & heterogeneity

Analytical Models				95th Confidence interval	
	K	g+	SE	Lower limit	Upper limit
Random Effects Model	28	0.266*	0.09	0.09	0.44
Fixed Effects Model	28	0.281**	0.05	0.19	0.37
Heterogeneity	Q-total = 99.67, df = 27, p < .001			I2 = 72.91%	Ʈ2 = 0.16

Z* =2.93, p < .001; Z** = 6.17, p < .001

Research design and methodological quality. It was necessary to ensure that the methodological quality of the included studies did not substantially affect the outcomes of this meta-analysis. The first thing was to ascertain if there was any difference in primary studies' research design that might have favored one category of study methodological quality over others. With this intention, research design of each study was reviewed. The study designs of the included studies in this meta-analysis were of a true experimental (randomized control trials) nature representing 25% of total effect sizes (the gold standard) and quasi-experimental (adjusted and refined by various means of statistical control) representing 75% of the total effect sizes. For the quality methodological checks, the author used Valentine and Cooper`s (2008) instrument called The Study Design and Implementation Assessment Device (Study DIAD). The moderator variable analysis comparing two types of research design (RCT vs. QE) was not statistically significant (Q-value = 3.47, p = .063). Though the *p* value ($p = 0.63$) was quite close to affect the design, the study research design did not differentially bias the findings of the meta-analysis.

The second concern was to verify if the psychometric quality and representativeness of the assessment tool had not affected the outcomes of the meta-analysis. For the analysis, two major measure types – single cumulative and calculated average of several complementary measures – were used. In this regard, the mixed effects analysis by measure source was conducted. The results showed no statistical significance ($p = .152$ with one study removed) regarding any effect on overall effect size. One selected single measure which did not belong to either category of measure type was removed. This did not affect any significant change in the outcomes (*Q*-value = 2.05). Similarly, the effect size extraction procedure ($p = .562$ with one study removed) proved non-significant. The ES extraction involved two procedures – ES precisely calculated from reported descriptive or inferential statistics and ES estimated with various reasonable assumptions.

Furthermore, there might have been other factors that could have affected the outcomes given the methodological quality of the analysis (e.g., instructor's equivalence, the equivalence of content materials). Considering these factors, the author collapsed the several aspects of the methodological qualities (including previously described individually tested qualities of research design, assessment instruments, and extraction procedures) of the studies to enable a composite analysis. The mixed effects analysis of this composite also reported a non-significant effect *Q*-value = 0.24 ($p = .624$).

Publication Bias Analysis. According to Rothstein, Sutton and Borenstein (2005), an analysis of publication bias helps to ensure if a sizeable number of studies might have been missed which could have otherwise been included in the analysis. Given the inclusion of these missing studies and considering the fact of accommodating them with more positive findings by journals (Polanin, Tanner-Smith, & Hennessy, 2015), some statistical procedures became warranted to examine if there is any reason to question the robustness of the average effect size.

In this case, the author used the funnel plot for the visual inspection and two statistical procedures – classic fail-safe analysis (for nullifying the average effect) and Orwin`s fail-safe (for the trivial effect of magnitude) – to verify the publication bias for the current meta-analysis. The visual inspection of the funnel plot depicted in Fig. 1, the Funnel plot, showed the symmetrical dispersion of effect sizes around the mean of the distribution ($g^+ = 0.266$). The following analytical statement about publication bias analysis appears in Comprehensive Meta-Analysis™.

> This meta-analysis incorporates data from 28 studies, which yield a *z*-value of 5.72 and corresponding 2-tailed *p*-value of 0.000. The fail-safe N is 212. This means that we would need to locate and include 212 'null' studies for the combined 2-tailed *p*-value to exceed 0.050. The Orwin fail-safe N is 51. This means that we would need to locate 51 studies with mean Hedges' *g* of 0 to bring the combined Hedges' *g* under 0.1. The Trim and Fill (Duval and Tweedie, 2004) analysis also reported the same pattern of inclusiveness. Using these parameters, the method suggests that no studies are missing. Under the fixed effect model the point estimate and 95% confidence interval for the combined studies is 0.28123 (Lower 95^{th} = 0.19184, Upper 95^{th} = 0.37061). Under the random effects model the point estimate and 95% confidence interval for the combined studies is 0.26566 (Lower 95^{th} = 0.08815, Upper 95^{th} = 0.44317). Using Trim and Fill these values are unchanged ($p = 0.000$).

The author judged and concluded that there was no publication bias present which might impact the effect size adversely resulting in skewness or any negotiation with the results.

