Canadian Journal for New Scholars in Education/
Revue canadienne des jeunes chercheures en education

Volume 5, Issue 1
March 2014

The Effects of Metacognitive Treatments on the Academic Performance of Students with Learning Disabilities: A Meta-Analysis

Bruce R. Maxwell

Concordia University
maxwellb@selwyn.ca

Kian Grenier

Concordia University
kiangrenier@hotmail.com

Abstract

This meta-analysis explores previous research involving metacognitive interventions and the academic performance of students with learning disabilities. From our initial review of over 120 studies, six studies met our inclusion criteria and were included in this study, and moderator variables were explored. Overall, a medium-to-large effect size of 0.79 was found using Cohen's *d* for these final selected six research articles. These results provide preliminary support for the use of metacognitive treatments for students with learning disabilities throughout elementary to postsecondary settings.

Introduction

School success and meeting the academic needs of current generations of students remains an important issue for families, educators, schools, and governments. North American governments have created a number of initiatives, and they have devoted considerable resources, in an effort to improve the quality of public education and to meet the needs of our students (Slavin, 2002).

In Canada, in the 1980's, Canadian teacher's associations created reform movements so that students with special needs could be included in the classroom (Winzer & Mazurek, 2011). Due to federal funding, the research possibilities associated with learning disabilities is increasing in the Education field at Canadian Universities. For the 2010-2011 academic year, the amount in funding was over $27 million, and this ranked fourth out of 12 subject areas (ElAtia, Ipperciel, & Hammad, 2012). In Canada, there is a growing awareness of how government and public funds can together work in order to benefit students with learning disabilities by establishing better structure and support for local governments, institutions, and teachers (Stacey, 2013).

The province of Quebec has demonstrated that there is a smaller proportion of students with disabilities in post-secondary institutions than in comparable institutions in the rest of Canada (Fichten et al., 2003). This is misleading information because it does not represent higher academic standards in the province of Quebec, but a greater tendency for students with learning disabilities to be forgotten, not encouraged to continue, or not enough accommodation in their formative schooling prior to post-secondary years. This suggests that other Canadian provincial governments are more accommodating towards students with learning disabilities.

School failure represents a serious societal concern as "quitting before earning a diploma, certificate, or degree, leaves young people ill-equipped for the world of worry and at high risk for a life of poverty" (Macionis, Jansson, & Benoit, 2012, p. 388). Dropping out is potentially associated with challenges that threaten the individual's financial security and psychosocial well-being, since students who are unable to complete high school successfully tend to have a much more challenging time fitting into society. Students with learning disabilities might be inclined to quit school, and subsequently, have less success when they are adults.

The Supreme Court of Canada ruled in November of 2012 that all students with learning disabilities must have meaningful access to the same education as all other students in British Columbia. This was based on Jeffrey Moore not receiving adequate support for his dyslexia within the North Vancouver School Board when it first came to light in the Fall of 2000. This is a huge victory for all students with learning disabilities (LDAC, 2013, pp.1-3).

A meta-analysis was used in this study to integrate the findings and results of many research studies by "deriving an overall numeric index of the magnitude of the results. The intent of this research is to summarize the results of many studies" (Creswell, 2008, p. 93). The numeric index to calculate the effect size used Cohen's *d*.

Metacognition

Metacognition involves knowledge of the task both in terms of content (i.e. what is learned) and process (i.e. when and how a task is learned), as well as knowledge of personal capabilities, interests, and attitudes. According to Schunk (2012), metacognitive awareness involves procedural knowledge where the learner is actively "monitoring one's level of learning, deciding when to take a different approach, and assessing readiness for a task" (p. 415). The author states that metacognition and self-regulation are roughly equivalent as the learner monitors, directs, and regulates actions towards their goals. Metacognition is governed by the frontal lobe of the brain (e.g. Schunk, 2012, p. 36).

There are many different understandings and interpretations of metacognition which include declarative knowledge or the understanding one's own capabilities (Camahalan, 2008), task knowledge or procedural knowledge which is how one perceives the difficulty of a task which is the content, length, and type of assignment (Borkowski & Cavanugh, 1979; Schunk & Rice, 1993), and strategic knowledge or conditional knowledge, which is one's own capability for using strategies to learn information (Schunk & Zimmerman, 1994, 1998). Other interpretations of metacognition might involve monitoring versus control (e.g. research by Nelson, Narens, & Dunlosky, 2004) or meta-memory, in terms of student recall and subsequent application, which can be connected to other areas of applied research through further studies (e.g. Flavell & Wellman, 1977).

Metacognitive skills involve prediction, planning, monitoring, evaluation, and reflection skills (Desoete & Roeyers, 2006), and students may need help acquiring such skills (Haywood, 2004). Metacognitive awareness can be improved through self-report questionnaires, interviews, stimulated recall situation, and adaptive feedback (Chen, Ho, & Yen, 2010). The researchers in this study assessed metacognition based on task or procedural knowledge, in the sense that a student can understand their own learning through the challenges that are associated with the treatment or required task at hand. This model of self-regulation has the potential through metacognition awareness to allow students to monitor, direct, and regulate actions towards a goal (Paris & Paris, 2001).

Purpose of the Study

The purpose of this study is to examine how metacognitive treatments affect the academic performance of students with learning disabilities. The instruction of learning strategies has been identified as a fundamental way that students with learning disabilities can overcome their needs to achieve academic success (Chan, 1994). Given the potential benefits associated with metacognitive strategy instruction, it is of central importance to understand metacognition and how it can be used to support students with various learning profiles.

Overview of Learning Disabilities

The field of learning disabilities has attracted considerable attention within the research community over the past several decades. The Learning Disabilities Association of Canada (est. 1963) states that learning disabilities involve a number of disorders affecting "the acquisition, organization, retention, understanding or use of verbal or nonverbal information" (LDAC, 2002). According to the Learning Disabilities Association of Canada:

> Learning disabilities are due to genetic and/or neurobiological factors or injury that alters brain functioning in a manner which affects one or more processes related to learning. These disorders are not due primarily to hearing and/or vision problems, socio-economic factors, cultural or linguistic differences, lack of motivation or ineffective teaching, although these factors may further complicate the challenges faced by individuals with learning disabilities. (LDAC, 2002)

Learning disabilities can involve a wide range of cognitive functions and processes involving language, phonology, memory, attention, and executive functioning may be affected. Individuals with learning difficulties, according to the LDAC (2002), may experience challenges with communicating orally with others (e.g. listening, speaking, understanding), reading (decoding words, knowledge of phonemes & morphemes, word recognition, comprehension), written language (spelling, organizing text, expressing thoughts as words), or mathematics (mental sequences, computation, problem solving).

Individuals with non-verbal learning disabilities may have poor motor skills, so handwriting may be uncoordinated and effortful. Tasks involving visual-spatial information and organization may be challenging and students with non verbal learning disabilities may have trouble following multi-step directions, they might have difficulty generalizing previously learned information, and they may ask many questions that are repetitive or inappropriately timed. Socially, such individuals may struggle when recognizing nonverbal cues such as facial expressions, body language, personal space, or appropriate social etiquette.

Verbal and nonverbal learning disabilities are not mutually exclusive and learning disabilities can greatly vary from individual to individual. Individuals with learning disabilities are of average to above average intelligence, are capable of learning and that learning disabilities "are distinct from global intellectual deficiency" (LDAC, 2002). Learning disabilities are lifelong and are highly comorbid with behavioural, emotional, and attentional disorders. Learning disabilities can follow an individual beyond formal educational settings.

In 1980, studies associated with ADHD and learning disabilities were initially reviewed because it coincides with the release of the *Diagnostic and Statistical Manual of Mental Disorders* (*3rd ed*, p. 191). The complexity of learning disabilities have made it so that this work has been re-edited again in 1987, and a new *4th* edition was released in 1994. That *4th* edition was reviewed and updated again in 2000, demonstrating the difficulty with defining and classifying mental disorders.

Statement of the Problem

Given that students with learning disabilities face a set of unique challenges and problems, it is of central importance that we continue to identify and implement research-based practices to support students with various learning profiles become more independent learners, encounter greater school success, and explore postsecondary options (Trainin & Swanson 2005).

Rationale for Our Study

The importance of this study becomes evident when considering current societal issues and recent trends within the academic literature. There has been growing debate regarding how to address the plight of North American public school systems (Slavin, 2002). Given that many students with learning disabilities are not receiving the required support, guidance and resources required for them to encounter academic success. In addition, there are a number of other challenges associated to having a learning disability, and such students may encounter problems in motivation, attributions, self-esteem, and affective responses that can further exacerbate academic difficulties (Hall & Webster, 2008). According to the Learning Disabilities Association of Canada, students with LD require specialized interventions that “need to be appropriate for each individual’s learning disability subtype and, at a minimum, include provision of specific, skill instruction, accommodations, compensatory strategies, and self-advocacy skills” (LDAC, 2002). Consequently, helping students develop metacognition is important because metacognitive skills will help the student identify, reflect upon, assess, practice, and master learning strategies that will help them overcome academic challenges and address their areas of need. This process may be of particular importance because, as Short (1992) notes, some learners inadequately complete academic tasks despite having the required intellectual faculties “either because of a failure to use effective strategies or because of an inability to recognize the utility of a known strategy” (p. 230). In short, the development of metacognitive awareness and strategies that help improve metacognition may provide students with learning disabilities with vital skills that can lead to greater autonomy and improved student success (Stevens & Shenker, 1991).

Research Question

The research question for the current project is as follows: *What are the effects of metacognitive treatments upon the academic performance of students who have learning disabilities?*

Definition of Terms

Independent Variable

Metacognitive treatments refer to interventions, such as a course, workshop, or support sessions, that focus on the instruction of material through the use strategies designed to improve metacognition skills and metacognitive awareness of the learner. Given that the acquisition of metacognition occurs slowly over time, treatments occurring within a single session or with less than a total duration of at least 60 minutes were not considered.

Dependent Variable

Academic performance refers to the student’s performance on an academic task (e.g. comprehension during a reading task, solving a word problem in algebra, asking and answering questions, etc.). Given that metacognition focuses on the process of learning, rather than the content of learning, academic subjects in a variety of educational settings are acceptable.

Population

Elementary school to university students with at least average intelligence who are identified as having a mild to severe learning disability (verbal or non-verbal) that did not result from physical trauma. This excludes students with special needs, such as mental or physical handicaps, diagnosed anxiety and mood disorders, as well as developmental and autism spectrum disorders.

Methods

Research Process

Initial database searches included a hundred and twenty articles that were coded independently. As criteria were refined over time, it was necessary to revisit the databases to ensure that thorough searches of the databases were conducted. Our research design outcomes had to be absolutely compatible in terms of calculated effect sizes that examined the gain scores of LD and NLD samples. This contributed to our decision to consider two group quasi-experimental and true experimental designs. Our final selection for the six studies can be seen in *Appendix A: Information Chart on Meta-Analysis Studies.*

Initial Excel spreadsheet database codes included organizational codes such as *article number*, *Database* origin, *Type* of document (i.e. scholarly article, dissertation, books, conference paper, resource manual, presentation, etc.), locators

(*Researcher 1* or *Researcher 2*), *quantitative* or *qualitative* studies (our focus was to locate the former given the nature of the meta-analysis study), as well as research *design* as identified by Campbell and Stanley (1963). We also included bibliographical codes such as *Author, Year, Title, Journal, Volume,* and *Page #*'s. Finally, we included codes to help ourselves monitor the databases themselves, such as *Accept /Reject* criteria to help identify differences of opinion as well as a *Notes* criteria allowing us to record any observations, reminders, and references. This process helped us monitor inter-reliability and helped us determine if initial searches qualified for full article retrieval.

Statistical items included items, such as *Design Type*, *Label* (of intervention), Group (*Control* or *Treatment*), Number of participants (n size), *Pre-Test Mean, Pre-Test Standard Deviation, Post-Test Mean, Post-Test Standard Deviation, Gain Scores, Pooled Standard Deviation, Es_d* (Effect Size- Cohen's d), *Es_ g* (Effect Size- Hedges's g), *Var_ g* (variance), *Weight_* g, and *w*g* (Hedges's g times the corresponding weight).

Both researchers were coders and database searchers so that our searches could become more rigorous when searching through databases or identifying grey materials. Each of our article databases made use of extensive color coding systems that greatly facilitated our ability to include or reject articles (included in the legend following each database). For instance, an article might have been rejected simply because it did not involve a treatment, or we included an article in the event that we wished to use the document as a potential source of reference. Specifically, the primary colors used included yellow to identify issues that we wished to revisit, green to accept an article, red to reject an article for failure to meet either our inclusion or exclusion criteria, as well as pink to identify a document to be used as a reference.