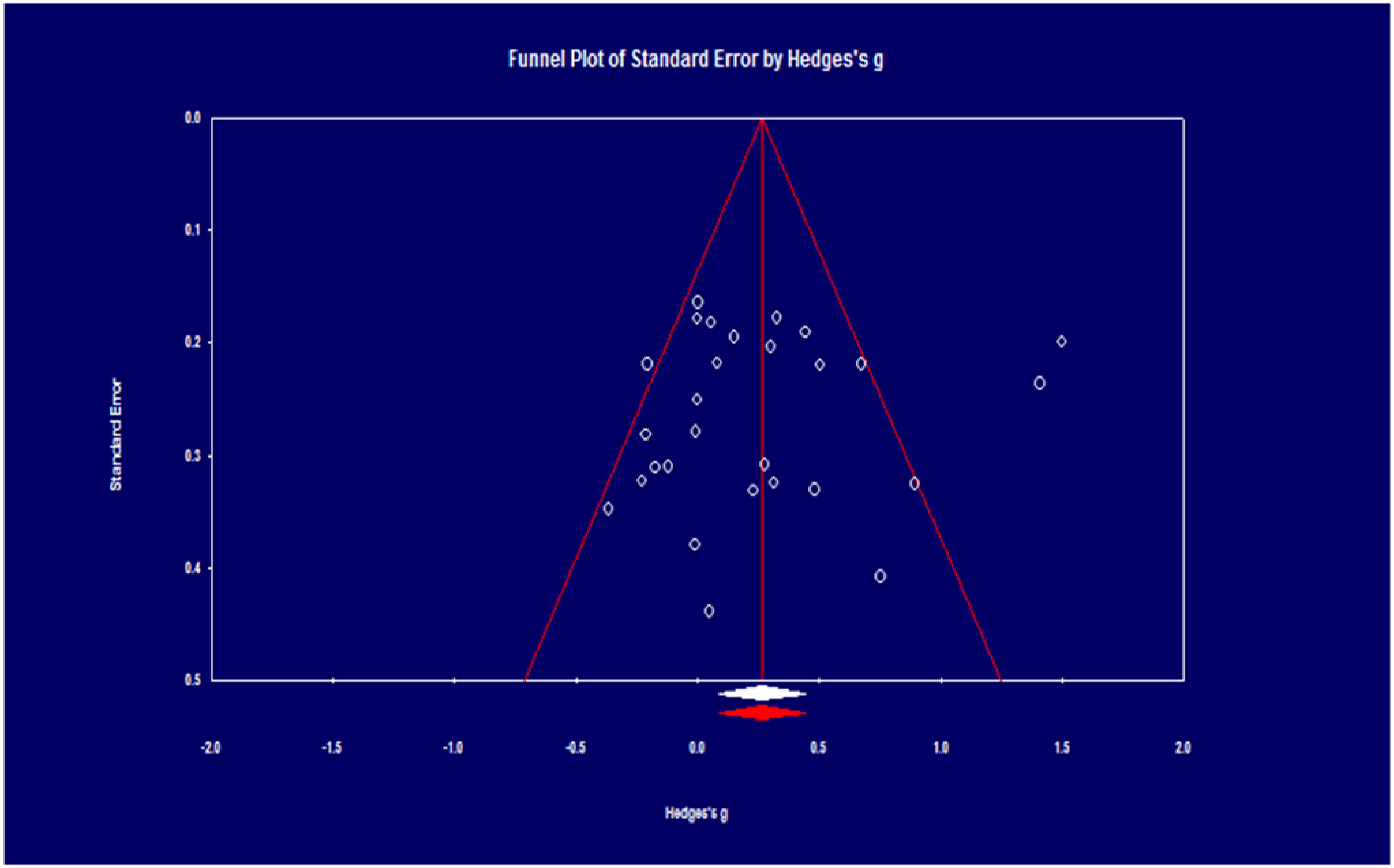

Figure 1. Funnel plot with effect size for using random model

Sensitivity Analysis. The sensitivity analysis aimed to eschew the distorting effects (both overall mean and variability) due to the presence of any outlier in the included studies. The author encountered one study (Hernández-Ramos, & Paz, 2009) with a comparatively higher positive ES (g = 2.534). The reason for this outlier was not known given the missing information in the study. Therefore, the author reduced the magnitude of this aberrant effect size to the second highest effect size, within the range of other large ES by using Comprehensive Meta-AnalysisTM. This adjustment of ES resulted in g = 1.409 which was in the range of other effect sizes in the distribution. After the outlier adjustment, the newly calculated averages fall within the confidence interval of the total collection g+ = 0.266 (k = 28, SE = 0.09, Lower 95th = 0.09 and Upper 95th 0.44). After this, the data were found quite stable and unaffected by any outliers for the further analysis.

Figure 2. depicts the Forest plot with individual and overall ESs for the included studies. On the left side of the figure are the study identifications, in this case, the author names. In the center are the study-level statistics for the twenty-eight ES : Hedges *g*, the standard error, the variance, the upper and lower boundaries of the 95^{th} confidence interval, the *z* value, and its relative probability (*p*-value). A visual representation called a Forest plot is on the right side of the figure. The ES for each study are depicted in the shape of squares. The lines around squares show the width of the 95th confidence interval for each study. The *z*-test of these effect sizes was significant ($p > .05$). Furthermore, the smaller dots represent, the lower leverage effect sizes (i.e., smaller contributors to the weighted averages ES), while larger dots demonstrate the higher leverage effects characterized by larger sample size.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Hernández-Ramos, P., & Paz, S. D.	1.500	0.200	0.040	1.109	1.891	7.518	0.000
Olgun, O. S., & Adali, B.	1.409	0.236	0.056	0.946	1.872	5.965	0.000
Faro, S., & Swan, K.	0.895	0.325	0.106	0.258	1.532	2.752	0.006
Nugent, G., Kunz, G., Levy, R., Harwood, D.et al.	0.754	0.408	0.166	-0.045	1.554	1.849	0.064
Lin, J. M., Wu, C., & Liu, H.	0.676	0.219	0.048	0.247	1.104	3.091	0.002
Lin, J. M., Wu, C., & Liu, H.	0.503	0.220	0.048	0.072	0.934	2.290	0.022
Gersten, R., Baker, S. K., Smith-Johnson, J. et al.	0.482	0.331	0.109	-0.166	1.130	1.457	0.145
Pariseau, S. E., & Kezim, B.	0.445	0.191	0.036	0.071	0.819	2.333	0.020
Hoon, T. S., Chong, T. S., & Ngah, N. A. B.	0.328	0.178	0.032	-0.020	0.677	1.847	0.065
Faro, S., & Swan, K.	0.316	0.325	0.105	-0.321	0.952	0.972	0.331
Hodges, T. L.	0.303	0.204	0.041	-0.097	0.702	1.485	0.137
Bames, L. J.	0.277	0.308	0.095	-0.326	0.880	0.900	0.368
DePriter, T. N.	0.230	0.332	0.110	-0.420	0.880	0.694	0.487
Freeman, S., O'connor, E., Parks, J. W. et al.	0.152	0.195	0.038	-0.230	0.534	0.779	0.436
Freeman, S., O'connor, E., Parks, J. W. et al.	0.081	0.218	0.048	-0.346	0.509	0.373	0.709
Hosal-Akman, N., & Simga-Mugan, C.	0.056	0.182	0.033	-0.300	0.413	0.309	0.757
Hummel, H. G., Paas, F., & Koper, R.	0.051	0.439	0.193	-0.810	0.911	0.115	0.908
Hosal-Akman, N., & Simga-Mugan, C.	0.003	0.164	0.027	-0.318	0.324	0.017	0.986
Freeman, S., O'connor, E., Parks, J. W. et al.	0.000	0.251	0.063	-0.492	0.492	0.000	1.000
Cavalier, J. C., & Klein, J. D.	0.000	0.178	0.032	-0.350	0.350	0.000	1.000
Zumbach, J., Kumpf, D., & Koch, S.	-0.006	0.279	0.078	-0.552	0.540	-0.022	0.983
Arts, J. A., Gijselaers, W. H., & Segers, M. S.	-0.011	0.380	0.144	-0.756	0.733	-0.030	0.976
Gersten, R., Baker, S. K., Smith-Johnson, J. et al.	-0.120	0.310	0.096	-0.728	0.488	-0.386	0.700
Tsai, M.	-0.175	0.311	0.096	-0.784	0.433	-0.564	0.573
Freeman, S., O'connor, E., Parks, J. W. et al.	-0.207	0.219	0.048	-0.636	0.223	-0.944	0.345
Priebe, R.	-0.211	0.282	0.080	-0.763	0.342	-0.747	0.455
Tsai, M.	-0.227	0.323	0.104	-0.860	0.406	-0.703	0.482
Wenk, M., Waurick, R., Schotes, D. et al.	-0.366	0.348	0.121	-1.047	0.315	-1.053	0.292
	0.281	0.046	0.002	0.192	0.371	6.167	0.000

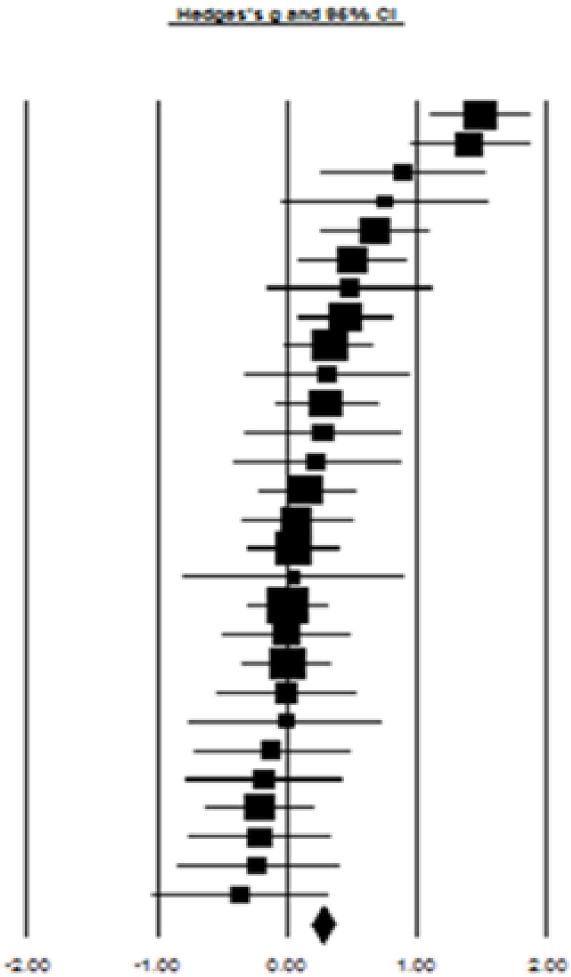

Figure 2. Forest plot of 28 effect size from the distribution (*k* =20) and study-level statistics

Table 11. shows the mixed effects analysis by treatment (i.e., Collaboration) strength. It relates the degree of difference between two groups in collaboration. It clearly depicts that average effect size with a high degree of difference in collaboration (k = 9) with g = 0.35 which is significantly different from zero and average effect size g = 0.29 (k = 12) that is moderately different while an average effect size g = 0.08 (k = 7) which is at a low level of difference in collaborative work. Q-Between = 2.46 was not statistically significant (p = .293). This analysis shows an increasing trend in effect size magnitude with the difference in collaboration in two conditions across groups.

Table 11. Mixed-effects analysis of collaboration strength

Level of Collaboration	Effect size			95th Confidence interval	
	k	*g*+	SE	Lower limit	Upper limit
Low	7	0.082	0.10	-0.12	0.28
Moderate	12	0.299	0.15	0.00	0.60
High	9	0.357	0.18	0.00	0.71

Q-Between = 2.46 (df = 2, p = 0.293)

To further explain the detected variability in g^+, four moderators – technology (both regarding its relevance to collaborative work and its functionality), subject matter, grade levels, and treatment duration – were analyzed. The moderator analysis of technology types was conducted using the mixed model analysis (Borenstein et al., 2009). Two qualitative categories about the added instrumental value were created – Yes and No. Yes (k = 22) favored the instrumental value of technology in collaboration (i.e., technology was an important factor in enabling and supporting collaborative learning activities), while No category (k = 6) (merely reflected some contextual presence of technology). However, an upward trend for the instrumental value of technology was discernible in the analysis (see Table 12). The moderator analysis of the instrumental value of technology in collaboration was not statistically significant (*Q*-Between = 0.34, *p* = .558). This trend may inform that though technology had added value in collaboration, it was not statistically significant to affect the overall effect.

Table 12. Mixed effects analysis by the instrumental value of technology

				95th Confidence Interval		
Groups	Number Studies	*g+*	SE	Lower limit	Upper limit	Q-Between
No	6	0.17	0.17	-0.16	0.51	
Yes	22	0.29	0.10	0.08	0.50	
Overall	28	0.26	0.09	0.08	0.43	
Between groups	*df* = 1 *z* = 2.87					*0.34* *p* = 0.558

Legend: No = No use of technology in collaboration,
Yes = Use of technology in collaboration

To dig deeper into the matter, another analysis for the types of cognitive support provided by technology was conducted (see table 13). By cognitive support, the author means the use of technology for two types of supports – cognitive support for distributed cognition (e.g. Word, Excel, and SPSS) and cognitive support for deep learning/understanding (e.g. simulations and knowledge creation). The *Q*-Between = 0.17 with *p* = 0.681 was not statistically significant for the type of cognitive support. When tested for the presence of technology for cognitive support (distributed and deep learning), the trend refers that *k* = 17 did not favor both distributed and deep learning and only *k* =10 favored the value of technology for both types of cognitive support.

Table 13. Mixed-effects analysis of cognitive support for Distributed Cognition & Deep Learning

				95th Confidence Interval		
Groups	Number Studies	*g+*	SE	Lower limit	Upper limit	Q-Between
No	10	0.31	0.16	0.00	0.63	
Yes	17	0.23	0.12	0.00	0.47	
Overall	27*	0.26	0.10	0.08	0.45	
Between groups,	df = 1 z = 2.76					0.17 p = 0.681

* ES 27 since one study was excluded from this analysis given the missing information (999)
Legend: No = No cognitive support of the use of technology in collaboration
Yes = Cognitive support offered by the use of technology in collaboration

Next, the author decided to explore the use of technology in collaboration for cognitive support for deep learning only. To perform the analysis, the frequency were calculated. This analysis was also not statistically significant (*Q*-Between = 0.02, $p = 0.893$). Table 14 demonstrates the use of technology for the cognitive support for deep learning only.

Table 14. Mixed effects analysis using technology for the cognitive support for deep learning

				95^{th} Confidence Interval		
Groups	Number Studies	*g+*	SE	Lower limit	Upper limit	Q-Between
No	15	0.25	0.11	0.03	0.47	
Yes	12	0.27	0.16	-0.03	0.58	
Overall	27*	0.26	0.09	0.08	0.44	
Between groups,	*df* = 1 *z* = 2.81					0.02 *p* = 0.89

* ES 27 since one study was excluded from this analysis given the missing information
Legend: No = No cognitive support, Yes = Cognitive support offered for deep learning

Furthermore, the author investigated if the subject domains had any moderating effects on the overall effect of collaboration. Table 15 shows that there was no difference in learning across STEM and Non-STEM subjects. Collaborative learning is equally effective across all subject domains. The Q-between = 0.00 with $p = 0.992$ inferred a statistically non-significant effect of subject domains. This result informed that collaboration is not limited only to STEM domains. Table 15 portrays an equal effect of collaboration across Non-STEM ($g = 0.26$ with $k = 10$) and STEM ($g = 0.26$ with $k = 18$) subject domains. Notwithstanding this outcome, it is important to note the the classification of a course being STEM versus non-STEM is problematic, so any interpretation of this result should be qualified.

Table 15. Mixed effects analysis by moderator variable subject matter

				95th Confidence Interval		
Category	Number Studies	*g+*	SE	Lower limit	Upper limit	Q-Between
Non STEM	10	0.26	0.18	-0.09	0.62	
STEM	18	0.26	0.10	0.06	0.46	
Overall	28	0.26	0.09	0.08	0.44	
Between groups,	df = 1 z = 2.92					0.00 p = 0.992

Non-STEM = Social Sciences, Humanities and STEM: Science, Technology, Engineering and Math

Furthermore, the author tried to explore the effect of duration on collaboration. Three groups of duration were formulated to measure the impact of duration on overall effect size. As Table 16 depicts the effect of duration on collaborative learning was not statistically significant

(Q-between = 3.27, p = 0.19). The In Between group (k = 14) produced showed the highest impact of collaboration (g = 0.43) in comparison to short (k = 3) and semester length (k = 11).

Table 16. Mixed effects analysis by treatment duration

				95th Confidence Interval		
Category	Number Studies	***g+***	SE	Lower limit	Upper limit	Q-Between
In Between	14	0.43	0.17	0.10	0.76	
Semester	11	0.09	0.07	-0.04	0.23	
Short	3	0.13	0.12	-0.09	0.36	
Overall	28	0.14	0.06	0.03	0.25	*3.27*
Between groups,	*df* = 2 *z* = 2.55					*p* = 0.195

Legend: More than three days but less than eight weeks = In Between
Nine weeks or more = semester, Three days or less = short

Next, the author had committed to conducting a moderator analysis of grade level to investigate its effect on collaboration. The analysis included grades from Elementary school to Undergraduate levels. There were no eligible studies found from Kindergarten and Graduate levels in the collection according to the inclusion and exclusion criteria. Even though the variation in grade level of education significantly differentiated student achievement outcomes (Q-Between = 11.18, p = 0.011), the small k sizes (high school - k = 6; elementary k = 3; and middle school - k = 5) were such that any further interpretation of this outcomes was abamdoned due to lack of statistical power.

Chapter Five : Discussion

This section involves the interpretations of the evidence in the form of discussion about the results obtained. The author here discusses the results to inform the three research questions which guided this meta-analysis and to underline any possible conceptual, theoretical and/or practical implications of the findings. This discussion is orchestrated in the backdrop of the literature of the domain and in the light of previous studies.

The purpose of the current meta-analysis was to measure the effect of collaborative learning on student achievement in the context of formal education across multiple subject domains and grade levels. As mentioned earlier, it included a total of 28 effect sizes from a collection of 20 studies. The main research question was: Does collaborative learning have any statistically significant effect on student achievement outcomes when compared learning without (or under lesser degree of) collaboration? The most explicit statement drawn from the analysis is that the effect of collaborative learning on student achievement is positive but low, though significantly greater than zero. Regarding percentile difference (i.e. U_3 minus 50%), 60% of students yielded an increase in scores, or a person with an average (50th percentile) could expect to move to the 60th percentile after participating in a collaborative learning group. The average effect size of $g^+ = 0.266$ ($k = 28$) is a little higher than the low category in terms of Cohen's (1988) qualitative effect size magnitude (i.e. $g^+ > 0.20 < 0.50$). However, given the considerable degree of heterogeneity among studies, it was difficult to fix the exact location of the population mean aside from the probability that it is located between $g^+ = 0.09$ and 0.44 (i.e. lower and upper levels of 95th confidence interval respectively).