We also used other colors for issues we wished to revisit at a later date, including light blue to identify links to self-regulation, light grey to identify grey material or unpublished material, orange for consideration of other variables, light purple to denote exclusion criteria, such as special needs (e.g. Pervasive Developmental Disorder (PDD), Autism and Asperger's), and maroon to categorize articles that examined school failure, rather than measures of academic achievement. This coding system is important as it provides us with validation criteria that we needed to agree upon to know how to proceed with each article and we discovered that this process easily allowed us to reduce our list to about twenty articles. We excluded all studies that dealt with self-regulation, learning disabilities that included behaviour problems (e.g. Oppositional Defiance Behaviour), Emotional Disorders (e.g. depression or anxiety disorders), mental intellectual deficits (e.g. PDD), as well as autism and Asperger's syndrome spectrum disorders, since we did not wish to include any of these within our categorization of learning disabilities.

Another important discovery was that the term 'achievement' became problematic. Several studies did indeed examine achievement; however, given that the decision was made to examine students with LD from all levels of formal education, some indicators of success, such as grade point average (GPA) were not applicable outside of postsecondary settings. Given that some metacognitive interventions examined skills, such as reading comprehension (Raskind & Higgins, 1999) or problem-solving (Barton, 1988), while others examined performance in specific subjects, such as science or math, and that all the treatments were occurring in qualitatively different settings (e.g. elementary classrooms, summer school for secondary students, a college learning strategies course), we decided to expand our focus from academic achievement to performance upon academic tasks. Consequently, we decided to focus on articles that examined metacognition from a procedural knowledge point of view so that students with learning disabilities understand their own learning through the required task.

Branching techniques were also fruitful as we felt that our short-listed articles were the best example of studies we wished to include for aggregated meta-analysis. Consequently, we were able to locate other possible sources; for example, the article by Burchard and Swerdzewski (2009) entitled "Learning effectiveness of a strategic learning course," helped us identify five potential sources, one of which, Hutchinson (1993): "Effects of cognitive strategy instruction on algebra problem solving of adolescents with Learning Disabilities," met our inclusion and exclusion criteria for our final article database.

Moderator Variables

Nine moderator variables were identified to help analyze our results. The order is as follows: *Design Type (*True Experimental or Quasi-Experimental*), Topic (*subject areas included language arts, mathematics, general learning strategies), *Location of Study* (Geographical), *Level of Education*, *Gender*, *Ethnicity*, *Socio-Economic Status*, *Metacognition Measurement* (standardized or non-standardized*)*, and *Treatment Administration* (i.e. was treatment provided through the primary researcher, a trained other, an untrained other, or simply unspecified?). These variables were coded for each of our final six articles included Lederer (2000); Vaurus, Kinnunen, & Rauhanummi (1999); Welch and Jensen (1991); Berkeley, Mastropieri, & Scruggs (2011); Burchard & Swerdzewski (2009); and Hutchinson (1993).

Design Type

Based on Campbell and Stanley's (1963), *Experimental and quasi-experimental designs for research,* we discovered that four of the final six articles were Quasi-Experimental (non-random assignment of participants), including Lederer (2000); Vaurus, Kinnunen, & Rauhanummi (1999), Welch & Jensen (1991), and Burchard & Swerdzewski (2009), and two of the final six studies were True Experimental (random assignment of participants), including Berkeley, Mastropieri, & Scruggs (2011), and Hutchinson (1993). They were both classic examples of the Pretest-Posttest Control Group Design with random controlled Trials. We excluded this category because there were not enough articles in each item to properly produce significant findings.

Topic

For this item, we coded three items as different topics through the final six studies. There were two studies and four results that dealt with Language (coded as 1) as a focus. They were Welch & Jensen (1991) and Berkeley, Mastropieri, & Scruggs (2011). There was one study that dealt with mathematics (coded as a 2) Hutchinson (1993), and there were three studies with eight results that dealt with learning strategies (coded as 3), including Burchard & Swerdzewski (2009), Lederer (2000), and Vaurus, Kinnunen, & Rauhanummi (1999). The final six studies involved three different topics of areas of study. We decided to not remove this moderator variable because power amongst the group increased the scores of students with learning disabilities. The topic produced beneficial results, and it encouraged future research in diverse topics affiliated with our research question.

Location of Study (Geographical)

We excluded this category as a moderator variable due to a lack of cohesion. For the final six studies, five of them, or the majority of the studies fit into the classification of North America that we used to distinguish it from *Other* (inferring all other places other than North America). The one study that did not fit into the North American criteria was from Finland by Vaurus, Kinnunen, & Rauhanummi (1999).

Level of Education

We excluded this category as a moderator variable because it would demonstrate serious ramifications. We excluded level of education in the six studies because there were five different levels of education represented. The only way that we might have considered education is if we separated them into before college (4 studies) and after starting college (2 studies).

Gender

We excluded this category in our study as a moderator variable because of the difficulty associated with exact gender representation, and then to compare them with treatments for the number of participants. This category was coded into four categories including Boys Only, Girls Only, Both Genders, and Unspecified. This variable was excluded by the researchers due to the fact that two studies producing four statistical results were unspecified with gender, and the other four studies producing 12 results were both.

Ethnicity

We excluded this category in our study as a moderator variable as it was complex and coded into six categories. The six categories were Caucasian (1), Black (2), Hispanic (3), Asian (4), Mixed (5), and Unspecified (6). This variable was excluded because we could not decipher any significant results that would relate to ethnicity. Four of the studies were unspecified with regards to ethnicity. The two other studies were mixed participants, but it would be impossible for us to include serious results based on just two studies that would relate to this variable.

Socio-Economic Status

We excluded this category in our study as a moderator variable. This category included four coded areas for the six research studies. They included 1= Lower, 2= Middle, 3= Upper Socio-Economic Status, and 4= Unspecified. This variable was excluded, since four of the six research studies were classified as unspecific, one study involved lower socio-economic status participants, and one study involved middle socio-economic status participants. This suggests that like gender and ethnicity, it may be an area that researchers wish to explore more in the future if they can connect it to metacognition and students with learning disabilities with similar research designs and results.

Metacognitive Measurement

We included this category as a moderator variable in our study. The six research articles were coded in two ways. These included 1= standardize measure and 2= non-standardized measure. This category was a good variable in that all six of the final studies fit category 2 which means that all six studies and all sixteen results were non-standardized measures. Our final articles were not confined through the analysis by set designs. The reliance on non-standardized tests may suggest that the researchers prefer to be able to interpret their results in their own fashion. The fact that all six of our final list of studies utilized metacognitive measurements that were non-standardized demonstrates 100% similarity across all six studies.