The next few paragraphs discuss the findings in the light of previous studies undertaken to measure the effects of collaborative/cooperative learning on student achievement . The findings of the current study are consistent with Johnson et al. (1988, 1998) who found high positive effects $g^+ = 0.78$ and $g^+ = 0.49$ ($k = 122$ and $k = 168$ respectively) of cooperative/competitive conditions on student achievement. In comparison to the Bowen's study

(2000) that found a statistically significant effect ($g^+ = 0.51$, $k = 37$) of cooperative learning on student achievement in high school and college chemistry courses, the current study relates the same positive effect i.e. findings are consistent. Springer's (1999) findings suggested that various forms of small-group learning are effective in promoting greater academic achievement in case of STEM courses and reported a statistically significant effect of small-group learning on student achievement ($g^+ = 0.51$, $k = 37$). Here, again the results of the current study are in line with Springer's findings. Further, the findings of the current analysis are consistent with Lou et al. (2001) who found that small group learning had significantly more positive effects than individual learning on student individual achievement ($g^+ = 0.15$, $k = 122$). Therefore, this consistent positive trend indicates that collaborative learning helps enhance student achievement considerably.

Comparatively, the average effect size of the current study ($g^+ = 0.26$) is in the middle of achievement effect of $g^+ = 0.17$ for Lou et al. (1996) and Bernard et al. (2009) which is $g^+ = 0.38$. This difference may be accounted for by the number of studies and the variables incorporated in those meta-analyses. For example, Bernard et al. (2009) were interested in the three types of interactional conditions in distance education. However, the current study aimed at the investigation of the effect of collaborative learning in all forms of instruction in multiple subject levels across all grade levels. Also worth noting are the findings of Borokhovski et al. (2016) on the use of collaborative activities to support student-student interaction in a technology rich environment, i.e., $g^+ = 0.52$ ($k = 25$) is much higher in magnitude. The possible explanation for this difference may be again be the number of studies included with multiple collaborative conditions, use and purpose of technology in instructional design, grade levels, and the primary research questions asked.

The second research question was: Do different types of technology have varying effects in collaborative activities when used to enhance/support/promote collaboration? A study by Schmid et al. (2014) reported a statistically significant effect of technology for cognitive support. However, the findings of the current study are not consistent with Schmid et al. (2014). The use of technology for collaborative activities did not favor cognitive support in distributed cognition and deep learning, collectively. This difference in the findings could be attributed to the difference in the use and degree of technology in classroom. For example, in the current study, technology as moderator was analyzed to measure its effect on collaborative learning rather than the effect of technology on student achievement. In addition, small *k*-sizes and varied treatment conditions might have resulted in different outcomes in both of the studies.

Interestingly, the use of technology for deep learning was also not statistically significant in the current study. On the contrary, Tamim, Bernard, Borokhovski, Abrami, & Schmid, (2011), found an average effect size $g^+ = 0.28$ ($k = 25$) while investigating the effect of technology on student achievement. They tried to map out the exclusive effect of technology on student achievement. However, unlike both Tamim et al., (2011), and Schmid et al., (2009, 2014), the current study aimed at the exploration of technology as a moderator to enhance/support/promote collaborative learning affecting student achievement . The possible explanation for the different findings in these studies could be attributed to the insufficient sample size, lack of training for both teachers and students, and use of technology for secondary purposes other than the collaborative activities.

In some studies, technology was used for cognitive support in the treatment condition. For example, Wenk, Waurick, Schotes, Gerdes, Aken, & Pöpping, (2008) used an electronic mannequin (a life size simulator) to enhance deep learning on the processes of medical treatment. Similarly, in other studies (Hoon et al., 2010 ; Hernández-Ramos et al., 2009; Tsai, 2002; Priebe, 1997; Freeman et al., 2007; Faro et al., 2006) technology was used to support the deep learning and for information resources.

In the context of the current findings, it would be interesting to reflect on the ongoing great debate between Clark (1983, 1994) and Kozma (1994). What Clark (1983, 1994, 2009) claims is that role of technology in learning is minimal (or negligible). Instead, it is the nature of pedagogy (for example deep learning) and learning design (learning environment) that matter in teaching and learning process irrespective of any mode of technology use. Kozma claims that technology is helpful for learners to remember, seek information, and to collaborate. However, these claims are too big to generalize from the findings of current studies given the lack the statistical power.

The third research question was: Do grade levels, and subject domains have any moderating effects on student achievement? As noted above, the lack of statistical power for this analysis resulted in abandoning any further discussion that might misinform the literature. Hence, there is a need for further exploration to understand what factors influence collaboration at various grade levels.

Regarding the subject domain, there was no significant difference found between STEM and Non-STEM comparisons groups. Students achieved almost equally in both domains. Therefore, the findings relate that collaborative activities can impact student learning outcomes across all subjects as opposed to the findings of Lou et al., (1996) and Qin et al., (1995), who reported that STEM influenced more significantly student achievement. The rationale for the outcome of current study is that collaborative activities are employed almost equally to enhance interactions among learners across STEM and Non-STEM subject domains. Therefore, collaborative activities serve as a means to create learning environment rather than as an end to maximize any subject specific content.

Further, the author decided to investigate if the duration of treatment had any impact on the determination of the degree of collaboration. The findings of the current analysis indicated that duration as moderator was not statistically significant which differed from the findings of other studies. In the between-study comparison group, however, there was an indication that students liked moderate duration (In Between) for collaborative activities. This outcome is consistent with what Fisher (1981) says that students' interest and choice of the content may determine their inclination toward specific duration. In this regard, the construct of academic learning time is important to predict student achievement. It is so because, for example, allocated time, engagement rate, and success rate of school activities are all associated with student achievement. This signifies that more academic learning time can be interpreted as helpful to an ongoing measure of student learning.

Limitations of the Study

This study has encountered some general and specific limitations. First, only a small number of the studies qualified for the final sample meant for the analysis. Here is the biggest limitation of this study because the number of samples included (*k*) were very low. Consequently, the low *k* affects the power of the study to find significant effects, especially in moderator variable analysis where the total number of samples is split by the number of levels of the moderator variable. Therefore, the generalizability of the findings of this study cannot be established given this small sample.

Second, the common critique on meta-analyse is the heterogeneity of the included studies; each study is often conducted in different conditions (for example different treatment, different variables, and a diverse population of participants). However, the results of a meta-analysis are combined as if they were similar (Eysenck, 1994). In this situation, it is not always easy to compare the results included in a reliable manner and the interpretation may be biased. As Higgins (2008) says, the heterogeneity between studies is to be expected and is acceptable ‘‘providing both that the predefined eligibility criteria for the meta-analysis are sound and that the data are correct’’ (p. 1158). The heterogeneity can be attributed to the variability found between studies regarding the participants, interventions effects and outcomes studied (diversity and statistical heterogeneity). This heterogeneity affects the overall average effect size of an analysis. In the current analysis , the author encountered one study (Hernández-Ramos, & Paz, 2009) with a comparatively higher positive ES (g = 2.534) which was adjusted to make the distribution more representative.

Third, the study might have publication bias regarding the exclusion of studies published in the languages other than English. In addition, the accommodation of the studies with more positive findings in journals could have affected the representativeness of the sample (Polanin, Tanner-Smith & Hennessy, 2015).

Implications and Future Directions

The primary purpose of the current meta-analysis was to explore the effect of collaborative learning on student achievement. In general, the collaborative learning was found favourable to enhance student achievement. The analysis reported some implications as to how collaborative activities with what combination of conditions yield better learning outcomes for students. In collaboratively designed instruction, students outperformed their control counterparts. Collaborative learning activities are beneficial in that these help enhance student achievement and persistence, (e.g. Bowen, 2000) change attitude and self-concept, and support those students who feel fearful while participating in classroom activities (e.g. Kyndt et al. 2013).