Treatment Administration

We excluded this category as a moderator variable in our study. This involved four unique coded variables in our final Excel Spreadsheet for our six articles. They were coded as 1= Researcher (meaning the researcher conducted the treatment), 2= Trained Others (meaning the researcher(s) trained someone else to conduct the treatment such as a teacher), 3= Untrained Other (meaning the researcher(s) used someone else to conduct the treatment without any advice or coaching on how to administer the treatment), and 4= Unspecified (meaning that there is no indication in the research study of how the treatment was administered). Four of our studies involved the treatment by the researcher, and two of the studies involved the researcher training others to conduct the treatment. The Untrained Other and Unspecified category did not have any studies. It is difficult to find valid or common trends because of items such as bias, level of researcher involvement, and proper training with the administrators of two different treatments.

The three items that were included in our final analysis were: Design Type, Topic, and Metacognitive Measurement. These three items play important roles in assessing the dependent and independent variables. The six items that were excluded as variables were: Location of Study (Geographical), Level of Education, Gender, Ethnicity, Socio-Economic Status (SES), and Treatment Administration.

Meta-Analysis

For the complete results of our meta-analysis, refer to *Appendix B: Aggregate Meta-Analysis of Effect Sizes*. For the overall effect there was some variance between the Fixed Effect Size of 0.86 to the Random Effect Size of 1.42 with calculations of the data based on Number of Studies, Effect Size, Standard Deviation, Variance, Lower Limit, Upper Limit, Z-value, P-value (the previous two dealt with the Testing of the Null-Hypothesis- 2 Tail), Q-Value, Degrees of Freedom, P-Value, and I-squared through the Heterogeneity Section. This very high random effect size of 1.42 in comparison with the fixed effect size of 0.86 dictated that there were one or more studies that could be classified as an outlier. The variance between the Overall Fixed effects size (0.86) and Overall Random effects size (1.42) was very large (*diff.* = 0.56). Also, this analysis produced a P-value score of 0.00 which would suggest that our sample represents a different population. This is problematic because through our meta-analysis we are seeking to look for similarities between each of the six studies, and also within each study in that the treatment and control groups can be compared with each other. In other words, we are searching for reasons to believe that interventions are consistent across studies with aspects, such as age of participants, setting and all experimental conditions to represent the typical representation for this population.

The Vaurus, Kinnunen, & Rauhanummi (1999) study was the excluded outlier with the most prominent random effect size of 0.782 suggesting there may have been problems with the intervention or group. This outlier must be carefully considered as only six studies were included in the meta-analysis, and it was the only one to have an effect size result below 1. Extreme caution should be exercised when including an outlier with such a high effect size as it could misrepresent the study by skewing all of the distributions. The problem was that both the experiment and control groups had only 22 participants, and thus, was not a large enough sample to present valid findings in comparison with the other studies. Furthermore, the study showed a unique result through a variance of 0.045, and a Z-value of 0. However, this study is included because it is similar to the remaining five studies in that they all share a successful intervention on the experiment group indicating that a successful treatment was used to increase academic achievement for the participants with learning disabilities in the study.

Another study that raises some flags about the statistical data is Burchard and Swerdzewski (2009), since it produced the next highest effect size of 1.2 in comparison to Vaurus, Kinnunen and Rauhanummi (1999). The 1.2 effect size was closer to the remaining four groups, and it increased the validity through keeping six studies. To provide a glimpse of how tight the effect sizes were for the other four research articles by this measurement, they are 1.624, 1.594, 1.624, and 1.553, indicating that they were much closer values than 0.782 and 1.2 for the two other studies. We agreed with *Urdan* (2010) that effect size is a more valuable statistic than $p \leq .05$ because effect size is not too heavily influenced by the sample size.

There were essentially three statistical options available with finding the best representation with our final six articles. These were One study removed (which essentially removes one study and removes each effect one at a time to provide the results with each study removed), Outlier truncated to the next highest point, and Outlier truncated to the mean.

One study removed was excluded because it entailed only five studies, and a z-value average of 0.002. The averages for every study removed indicated that it would not be the best method to proceed with because the average data could not be further analyzed.

When we analyzed our statistical data through Outlier truncated to the next highest point we noticed that the fixed effect size was 0.79 and the random effect size was 1.24. When we did it through Outlier truncated to the mean (assigned values to the mean), the effect size for the fixed sample was 0.63, and for the random effect sample, it was 0.76. The method of analysis known as Outlier truncated to the next highest point and Outlier truncated to the mean are adjusted by assisting a value as the next highest, and mean effect size.

We placed greater emphasis on the fixed model as it is the best measure for different treatments for the sake of this study. The fixed model deals with one average that really exists in the population. Despite a variety of treatment constellations, there was a significant effect size at (d=0.79) with Outlier truncated to the next highest point for the metacognitive treatments for students with learning disabilities through random effect size. Our findings provide support that our research question has been successfully answered through this meta-analysis.

Both Outlier Truncated to the mean (d= 0.76), and Outlier Truncated to the highest point (d= 0.79) are valuable indicators of successful treatments. We sided with Outlier Truncated to the highest point because it provides a positive result in which the direction does not change. It does not change the mean while retaining the direction restricting variability.

Discussion

Interpretation of Meta-Analysis

According to Cohen's (1988) qualitative descriptors regarding the interpretation of percentile differences, we managed to find a medium to large effect size at 0.79 using random effects because of the constellation of different treatments. This provides reasonable evidence that treatments focusing on developing the metacognitive strategy use can effectively help students with learning disabilities enhance their performance on a variety of academic tasks.

Synthesizing the Findings of our Studies

As a whole, there are a number of important features regarding the main studies examined in this project. First, the studies varied considerably in terms of size. While some of the sample sizes had noticeably larger samples of participants (Lederer, 2000; Welch & Jensen, 1991), the studies conducted by Hutchinson (1993) and Vauras, Kinnunen and Rauhanummi (1999) used small samples of participants. While most of the groups had roughly equal control and treatment group sizes, there were noticeable between group differences for both Welch and Jensen (1991) and Burchard and Swerdzewski (2009).

Furthermore, the studies varied greatly in terms of how the treatment and control groups were created. While some randomly assigned individuals with LD to groups (Berkeley, Mastropieri, & Scruggs, 2011; Hutchinson, 1993), others chose to match their groups to ensure that their samples matched in terms of socioeconomic factors (Vauras, Kinnunen & Rauhanummi, 1999). Conversely, Lederer (2000) as well as Burchard and Swerdzewski (2009) used students both with and without learning disabilities within their treatment groups. Interestingly, Welch and Jensen (1991) had the least amount of control over their groups as they opted to compare students from two different sites.

All of the interventions were carried out in traditional classroom settings or within small group instruction. In addition, the consistency of treatments was similar in that all the students attended the interventions two to four times each week. However, intervention periods differed greatly; most persisted over a three to four month period, however, some lasted only four weeks (Berkeley, Mastropieri, & Scruggs, 2011; Lederer 2000;), while another provided intervention over a seven month period (Vauras, Kinnunen & Rauhanummi, 1999). These differences likely reflected the settings in which the interventions would occur and the researchers seemed to try providing intervention over as long a period of time as they could (e.g. several weeks of summer school, a fifteen-week postsecondary term, or an academic year spanning nine months).

One of the major concerns that we faced is the problem of standardization across all studies. For instance, when reviewing articles, we wanted to reduce article selection bias as much as possible; however, the quasi-experimental designs lacked random controlled trials so we were concerned that some groups were not equal (e.g. consider outcome comparisons between a regular high school classroom and one comprised uniquely of individuals with LD). Furthermore, treatments between studies were not completely identical. For example, while several studies examined metacognitive interventions involving general learning strategies (Burchard & Swerdzewski, 2009; Lederer, 2000; Vauras, Kinnunen, & Rauhanummi, 1999), others examined metacognitive interventions within the context of specific academic subjects including math (Hutchinson, 1993) and English (Raskind & Higgins, 1999; Welch & Jensen, 1991).

Despite a lack of standardized tests, researchers used a considerable variety of measures to assess metacognitive development. For instance, while some researchers opted for interviews (Hutchinson, 1993), others used questionnaires, such as the *Meta-Comprehension Strategy Index (MSI)* (Berkeley, Mastropieri, & Scruggs, 2011) or the *Metacognitive Awareness Inventory (MAI)* (Burchard & Swerdzewski, 2009). Interviews are possibly subject to researcher bias and, unfortunately, information regarding the interview process was often lacking (e.g. whether or not interviews were conducted by the primary researchers, inter-rater reliability, etc.). Conversely, the validity of self-report measures, such as the MSI or MAI is also questionable, especially since there exists at least partial evidence that students with learning disabilities overestimate their academic skills (Stone & May, 2002), as well as their performance on both academic and non-academic tasks (Job & Klassen, 2012). There were some concerns regarding construct validity and whether or not some of the instruments were stronger measures of metacognition than others. Moreover, there were slight variations in the types of constructs being measured (e.g. types of metacognitive questions being asked vs. quality of questions). As a result, we were conscientious that the studies may be examining slightly different aspects of metacognition, and we made a judgment to include aggregate meta-analysis statistical procedures, as well as to perform a descriptive analysis on each article.

The research literature is characterized by a lack of specificity regarding implementation descriptions of metacognitive treatments. While many studies identified the metacognitive strategies or instruction provided during treatments by providing brief descriptions, it was difficult to know exactly how such information was presented during treatments and this could be an issue of implementation fidelity that could also affect later attempts of replicating findings.

Another concern was instruction validity. In some of the studies (Hutchinson, 1993; Lederer 2000), the primary researcher conducted all the intervention sessions, and there is a possibility that the researchers' predispositions and attitudes influenced the outcomes of those interventions. In studies where researchers trained others, it was not clear whether there were noteworthy differences between instructors, particularly between the control and treatment conditions. The experience and level of involvement of the treatment trainers is difficult to assess for obvious reasons, including implications associated with the results. While Welch and Jensen (1991) addressed concerns of instruction validity by monitoring instructors through "periodic observation by the investigator and summer school supervisor and review of a daily log/journal" (p.44), other researchers did not provide such information, despite having interventions involving considerably more instructors (Berkeley, Mastropieri, & Scruggs, 2011), and great differences in instructor teaching experience (Lederer, 2000).

Finally, the results of our studies were often similar to the findings of other research that was not included in the meta-analysis statistical procedures of the current project. For instance, the benefits identified by Hutchinson (1993) of using think aloud strategies when solving algebra word problems supports other research that metacognitive training can lead to improved performance both in math (Lucangeli, Cornoldi & Tellarini, 1998), as well as other specific content courses such as science (Aykyol, Sungur, & Tekkaya, 2010).

Implications for Practice

The current findings have clear implications for practice. Although, the current project examined learners at various stages of their formal education, educators need to consider whether the strategies presented to different age groups of students are developmentally appropriate. For developing reflection skills regarding their work process, some students may benefit more from a structured reflection sheet with prompting questions and self-report criteria, while other students may prefer to use thought journals where they can describe learning strategies, revisit previous learning sessions to establish connections and monitor progress.

Although students with LD may require scaffolded instruction when acquiring new strategies; specifically, they may particularly benefit from systematic prompts and detailed format feedback. Consequently, given that students with LD may particularly benefit from reciprocal teaching, peer-tutoring, and small group learning, educators should carefully consider their instructional approach within the context of the learning setting (e.g. Are students attending a college-level learning strategies course, receiving support from a learning center, or attending individual support sessions?). Furthermore, students must be supplied with sufficient opportunities to practice, revisit, and reflect upon strategy use until these can be used more effectively and be applied more independently.

It is important to remember that metacognitive skills are not acquired automatically throughout the developmental process. Ssome students, particularly those with learning disabilities may struggle when attempting to learn metacognitive strategies. Moreover, they may particularly benefit from consistent exposure to such strategies across different subjects and learning settings. Consequently, the implementation of metacognition instruction needs to be carefully considered both inside and outside of the classroom by teachers, supporting educators, school administrators, and families alike.

In short, practitioners working with LD populations need to consider both the content of formal curriculum (i.e. the actual strategies that students will be learning), as well as the process of instruction (i.e. how the students learn, practice and use specific strategies) over time.

Recommendations for Future Research

Initial searches within the academic databases revealed a great degree of variability in terms of the quality of studies within the field of educational psychology research. Articles that initially seemed interesting often turned out to be more descriptive in nature by examining measures metacognition, achievement and learning disabilities without any formal treatments. Perhaps databases should include summative information including study design as identified by Campbell and Stanley (1963). Furthermore, specific information identifying independent and dependent variables, study questions and hypotheses, sample demographics, nature and duration of treatment conditions, data collection instruments, etc. would greatly facilitate and refine database searches.