Next ,the differential use of technology in collaborative activities was not found to matter. Though technology was used in both groups for various purposes, the impact of different technology on student achievement was not significant. This outcome may provide opportunities for future researchers to explore the questions of what types of technology and what contexts help teachers design collaborative activities which can enhance student achievement.

Among the other future implications include, firstly, the cultural differences among learners impact their degree of collaboration considerably (Kyndt et al., 2013). Culture, as a moderator which has not been explored in this analysis, may be added for future exploration. It is so because the exploration of culture as moderator will add more insights into the factors which are conducive/detrimental for collaborative learning. The cultural differences between Western/individualistic and Eastern/collectivistic cultures may have significant influences on the ways students cooperate in the learning activities. Studies have found that ‘‘cross-cultural studies have shown that Northern and Western Europeans and North Americans tend to be individualistic and that Chinese people, other Asians, Latin-Americans, and most East and West

Africans tend to be collectivists'' (Cox, Lobel, & McLeod , 1991, p. 828). This means that the Western approach of cooperative learning embraces critiquing opinions by challenging each other's reasoning and dealing with conflicts which may be culturally inappropriate for collectivistic cultures. For example, Thanh et al. (2008) found that the incongruity between cooperative learning principles and Asian culture accounted partially for the failure of cooperative learning.

Secondly, the aspects such as the forms, contexts and purposes of any selected collaborative activities and the roles of a teacher (e.g. either facilitator, partner or observer) during the process of collaboration are other major areas to guide how cooperative learning can be used to enhance student achievement.

Thirdly, the investigation of the use of technology for secondary purposes will be helpful, as these secondary purposes trigger students' interests to participate in classroom activities. This means that how students' previous exposures to various forms of social technology such as Facebook, Twitter, Instagram, or Snapchat may help them collaborate in their learning activities (Phua, Jin, & Kim, 2017; Kim & Kim, 2017).

A fourth direction for future researchers may be to verify whether collaborative activities are favourable among lower grade levels in comparison to higher educational levels. It would be interesting to unpack the favourable conditions and types of tools which enhance participation among post-secondary students for collaborative activities. An extensive study is warranted to investigate the factors that influence collaboration at all levels.

Lastly, Bowen (2000) collected data on persistence from nine studies and found that cooperative learning has a significant and positive effect on student attitudes towards STEM courses. There is no clear cut definition of the construct of persistence. In the educational context, a student persistence entails the capacity which allows students to continue his efforts through self-regulation, motivation, and positive affirmation to the achieve the set goals. The degree of persistence dictates the student achievement. Bowen's (2000) meta-analysis reported that persistence for continued study in STEM courses taught with cooperative learning approaches was 22% greater than persistence of students taught by traditional approaches. Students in cooperative learning classes also had more positive attitudes toward their classes (p.11). Therefore, it would be interesting to investigate other personality characteristics such as "persistence or fear" (e.g. Bowen, 2000) which affect student collaboration.

Chapter Five : Conclusion

The overall effect of collaborative learning on student achievement was positive and significant despite some limitations of the study. However, there are some moderators which impact student achievement. The analysis found that collaborative activities organized at different school levels may affect student achievement. Technology in its various forms are used in classroom, however, these forms are integrated purposefully in instructional and curricular designs (Borokhovski et al., 2016). Embeding technology in pedagogy may help improve student collaboration and thereby their social, and academic success. Instead of incorporating technology as a mere extension or ancillary in instructional designs, technology might be added in curricula to improve student collaboration.

Furthermore, the subject domains and treatment duration are important to understand the impact of collaborative learning on student achievement. STEM and Non-STEM subjects can be taught equally successfully when one uses collaborative methodology. While duration dictates the level of collaboration among learners, however, the reasons for favouring small duration over longer duration depend on learners' interests, teacher, available time, and nature of content. Again, an understanding of learners' needs and interests may help them choose their best options in this regard.

On the whole, the findings of this meta-analysis are valuable for teacher-educators and curriculum designers to take informed decisions on the conditions and forms of collaborative activities to be included while planning, designing, and implementing effective instructional strategies. For example, a teacher may employ collaborative activities based on arts, culture and local issues related to science, environment and health to develop group projects. These exercises will serve two purposes - knowledge creation and development of critical skills.

Further, these findings may guide teachers in making choices of types and purposes of technology use in their classroom. For instance, the use of augmented and virtual reality can help create an environment to simulate scientific and natural phenomena such as study of stars and galaxies, earthquakes and earth. Findings on the other contextual factors such as grade levels and subject domains along with duration of treatment will help inform subject experts and researchers to improve student learning. Finally, these findings have added to the knowledge of the domains and have replicated the results of the previous studies.

References

Note: A leading asterisk (*) indicates the primary studies included for the meta-analysis.

Abrami, P. C., Bernard, R. M., Bures, E. M., Borokhovski, E., & Tamim, R. (2011). Interaction in distance education and online learning: Using evidence and theory to improve practice. *Journal of Computing in Higher Education*, 23(2/3), 82-103.

Adelskold, G., Alklett, K., Axelsson, R., & Blomgren, G. (1999). Problem-based distance learning of energy issues via computer network. *Distance Education*, 20(1), 129-143.

Anderson, T. (2003). Getting the mix right again: an updated and theoretical rationale for interaction. *International Review of Research in Open and Distance Learning*, 4(2), 9-14.

*Arts, J. A., Gijselaers, W. H., & Segers, M. S. (2006). Enhancing problem-solving expertise by means of an authentic, collaborative, computer supported and problem-based course. *European Journal of Psychology of Education, 21*(1), 71-90.

Arts, J. A. R., Gijselaers, W. H., & Segers, M. S. R. (2002). Cognitive effects of an authentic computer-supported, problem-based learning environment. *Instructional Science*, *30*, 465-495.

Baeplar, P., Walker, J.D., & Driessen, M. (2014). It's not about seat time: Blanding, flipping, and efficiency in active learning classrooms. *Computers & Education*, 78, 227-236.

Bangert-Drowns, R. L., & Rudner, L. M. (1991*). Meta-Analysis in Educational Research. ERIC Digest.* Retrieved from ERIC database. (ED339748)

*Barnes, L. J. (2008). Lecture-free high school biology using an audience response system. *The American Biology Teacher, 70*(9), 531-536. doi: 10.2307/27669338

Beldarrain, Y. (2006). Distance education trends: Integrating new technologies to foster student interaction and collaboration. *Distance Education*, 27(2), 139-153.

Bernard, R. M., Abrami, P. C., Borokhovski, E., Wade, A., Tamim, R., Surkes, M., et al. (2009). A meta-analysis of three interaction treatments in distance education. *Review of Educational Research*, 79(3), 1243-1289.

Bernard, R. M., Abrami, P. C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., et al. (2004). How does distance education compare to classroom instruction? A Meta-analysis of the empirical literature. *Review of Educational Research*, 74(3), 379-439.

Bernard, R. M., Rojo de Rubalcava, B., & St-Pierre, D. (2000). Collaborative online distance education: Issues for future practice and research. *Distance Education,* 21(2), 260-277.

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.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. (2005). *Comprehensive meta-analysis*. Englewood, NJ: Biostat.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, *1*(2), 97-111.

Borokhovski, E., Bernard, R. M., Tamim, R. M., Schmid, R. F., & Sokolovskaya, A. (2016). Technology-supported student interaction in post-secondary education: A meta-analysis of designed versus contextual treatments. *Computers & Education, 96*, 15-28.

Borokhovski, E., Tamim, R. M., Bernard, R. M., Abrami, P. C., & Sokolovskaya, A. (2012). Are contextual and design student-student interaction treatments equally effective in distance education? A follow-up meta-analysis of comparative empirical studies. *Distance Education*, 33(3), 311-329.