Database searches also yielded a considerable number of relevant studies that suffered from weak pre-experimental designs that we ultimately chose not to include (Ellis, 1989). Other studies were rejected because they had more rather unsystematic reviews that seemed to present information to provide support for subjective perspective. There is a need for additional studies with more robust research designs. Given that much of the research compares the performance of LD and NLD students, random assignment between treatment groups and control groups making it difficult to obtain true experimental designs, as identified by Campbell and Stanley (1963). An example of a stronger research design includes two-group, pre-test/post-test designs that provide metacognitive treatments to both LD and NLD participants and can provide valuable gain scores or change scores of each group over time. Conversely, regression discontinuity designs can be used with a single group in order to measure the effectiveness of specific interventions by examining the experimental group both before and after treatment.

Our database searches also revealed a lack of description regarding learning disabilities. Learning disabilities can greatly vary in terms of area of impairment and severity. Were participants simply placed together? How were the students being assessed? Did the researchers consider how the learning difficulties of the students may have affected their ability to understand instructions, process information presented during treatments? Were students able to advocate their needs to the researchers? Were researchers able to provide any accommodations or supports to the participants (i.e. providing readers, adaptive technology, extra time, quiet individual work spaces, etc. are accommodations that are often provided to students during test-taking situations)?

With regards to the specific field of metacognitive research, the development of a standardized instrument to measure the psychometric constructs associated with the development and different types of metacognition would greatly benefit research within this area that could help future researchers within this field establish a best quality standard.

The studies were far less skilled at describing the LD aspects of their population. For instance, did participants with LD have mild, moderate or severe impairments? Were researchers considering verbal as well as nonverbal learning disabilities when examining interventions that target language skills or math ability? Were subtypes of learning disabilities considered such as attention deficit disorder or dyslexia? Given the diverse nature and different types of learning disabilities (Meltzer, 1991; APA, 2000; LDAC, 2002), these are important considerations given that some learning strategies may be more effective with some groups (Moore, Alvermann & Hinchman, 2000) and that practitioners need to carefully consider the structure of learning sessions (Nevin & Renne, 2001).

Given the considerable body of literature in similar areas of research, future studies should attempt to examine the link with diverse forms of metacognition and other cognitive frameworks involving students with learning disabilities such as executive functioning (Meltzer, 2007), metamemory (Geary, Klosterman, & Adrales, 1990), learned helplessness (Valas, 2001), attribution theory (Berkeley, Mastropieri, & Scruggs, 2011), self-efficacy (Magno & Lajorn, 2008), and self-regulation (Camahalan, 2006).

There are also a need for more research to consider forms of metacognition in order to assess specific topics such as low-income, ethnicity, rural versus urban, francophone, Aboriginal, and first-generation citizens amongst other topics. (Smith & Gottheil, 2011). Also, the transitions for students with learning disabilities from grade school to post-secondary education require more research, and establishing better record-keeping at all education levels through educational research by what ElAtia, Ipperciel, & Hammad (2012) call “data mining” (pp. 106-107).

Conclusion

The aggregate effect size result of 0.79 indicates that treatments focusing on metacognition have a medium-to-large impact upon the academic performance of students with learning disabilities. Despite the small number of studies included in our meta-analysis, this finding is important because there are clear implications regarding how we support students with different learning needs and that both content (e.g. formal curriculum, individual subjects, learning strategies, etc.) as well as process (i.e. how students learn and monitor their progress) are both important factors that can lead to greater student autonomy and school success. Additional studies with robust research designs are required, emphasizing how we implement, monitor, and assess practical applications of metacognition.

References

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: 1-507. http://clues.concordia.ca/record=b2724185~SO.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: 1-598. http://clues.concordia.ca/record=b2724184~SO.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: 1-915. http://clues.concordia.ca/record=b2721645~SO.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed. Text rev.). Washington, DC: 1-943 http://clues.concordia.ca/record=b2784649~SO.

Akyol, G., Sungur, S, & Tekkaya, C. (2010). The contribution of cognitive and metacognitive strategy use to students' science achievement. *Educational Research and Evaluation: An International Journal on Theory and Practice*, *16*(1), 1-21.

Barton, J. A. (1988). Problem-solving strategies in learning disabled and normal boys: Developmental and instructional effects. *Journal of Educational Psychology*, *80*(2), 184-191.

*Berkeley, S., Mastropieri, M. A., & Scruggs, T. E. (2011). Reading comprehension strategy instruction and attributional retraining for secondary students with learning and other mild disabilities. *Journal of Learning Disabilities*, *44*(1), 18-32.

Borkowski, J. G., & Cavanaugh, J. C. (1979). Maintenance and generalization of skills of skills and strategies by the retarded. In N. R. Ellis (Ed.), *Handbook of mental deficiency, psychological theory and research* (2nd ed., pp. 569-617). Hillsdale, NJ: Erlbaum.

*Burchard, M. S., & Swerdzewski, P. (2009). Learning effectiveness of a strategic learning course. *Journal of College Reading & Learning*, *40*(1), 14-34.

Camahalan, F. M. G. (2006). Effects of self-regulated learning on mathematical achievement of selected southeastern asian children. *Journal of Instructional Psychology*, *33*(3), 194-205.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi experimental designs for research.* Boston: Houghton Mifflin.

Chan, L. K. S. (1994). Relationship of motivation, strategic learning, and reading achievement in Grades 5, 7, and 9. *Journal of Experimental Education*, *62*(4), 319-339.

Chen, L.-J., Ho, R. G., & Yen, Y. C. (2010). Marking strategies in metacognition-evaluated-computer-based testing. *Educational Technology & Society*, *13*(1), 246-259.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Creswell, J. W. (2008*). Educational research: Planning, conducting, and evaluating quantitative and qualitative research. (*3rd ed.*).* Upper Saddle River, NJ: Pearson Prentice Hall.

Desoete, A., & Roeyers, H. (2006). Metacognitive macroevaluations in mathematical problem solving. *Learning and Instruction*, *16*(1), 12-25.

ElAtia, S., Ipperciel, D., & Hammad, A. (2012). Implications and challenges to using data mining in educational research in the canadian context. *Canadian Journal of Education*, *35*(2), 101-119.

Ellis, E. S. (1989). Teaching adolescents with learning disabilities to generate and use task-specific strategies. *Journal of Learning Disabilities*, *22*(2), 108-119.

Fichten, C. S., Barlie, M., Asuncion, J. V., Fossey, M. E., Robillard, C. et al. (2003). Canadian postsecondary students with learning disabilities: Where are they? *The Canadian Journal of Higher Education*, *33*(3), 71-113.

Flavell, J. H. (1976). *Solving spatial perspective-taking problems by rule vs .computation: A developmental study*. Minnesota University: Minneapolis Institute of Child Development, pp. 42.