Bowen, C. W. (2000). A quantitative literature review of cooperative learning effects on high school and college chemistry achievement. *Journal of Chemical Education*, 77(1), 116-119.

Bruffee, K. (1995). Sharing our toys. *Change*, 27(1), 12-18.

*Cavalier, J. C., & Klein, J. D. (1998). Effects of cooperative versus individual learning and orienting activities during computer-based instruction. *Educational Technology Research and Development*, 33(3), 52-72.

Chou, S., & Min, H. (2009). The impact of media on collaborative learning in virtual settings: The perspective of social construction. *Computers & Education, 52*(2), 417-431.

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research,* 53, 445–449.

Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development*, 42(2), 21–29.

Cox, T., Lobel, S., & McLeod, P. (1991). Effects of ethnic group cultural differences on cooperative and competitive behaviour on a group task. *Academy of Management Journal*, 34, 827-847. http://dx.doi.org/10.2307/256391.

Cobb, T. (1997). Cognitive efficiency: toward a revised theory of media. *Educational Technology Research and Development*, 45(4), 21-35.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Coombs, P. with Ahmed, M. (1974). *Attacking Rural Poverty,* Baltimore: The John Hopkins University Press.

Cooper, J., Prescott, S., Cook, L., Smith, L., & Mueck, R. (1990). *Cooperative learning and college instruction*. California: California State University.

Curtis, D. D., & Lawson, M. J. (2001). Exploring collaborative online learning. *JALN, 5*(1), 21-34.

Dalton, D. W., Hannafin, M. J., & Hooper, S. (1989). The effects of individual versus cooperative computer-assisted instruction on student performance and attitudes. *Educational Technology Research and Development*, 37(2), 15-24.

Decuyper, S., Dochy, F., & Van den Bossche, P. (2010). Grasping the dynamic complexity of team learning. An integrative systemic model for effective team learning. *Educational Research Review,* 5, 111-133.

Dede, C. (1996). Emerging technologies and distributed learning. *American Journal of Distance Education,* 10(2), 4-36.

Dede, C. (2004). Enabling distributed learning communities via emerging technologies. *The Journal*, 32(2), 12-22.

*DePriter, T. N. (2008). *Individual or Collaborative Learning: An Investigation of Teaching Strategy in the Distance Learning Mathematics Classroom*. ProQuest.

Dillenbourg, P. (1999). What do you mean by 'collaborative learning?' In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and Computational Approaches* (pp.1–19). Oxford: Elsevier.

Dillenbourg, P., Baker, M., Blaye, A., & O'Malley, C. (1996). The evolution of research on collaborative learning. In E. Spada & P. Reiman (Eds.), *Learning in humans and machine: Towards an interdisciplinary learning science* (pp. 189-211). Oxford: Elsevier.

Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. (2002). Effects of problem based learning: A meta-analysis. *Learning and Instruction*,13, 533-568.

Duval, S. & Tweedie, R. (2000b). A nonparametric ‘‘trim and fill’’ method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, 95, 89-98.

Duval, S., & Tweedie, R. (2004). Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2), 455-463.

Eastmond, D. (1995). *Alone but together: Adult distance study through computer conferencing*. Cresskill, NJ: Hampton Press.

Eysenck, H. J. (1994). Systematic reviews: Meta-analysis and its problems. *British Medical Journal*, 309, 789-792.

*Faro, S., & Swan, K. (2006). An investigation into the efficacy of the studio model at the high school level. *Journal of Educational Computing Research, 35*(1), 45-59.

Fisher, C. W. (1981). Teaching behaviors, academic learning time, and student achievement: An overview. *Journal of Classroom Interaction*, 17(1), 2-15.

Fjermestad, J. (2004). An analysis of communication mode in-group support systems research. *Decision Support Systems*, 37(2), 239-263.

*Freeman, S., O'connor, E., Parks, J. W., Cunningham, M., Hurley, D., Haak, D., . . . Wenderoth, M. P. (2007). Prescribed active learning increases performance in introductory biology. *Cell Biology Education, 6*(2), 132-139.

*Gersten, R., Baker, S. K., Smith-Johnson, J., Dimino, J., & Peterson, A. (2006). Eyes on the prize: Teaching complex historical content to middle school students with learning disabilities. *Exceptional Children, 72*(3), 264-280.

Glass, G.V, McGaw, B., & M.L. Smith (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.

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

Hedges, L. V., Shymansky, J. A., & Woodworth, G. (1989). *A practical guide to modern methods of meta-analysis.* Washington, DC: National Science Teachers Association.

*Hernández-Ramos, P., & Paz, S. D. (2009). Learning history in middle school by designing multimedia in a project-based learning experience. *Journal of Research on Technology in Education*, 42(2), 151-173.

Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers & Education, 90*, 36-53.

Higgins, J. P. (2008). Commentary: Heterogeneity in meta-analysis should be expected and appropriately quantified. *Journal of Epidemology*, 37, 1158-1160.

Higgins, J. P. T., Thompson, S. G., Deeks, J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, 327, 557–560.

*Hodges, T. L. (2008). *Examination of gaming in Nursing Education and the effects on learning and retention*. ProQuest.

Hooper, S. (1992). Cooperative learning and computer-based instruction. *Educational Technology Research and Development*, 40(3), 21-38.

Hooper, S., & Hannafin, M. J. (1991). The effects of group composition on achievement, interaction, and learning efficiency during computer-based cooperative instruction. *Educational Technology Research and Development,* 39(3), 27-40.

*Hoon, T. S., Chong, T. S., & Ngah, N. A. B. (2010). Effect of an Interactive Courseware in the Learning of Matrices. *Educational Technology & Society*, *13*(1), 121-132.

* Hosal-Akman, N., & Simga-Mugan, C. (2010). An assessment of the effects of teaching methods on academic performance of students in accounting courses. *Innovations in Education and Teaching International, 47*(3), 251-260. doi:10.1080/14703297.2010.498176

Hoskins, S. L., & van Hoof, J. C. (2005). Motivation and ability: Which students use online learning and what influence does it have on their achievement. *British Journal of Educational Technology*, 36(2), 177-192.

*Hummel, H. G., Paas, F., & Koper, R. (2006). Effects of cueing and collaboration on the acquisition of complex legal skills. *British Journal of Educational Psychology, 76*(3), 613-631.

Hunter, J.E., & F.L. Schmidt (1990). *Methods of meta-analysis*. Newbury Park, CA: Sage.

Jehng, J. C. J. (1997). The psycho-social processes and cognitive effects of peer-based collaborative interactions with computers. *Journal of Educational Computing Research*, *17*, 19-46.

Johnson, D. W., & Johnson, R. (1989). *Cooperation and competition: Theory and research.* Edina, MN: Interaction Book Company.

Johnson, D. W., & Johnson, R. T. (1996). Cooperative learning and traditional American values: An appreciation. *NASSP Bulletin*, 80(579), 63-66.

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher*, 38, 365-379.

Johnson, D. W., Johnson, R. T., & Smith, K. A. (1998). Cooperative learning returns to college: What evidence is there that is works? *Change,* 30, 26–35.

Jonassen, D. H., & Kwon, H. I. (2001). Communication patterns in computer-mediated and face-to-face group problem-solving. *Educational Technology Research and Development*, 49, 35-51.

Johnson, D. W., Maruyama, G., Johnson, R. T., Nelson, D., & Skon, L. (1981). Effects of cooperative, competitive and individualistic goal structures on achievement: A meta-analysis. *Psychological Bulletin,* 89, 47–62.

Jonassen, D., Prevish, T., Christy, D., & Stavulaki, E. (1999). Learning to solve problems on the Web: Aggregate planning in a business management course. *Distance Education*, 20(1), 49-63.