Flavell, J. H., & Wellman, H. M. (1977). Metamemory. In R. B. Kail, Jr., & J. W. Hagens (Eds.), *Perspectives on the development of memory and cognition* (pp. 3-33). Hillsdale, NJ: Erlbaum.

Geary, D. D., Klosterman, I. H., & Adrales, K. (1990). Metamemory and academic achievement: Testing the validity of a group-administered metamemory battery. *Journal of Genetic Psychology*, *151*(4), 439-450.

Hall, C. W., & Webster, R. E. (2008). Metacognitive and affective factors of college students with and without learning disabilities. *Journal of Postsecondary Education and Disability*, *21*(1), 32-41.

Haywood, C. H. (2004). Thinking in, around, and about the curriculum: The role of cognitive education. *International Journal of Disability Development and Education*, *51*(3), 231-252.

*Hutchinson, N. L. (1993). Effects of cognitive strategy Instruction on algebra problem solving of adolescents with learning disabilities. *Learning Disability Quarterly*, *16*(1), 34-62.

Job, J. M., & Klassen, R. M. (2012). Predicting performance on academic and non-academic tasks: A comparison of adolescents with and without learning disabilities. *Contemporary Educational Psychology*, *37*(2), 162-169.

Learning Disabilities Association of Canada. (2002). Official definition of learning disabilities. Retrieved March 14, 2012 from: http://www.ldac-acta.ca/en/learn-more/ld-defined/official-definition-of-learning-disabilities.html.

Learning Disabilities Association of Canada. (2013). The Supreme Court of Canada makes unanimous milestone ruling on Canadians with learning disabilities. Retrieved June 26, 2013 from http://www.ldac-acta.ca/en/media-releases/91-the-supreme-court-of-canada-makes-unanimous-milestone-ruling.html.

*Lederer, J. M. (2000). Reciprocal teaching of social studies in inclusive educational classrooms. *Journal of Learning Disabilities*, *33*(1), 91-106.

Lucangeli, D., Cornoldi, C., & Tellarini, M. (1998). Metacognition and learning disabilities in mathematics.*Advances in Learning and Behavioral Disabilities*, *12*(1), 219-244.

Macionis, J. J., Jansson, S. M., & Benoit, C. M. (2012). *Society: The basics (Updated Fourth Canadian Edition).* Toronto: Pearson Canada.

Magno, C., & Lajorn, J. (2008). Self-regulation, self-efficacy, metacognition, and achievement goals of high school and college adolescents. *Philippine Journal of Psychology*, *41*(1-2), 1-23.

Meltzer, L. (1991). Problem-solving strategies and academic performance in learning-disabled students: Do subtypes really exist? In L. V. Feagans, E. J. Short, L. J. Meltzer (Eds.), *Subtypes of learning disabilities: Theoretical perspectives and research* (pp. 163-188). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Meltzer, L. (2007). Executive function in education: From theory to practice. *Guilford Publications in New York*, ED493739, pp. 308.

Moore, D. W., Alvermann, D. E., & Hinchman, K. A. (2000). Struggling adolescent readers: A collection of teaching strategies. I*nternational Reading Association, Newark, DE*, ED441223, pp. 344.

Nelson, T. O., Narens, L., & Dunlosky, J. (2004). A revised methodology for research on metamemory: Pre-judgment recall and monitoring (PRAM). *American Psychological Association*, *9*(1), 53-69.

Nevin, A., & Renne, R. (2001). Cooperative group learning: K-12 mathematics lesson plans. *Classroom Teacher Guide- ERIC- Resources in Education (52)*, ED455638, pp. 21.

Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist*, *36*, 89-101.

Raskind, M. H., & Higgins, E. L. (1999). Speaking to read: The effects of speech recognition technology on the reading and spelling performance of children with learning disabilities. *Annals of Dyslexia*, *49*(1), 251-281.

Schunk, D. H. (2012). *Learning theories: An educational perspective* (6th ed.). Pearson Education Inc. Boston, Massachusetts.

Schunk, D. H., & Rice, J. M. (1993). Strategy fading and progress feedback: Effects of self-efficacy and comprehension among students receiving remedial reading services. *Journal of Special Education*, *27*, 257-276.

Schunk, D. H., & Zimmerman, B. J. (Eds.) (1994). *Self-regulation of learning and performance: Issues and educational applications.* Hillsdale, NJ: Erlbaum.

Schunk, D. H., & Zimmerman, B. J. (Eds.) (1998). *Self-regulated learning: From teaching to self-reflective practice*. New York: Guilford Press.

Short, E. J. (1992). Cognitive, metacognitive, motivational, and affective differences among normally achieving, learning disabled, and developmentally handicapped students: How much do they affect school achievement? *Journal of Clinical ChildPsychology*, *21*(3), 229-239.

Slavin, R. E. (2002). Evidence-based educational policies: Transforming educational practice and research. *Educational Researcher*, *31*(7), 15-21.

Smith, C., & Gottheil, S. (2011). Increasing accessibility: Lessons learned in retaining special population students in Canada. *College and University*, *86*(4): 47-52.

Stacey, P. (2013). Government support for open educational resources: Policy, funding, and strategies. *The International Review of Research in Open and Distance Learning*, *14*(2), 67-80.

Stevens, R., & Shenker, L. (1991). A multidimensional treatment program for learning disabled adolescents: A preliminary report. *Canadian Journal of Special Education*, *7*(1): 60-66.

Stone, C. A., & May, L. A. (2002). The accuracy of self-evaluation in adolescents with learning disabilities. *Journal of Learning Disabilities*, *35*(4), 370-383.

Trainin, G., & Swanson, H. L. (2005). Cognition, metacognition, and achievement of college students with learning disabilities. *Learning Disability Quarterly*, *28*(4), 261-272.

Urdan, T. C. (2010). *Statistics in plain English* (3rd ed.). New York: Taylor & Francis Group.

Valas, H. (2001). Learned helplessness and psychological adjustment II: Effects of learning disabilities and low achievement. *Scandinavian Journal of Educational Research*, *45*(2), 101-114.

*Vauras, M., Kinnunen, R., & Rauhanummi, T. (1999). The role of metacognition in the context of integrated strategy intervention. *European Journal of Psychology of Education*, *14*(4), 555-569.

*Welch, M., & Jensen, J. B. (1991). Write, P.L.E.A.S.E.: A video-assisted strategic intervention to improve written expression of inefficient learners. *Remedial & Special Education*, *12*(1), 37-47.