Kanuka, H., & Anderson, T. (1998). On-line interchange, discord, and knowledge construction. *Journal of Distance Education*, 13(1), 57-74.

Kanuka, H. & Anderson, T. (1999). Using constructivism in technology mediated learning: Constructing order out of the chaos in the literature. *Radical Pedagogy*, 1(2).

Keeler, C. M., & Anson, R. (1995). An assessment of cooperative learning used for basic computer skills instruction in the college classroom. *Journal of Educational Computing Research, 12*(4), 379-393.

Kim, M.-S., & Kim, H.-M. (2017). The effect of online fan community attributes on the loyalty and cooperation of fan community members: The moderating role of connect hours. *Computers in Human Behavior, 68*, 232-243.

Kirschner, P. A., & Erkens, G. (2013). Toward a framework for CSCL research. *Educational Psychologist*, 48(1), 1-8.

Klein, J. D., & Doran, M. S. (1999). Implementing individual and small group learning structures with a computer simulation. *Educational Technology Research and Development*, 47(1), 97-110.

Kozma, R. (1991). Learning with media. *Review of Educational Research*, 61, 179-221.

Kozma, R. (1994). Will media influence learning: reframing the debate? *Educational Technology Research and Development*, 42(2), 7-19.

Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. *Computers in Human Behavior, 19*(3), 335-353. doi:http://dx.doi.org/10.1016/S0747-5632(02)00057-2

Kulik, C. & Kulik, J. A. (1991). Effectiveness of computer-based instruction: An updated analysis. *Computers in Human Behavior*, 7, 75-74.

Kyndt, E., Raes, E., Lismont, B., Timmers, F., Cascallar, E., & Dochy, F. (2013). A meta-analysis of the effects of face-to-face cooperative learning. Do recent studies falsify or verify earlier findings? *Educational Research Review, 10*, 133-149.

Lai, E. R. (2011). Collaboration: A Literature Review (Rep.). Retrieved from Pearson website: http://www.pearsonassessments.com/research

Lee, K. (2007). Online collaborative case study learning. *Journal of College Reading and Learning,* 37(2), 82-100.

*Lin, J. M., Wu, C., & Liu, H. (1999). Using SimCPU in cooperative Learning laboratories. *Journal of Educational Computing Research, 20*(3), 259-277.

Lou, Y., Abrami, P. C., & D'Appolonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research, 71*(3), 449-521.

Lou, Y., Bernard, R. M., & Abrami, P. C. (2006). Media and pedagogy in undergraduate distance education: A theory-based meta-analysis of empirical literature. *Educational Technology Research & Development*, 54(2), 141-176.

Lou, Y., Abrami, P. C., Spence, J. C., Poulson, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research*, 66, 423-458.

Mantri, A., Dutt, S., Gupta, J. P., & Chitkara, M. (2008). Design and evaluation of a PBL-based course in analog electronics. *IEEE Transactions on Education*, *51*, 432-438.

Margaryan, A., Bianco, M., & Littlejohn, A. (2015). Instructional quality of massive open online courses (MOOCs). *Computers & Education*, *80*, 77-83.

*Mastropieri, M. A., Scruggs, T. E., Norland, J. J., Berkeley, S., Mcduffie, K., Tornquist, E. H., & Connors, N. (2006). Differentiated curriculum enhancement in inclusive middle school science: effects on classroom and high-stakes tests. *The Journal of Special Education, 40*(3), 130-137.

Mayer, R. E. (2008). *Learning and instruction* (2nd ed.). Upper Saddle River, NJ: Merrill Prentice-Hall.

Moore, M. G. (1989). Three types of interaction. *American Journal of Distance Education*, 3(2), 1-7.

Nelson, C. E. (1994). Critical thinking and collaborative learning. *New Directions for Teaching and Learning,* 1994 (59), 45-58.

Newmann, F., & Thompson, J. A. (1987). *Effects of cooperative learning on achievement in secondary schools: a summary of research*. Madison, WI: University of Wisconsin-Madison, National Center on Effective Secondary Schools.

*Nugent, G., Kunz, G., Levy, R., Harwood, D., & Carlson, D. (2008). The impact of a field-based, inquiry-focused model of instruction on preservice teachers' science learning and attitudes. *Electronic Journal of Science Education*, *12*(2).

Ollendick, T., & Schroeder, C. S. (2003). Encyclopedia of clinical child and pediatric psychology. NY: Kluwer Academic/Plenum Publishers.

*Olgun, O. S., & Adali, B. (2008). Teaching grade 5 life science with a case study approach. *Journal of Elementary Science Education, 20*(1), 29-44.

Panitz, T. (1996). A definition of collaborative vs co-operative learning. Retrieved from http//:www.londonmet.ac.uk/deliberations/collaborative-learning/panitz paper.cfm

Paris, S. G., & Winograd, P. (1990). How metacognition can promote academic learning and instruction. In B. F. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (pp. 15-51). Hillsdale, NJ: Lawrence Erlbaum.

*Pariseau, S. E., & Kezim, B. (2007). The Effect of Using Case Studies in Business Statistics. *Journal of Education for Business, 83*(1), 27-31.

Pedró, F. (2005). Comparing traditional and ICT-enriched university teaching methods: Evidence from two empirical studies. *Higher Education in Europe*, *30*, 399-411.

Phua, J., Jin, S. V., & Kim, J. (2017). Gratifications of using Facebook, Twitter, Instagram, or Snapchat to follow brands: The moderating effect of social comparison, trust, tie strength, and network homophily on brand identification, brand engagement, brand commitment, and membership intention. *Telematics and Informatics, 34*(1), 412-424.

Polanin, J. R., Tanner-Smith, E. E., & Hennessy, E. A. (2015). Estimating the difference between published and unpublished effect sizes: a meta-analysis. *Review of Educational Research*. OnlineFirst, Available from http://rer.aera.net

Porta, M., (2008). (Eds). *A Dictionary of Epidemiology*. Oxford University Press: New York.

* Priebe, R. (1997). *The Effects of Cooperative Learning in a Second-Semester University Computer Science Course*. Paper presented at the National Association for Research in Science Teaching Annual Meeting 1997. https://www.learntechlib.org/p/82425

Puzio, K., & Colby, G. T. (2013). Cooperative learning and literacy. *Journal of Research on Educational Effectiveness, 6*(4), 339-360.

Qin, Z., Johnson, D. W., & Johnson, R. T. (1995). Cooperative versus competitive efforts and problem solving. *Review of Educational Research*, 65, 129-143.

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

Schmid, R. F., Bernard, R. M., Borokhovski, E., Tamim, R., Abrami, P. C., Wade, C. A., et al. (2009). Technology's effect on achievement in higher education: A stage I meta analysis of classroom applications. *Journal of Computing in Higher Education*, 21(2), 95-109.

Sherman, G. P., & Klein, J. D. (1995). The effects of cued interaction and ability grouping during cooperative computer-based science instruction. *Educational Technology Research and Development,* 43(4), 5-24.

Slavin, R. E. (1990). Research on cooperative learning: Consensus and controversy. *Educational leadership*, *47*(4), 52-54.

Smith, B. L. & McGregor, J. (1992). What is cooperative learning? In A. S. Goodsell, M. R. Maher,V. Tino, B. L. Smith, & J. McGregor (Eds.), *Collaborative learning: A sourcebook for higher education.*Washington DC: ERIC.

Springer, L., Stanne, M. L., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology. *Review of Educational Research, 69I1), 21-51.*

Susman, E. B. (1998). Cooperative learning: A review of factors that increase the effectiveness of cooperative computer-based instruction. *Journal of Educational Computing Research*, 18(4), 303-322.

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

Thanh, P. T., Gillies, R., & Renshaw, P. (2008). Cooperative learning (CL) and academic achievement of Asian students: A true story. *International Education Studies, 1*(3), 83-88. doi:10.5539/ies.v1n3p82

*Terwel, J., Oers, B. V., Dijk, I. V., & Eeden, P. V. (2009). Are representations to be provided or generated in primary mathematics education? Effects on transfer. *Educational Research and Evaluation,*15(1), 25-44.