Winzer, M., & Mazurek, K. (2011). Canadian teachers' associations and the inclusive movement for students with special needs. *Canadian Journal of Educational Administration and Policy*, *116*, 1-24.

* Note: References marked with an asterisk (*) indicate studies included in the meta-analysis.

Appendix A:

Information Chart on Meta-Analysis Studies

Study/ Design	Sample	Duration/ Treatment	Contrast	Measure	Concerns	Recommendations	Notes
Lederer (2000) Quasi-Experimental	128 students from 4th, 5th & 6th grade; 25 students have LD	The social studies classes over a four week period; 15-17 treatment days consisting of reciprocal teaching in small groups	Regular Social Studies curriculum set by the original teacher	Non standardized assessment: Four 45 minute assessments given each week; non-standardized assessment of question asking, comprehension questions and summaries	- No standardized measures. The scoring key is rather subjective. - Bias; experimenter taught all classes himself (only study) - Are groups truly representative given that they consist of LD and NLD students? - Lack of descriptive statistics of sample (Race & SES)	Why was a global comprehension assessment not administered in p/P? While many psychometric tests cannot be re-administered within such a short period of time, waiting a period of 6 -12 months could have been an excellent idea and given much more credibility to the effectiveness of intervention	How LD students were identified is not clarified. No discussion about the type of LDs.
Vauras, Kinnunen & Rauhanummi (1999) Quasi Experimental	44 students with learning problems in matched E/C Groups (22 Each)	Metacognitive intervention program consisted of 2 x 1-hour sessions over a 7 month period (38 total)	22 students in control group and 16 gifted students	Non standardized assessment: a series of 21 verbally and pictorially presented tasks	- Lack of descriptive statistics of sample (Gender, Race, SES) - A 'nothing' control treatment?	- Multiple comparison groups; Control (22 students w/ learning problems), 16 gifted students, and "other" category (N=130)	Are more recent studies are paying greater attention to stronger research: true experimental research designs? Can we correlate to 3 preexperimental designs in Database 3?
Welch & Jensen (1991) Quasi Experimental	114 students from two middle schools (E44/71 C)	P.L.E.A.S.E. intervention during the 1989 summer school session	71 summer school students at the second school site	Non standardized assessment: Evaluation of the P.L.E.A.S.E. Strategy	- Lack of descriptive statistics of sample (Gender) - Inefficient Learners is not LD	Not necessarily an LD group; the authors identify the students as inefficient learners who are "at risk"	Very well defined strategies in the PLEASE approach, however NO IMPLEMENTATION

Study/ Design	Sample	Duration/ Treatment	Contrast	Measure	Concerns	Recommendations	Notes
Berkeley, Mastropieri & Scruggs (2011) True Experiment	59 students with LD randomly assigned to 2 X E or C	360 minutes of instruction over a 4 week period (12 x 30 minute sessions); 20 minutes of primary instruction/10 minutes AR or read aloud	Read Naturally (RN) Control group & Attribution Retraining (AR) 2nd Treatment Group	Meta-Comprehension Strategy Index (MSI)	- Metacomprehension is not exactly metacognition; discuss in write up - "other mild disabilities"	-Link to Attribution Retraining	Well detailed process; discuss for implications regarding replication
Burchard & Swerdzewski (2009) Quasi Experimental	78 treatment participants (44 NLD; 34 LD); control of 1463	16-week strategic learning course (3 credit college course)	General population college students	Non standardized assessment: Metacognitive Awareness Inventory (MAI)	The control group receives a treatment of "nothing." Should they not have received some kind of support instead? - Lack of descriptive statistics of sample (Race & SES)	This is a MUCH description of strategies, setting, selection criteria, etc. Evidence of a trend in the right direction? Excellent link: metacognition, leaning strategies & success	All studies used non-standardized measures for metacognition; this used self-report (MAI) - LD being compared to NLD population (see table 1)
Hutchinson (1993) True Experiment	20 LD students, randomly assigned to E or C (12E/8C)	Individual 40 minute sessions on alternating days over a four month period	Other LD students attending resource course	Metacognitive Interview based on Flavell (1976)	- 24% of sample were low achievers (LD could still be a possibility) - Researcher is instructor; bias? - Interview bias?	True Experimental design with RCT	- Support for metacognitive math intervention - Used only LD students

Appendix B:

Aggregate Meta-Analysis of Effect Sizes

Overall effect

Model	Number S	Point estin	Standard e	Variance	Lower limi	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared
		Effect size and 95% confidence interval					Test of null (2-Tail)		Heterogeneity			
Fixed	6	0.86	0.09	0.01	0.69	1.03	9.88	0.00	102.01	5.00	0.00	95.10
Random effects	6	1.42	0.47	0.22	0.50	2.34	3.04	0.00				

One study removed

1.000	1.624	0.669	0.448	0.313	2.935	2.427	0.015
2.000	0.782	0.213	0.045	0.365	1.200	3.672	0.000
3.000	1.594	0.556	0.309	0.506	2.683	2.870	0.004
4.000	1.624	0.685	0.469	0.282	2.966	2.372	0.018
5.000	1.200	0.493	0.243	0.234	2.166	2.435	0.015
6.000	1.553	0.528	0.279	0.519	2.588	2.942	0.003
Average:	1.419	0.468	0.219	0.503	2.336	3.036	0.002

Model	Number S	Point estin	Standard e	Variance	Lower limi	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared
		Effect size and 95% confidence interval					Test of null (2-Tail)		Heterogeneity			
Fixed	5	0.61	0.09	0.01	0.44	0.79	6.76	0.00	12.89	4.00	0.01	68.96
Random effects	5	0.78	0.21	0.05	0.36	1.20	3.67	0.00				

Outlier truncated to the next highest point

Model	Number S	Point estin	Standard e	Variance	Lower limi	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared
		Effect size and 95% confidence interval					Test of null (2-Tail)		Heterogeneity			
Fixed	6.00	**0.79**	0.09	0.01	0.62	0.96	9.07	0.00	58.18	5.00	0.00	91.41
Random effects	6.00	1.24	0.36	0.13	0.54	1.95	3.47	0.00				

Outlier truncated to the mean

Model	Number S	Point estin	Standard e	Variance	Lower limi	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared
		Effect size and 95% confidence interval					Test of null (2-Tail)		Heterogeneity			
Fixed	6.00	0.63	0.09	0.01	0.46	0.80	7.21	0.00	13.16	5.00	0.02	62.00
Random effects	6.00	0.76	0.18	0.03	0.41	1.11	4.29	0.00				