Tomcho, T. J., & Foels, R. (2012). Meta-analysis of group learning activities: Empirically based teaching recommendations. *Teaching of Psychology*, 39(3), 159-169.

*Tsai, M. (2002). Do male students often perform better than female students when learning computers? A study of Taiwanese eighth graders' computer education through strategic and cooperative learning. *Journal of Educational Computing Research, 26*(1), 67-85.

Uribe, D., Klein, J. D., & Sullivan, H. (2003). The effect of computer-mediated collaborative learning on solving ill-defined problems. *Educational Technology Research and Development,* 51, 5-19.

Valentine, J. C., & Cooper, H. (2003). *Effect size substantive interpretation guidelines: Issues in the interpretation of effect sizes*. Washington, DC: What Works Clearinghouse.

Valentine, J. C., & Cooper, H. (2008). A systematic and transparent approach for assessing the methodological quality of intervention effectiveness research: the study design and implementation assessment device (Study DIAD). *Psychological Methods*, 13(2), 130-149.

Van Boxtel, C., Van der Linden, J., & Kanselaar, G. (2000). Collaborative learning tasks and the elaboration of conceptual knowledge. *Learning and Instruction, 10*(4), 311-330.

Vygotsky, L.S. (1962). *Thought and language*. Cambridge, MA: MIT Press.

Webb, N.M. (1991). Task-related verbal interaction and mathematical learning in small groups. *Research in Mathematics Education, 22*(5), 366-389.

Webb, N. M. (1993). Collaborative group versus individual assessment in mathematics: Processes and outcomes. *Educational Assessment, 1*(2), 131-152.

Webb, N. M., & Palincsar, A. S. (1996). Group processes in the classroom. In D. Berliner & R. Calfee (Eds.), *Handbook of educational psychology* (pp. 841-873). New York: Macmillan.

Wecker, C., & Fischer, F. (2014). Where is the evidence? A meta-analysis on the role of argumentation for the acquisition of domain-specific knowledge in computer-supported collaborative learning. *Computers & Education*, 75, 218-228.

*Wenk, M., Waurick, R., Schotes, D., Gerdes, C., Aken, H. K., & Pöpping, D. M. (2008). Simulation-based medical education is no better than problem-based discussions and induces misjudgment in self-assessment. *Advances in Health Sciences Education,*14(2), 159-171.

Wilczenski, F. L, Bontrager, T., Ventrone, P., & Correia, M. (2001). Observing collaborative problem-solving processes and outcomes [Electronic version]. *Psychology in the Schools,* 38(3), 269 – 281.

Williams, S. M. (2009). The impact of collaborative, scaffolded learning in K-12 schools: A meta-analysis. Retrieved from Cisco Systems website: https://www.cisco.com/web/about/citizenship/socio-economic/docs/Metiri_Classroom_Collaboration_Research.pdf

Wright, K. B., Kandel-Cisco, B., Hodges, T. S., Metoyer, S., Boriack, A. W., Franco Fuenmayor, S. E., & Waxman H. C. (2013). *Developing and assessing students' collaboration in the IB programme.* Retrieved from http://www.ibo.org/globalassets/publications/ib-research/developingandassessingstudentcollaborationfinalreport.pdf

*Zumbach, J., Kumpf, D., & Koch, S. (2004). Using multimedia to enhance problem-based learning in elementary school. *Information technology in childhood education annual*, *1*(2004), 25-37.

Appendix A

Codebook: based on the meta-analysis conducted by Schmid et al. (2014).

Categories and details of the codebook

Study ID number

Sequential number of the ES

Author(s)

Publication Data

Year of publication

Type of publication

1) Journal

2) Dissertation

3) Conference Proceedings

4) Report/Gray literature

Effect Size Extraction

n for the experimental group

n for the control group

Total *N* (the entire sample size)

Effect size (*d*)

Procedure of ES extraction

1) Calculated using reported descriptive statistics

2) Calculated using reported inferential statistics

3) Estimated from partial inferential statistics, e.g. reported *p*-value

4) Estimated from hypothesis ($p < a$) or assumption of equal sample size when only N is given.

5) ES reported by the authors (only used when no other information is available)

Outcome Information

Outcome type:

Achievements (Exam, GPA, Grades, and tests)

Form of outcome measure

1) Most representative (cumulative) one-time performance measure

2) Reported composite of several tests/evaluations

3) Calculated composite of several assessments reported in the paper

4) Individual measure/item selected to represent the corresponding outcome

Nature of comparison

Brief description of both, experimental and control, conditions and of the source of data for ES extraction (open-ended entry)

Methodological Quality

Research design

1) Quasi-experimental design (QED, non-equivalent groups with control for selection bias, etc.)

2) True experimental design (RCT, random assignment of participants to groups, etc.)

Learner Demographics

Academic level of learners

1 = Kindergarten (KG)

2 = Grades 1-5 (Elementary)

3 = Grades 6-8 (Middle school)

4 = Grades 9-12 (High school)

5 = Higher education: Undergraduate

6 = Higher Education: Graduate

7 = Combination: Specify (e.g., 7: 5+6)

8 = Other: Specify (e.g., 8: Military)

Subject matter

Open entry:

Specify subject matter as reported in the study

For comparison in analysis*:* STEM (1) and Non- STEM (2)

999 = Missing information

Nature of Treatment

Duration of the treatment

More than three days but less than eight weeks = In Between

Nine weeks or more = semester,

Three days or less = short

999 = Missing information

Delivery mode (Ge & Gc)

1 = F2F (Classroom Instruction)

2 = BL *a substantive mix of both modes are used simultaneously

3 = DE (Distance Education)

4 = Fixed Computer automated program: in lab, class or on campus, with or without the presence of a lab assistant

5 = Flexible Computer automated program with no instructor Lab-based (class or campus)

999 = Missing information

Technology Presence (Ge & Gc)

1 = Yes

2 = No

999 = Missing information (???)

Technology type (Ge & Gc)

Open entry:

Please, name technological tool(s) as reported in the study OR N/A - when none

Major purpose of technology use (Ge & Gc)

1 = Communication/interaction

2 = Cognitive support (distributed cognition, e.g. Excel, Word, SPSS)

3 = Cognitive support (deep understanding - e.g., simulations, knowledge creation)

4 = Informational resources

5 = Presentation

6 = A mixture of max two (should be really two major purposes where one cannot be singled out - e.g., 6: 2 & 5)

999 = Missing information (???)

Appendix B

Mixed effects analyses of the methodological features of the studies

Mixed effects analysis by research design

Groups	Number of Studies	Point estimate (g)	Standard error	P-value
QE	21	0.34	0.107	
RCT	7	0.04	0.116	
Total within				
Total between				0.062
Overall	28	0.20	0.079	

Mixed effects analysis by Measure source

(*k* =1 removed)

Groups	Number of Studies	Point estimate (g)	Standard error	P-value
Single cumulative measure	19	0.31	0.12	
Calculated average	8	0.07	0.12	
Total within				
Total between				0.152
Overall	27	0.20	0.08	

Legend: 1 = Single cumulative measure (Final score)
3 = Calculated average (e.g. Projects, assignments)
4 = Single selected measure which was removed)

Mixed analysis by ES extraction Procedure
***(k =1 removed*)**

Groups	Number of Studies	Point estimate (g)	Standard error	P-value
Calculated from Descriptives	21	0.29	0.12	
Estimated with assumptions	6	0.19	0.13	
Total within				
Total between				0.562
Overall	27	0.24	0.09	

Mixed effects analysis by MQ composite

Groups	Number of Studies	Point estimate (g)	Standard error	P-value
High	21	0.28	0.12	
Low	7	0.20	0.10	
Total within				
Total between				0.624
Overall	28	0.24	0.08	