# A Meta-analytic Review of Studies of the Effectiveness of Small-Group Learning Methods on Statistics Achievement

Sema A. Kalaian
Eastern Michigan University

Rafa M. Kasim
Statistical and Research Consultant

**Key Words**: Statistics Education; Active Learning; Cooperative Learning; Collaborative Learning; Meta-Analysis; Hierarchical Linear Modeling

## Abstract

This meta-analytic study focused on the quantitative integration and synthesis of the accumulated pedagogical research in undergraduate statistics education literature. These accumulated research studies compared the academic achievement of students who had been instructed using one of the various forms of small-group learning methods to those who had been instructed using lecture-based instruction. The meta-analytic results showed that cooperative, collaborative, and inquiry-based learning methods were used in college-level statistics courses. The results also showed that cooperative and collaborative learning methods supported the effectiveness of the small-group learning methods in improving students' academic achievement with an overall average effect-size of 0.60. In contrast, the effectiveness of inquiry-based learning was close to zero. This significant positive average effect-size indicated that using small-group learning methods in statistics classrooms could increase the achievement of college students, increasing the scores on a statistics exam from the 50th to the 73rd percentile. In addition, the multilevel analysis revealed that the effect sizes were influenced significantly by the publication-year of the studies, with the most recently published studies having lower effect sizes. The major implication of this study is that evidence-based research supports the effectiveness of active small-group learning methods in promoting students' achievement in statistics.

# 1. Introduction

Statistics has become an important academic subject across all levels of schooling from elementary school to college level and across a broad range of disciplines (American Statistical Association 2005; National Council of Teachers of Mathematics 2006). Because of its educational significance, most undergraduate majors across all scientific disciplines in the U.S. and abroad require statistics and research methods courses in their undergraduate programs of study. For example, the American Board of Engineering and Technology (ABET) recommended the inclusion of statistical learning and training in undergraduate education as one of the accreditation requirements for engineering and technology (Bryce 2005). Additionally, there have been increased calls from federal agencies (e.g., National Research Council 1996; National Science Foundation 1996; National Council for Teachers of Mathematics 2006) and professional organizations such as the American Statistical Association (2005, 2010) for pedagogical reforms in statistics education (Cobb 1993; Snee 1993). One of the key goals of the educational and pedagogical reform efforts has been the development of innovative student-centered pedagogies as alternatives to traditional lecture-based instruction.

The major aims of these alternative pedagogies are to help the students to: (a) develop effective statistical-reasoning and problem-solving skills; (b) develop reflective critical and higher-order thinking skills; (c) develop meta-cognitive skills; (d) develop effective communication, teamwork, and social skills; (e) retain the newly learned statistical concepts for subsequent and future applications; and (f) apply the learned statistical materials and concepts to new scientific problems and situations. In general, the necessity for students to be actively engaged in the statistics classrooms is grounded in the constructivist theory of learning. Constructivist theory views students as active learners engaged in constructing and restructuring their own newly learned concepts based on previously learned materials (Cooperstein and Kocevar-Weidinger 2004; Vygotsky 1978).

Garfield (1995) identified ten important principles for learning statistical material and concepts based on the context of constructivism theory, which focuses on the belief that learners in statistics courses continuously construct their own understanding of the statistical concepts. Generally, constructivism is guided by the following four general principles: (a) Learners construct their own meaning of the newly learned materials; (b) New conceptual learning builds on prior knowledge and real-life experiences; (c) Learning is enhanced by the social interactions between the learners in the active small-group learning environments; and (d) Learning develops through the performance of authentic tasks (Cooperstein and Kocevar-Weidinger 2004). The process of continuously constructing and restructuring newly learned statistical concepts to fit into students' existing cognitive framework helps the students to meaningfully and conceptually understand the newly learned statistical concepts and materials. Therefore, the main role of the instructor in the active small-group learning classrooms is to serve as a facilitator of learning rather than taking the "sage on the stage" role. The facilitator's main role is creating and structuring classroom environments that (a) provide engaging learning experiences, and (b) encourage cognitive conflict, critical thinking, creativity, meta-cognitive thinking, and self-directed collaborative learning.

To date, only a handful of non-quantitative reviews have been conducted about the status of teaching and learning in statistics education. These narrative reviews were conducted by Garfield (1993), Bryce (2002), Bryce (2005), Garfield and Ben-Zvi (2007), and Zieffler et al. (2008). These reviews provided practical recommendations for reforming and improving statistics education.  A key recommendation across these reviews was the use of various active small-group learning methods (e.g., cooperative, collaborative, inquiry-based, or problem-based learning methods) and activities to augment or replace the traditional lecture-based instruction (American Statistical Association 2005; Cobb 1993; Snee 1993) in statistics classrooms. Accordingly, based on these recommendations, the American Statistical Association (ASA) developed comprehensive guidelines for teaching and learning of statistics for K-12 education (American Statistical Association 2005) and college education (American Statistical Association 2010). The Guidelines for Assessment and Instruction in Statistics Education (GAISE) of the American Statistical Association (2010) recommended the following in relation to active small-group learning (recommendation #4):

> *"Using active learning methods in class is a valuable way to promote collaborative learning, allowing students to learn from each other. Active learning allows students to discover, construct, and understand important statistical ideas and to model statistical thinking. Activities have an added benefit in that they often engage students in learning and make the learning process fun. Other benefits of active learning methods are the practice students get communicating in the statistical language and learning to work in teams. Activities offer the teacher an informal method of assessing student learning and provide feedback to the instructor on how well students are learning. It is important that teachers not underestimate the ability of activities to teach the material or overestimate the value of lectures, which is why suggestions are provided for incorporating activities, even in large lecture classes."* (p. 18)

Based on these recommendations and guidelines, statistics educators have been actively seeking ways to develop, adopt, and implement innovative methods of learning and instruction as alternative pedagogies to the traditional lecture-based instruction. The different forms of active small-group learning methods such as cooperative learning, collaborative learning, problem-based learning, inquiry-based learning, peer-learning, and team learning are examples of such alternative small-group pedagogies that have been developed and implemented across various levels of schooling from elementary to college throughout the United States of America, South America, Australia, Europe, and in many other countries.

The increased use of various small-group methods in the United States and worldwide has been evidenced by the proliferation of research studies in Science, Technology, Engineering, and Mathematics (STEM) teaching/learning that evaluate the effectiveness of active small-group learning methods across all disciplines and fields of study including statistics. Similarly, there is a growing body of pedagogical research in statistics education that focuses on examining the effectiveness of various forms of active small-group learning methods, compared to lecture-based instruction in college statistics classrooms. These accumulated primary studies have been included in the present meta-analytic review. Zieffler, et al. (2008) noted that summarizing the entire body of accumulated literature that focuses on the teaching and learning of statistics is a

challenging and important endeavor. Tishkovskaya and Lancaster (2012) pointed to the need for continuous review of the teaching and learning strategies that are used in statistics classrooms.

Therefore, there is a need to integrate and synthesize the collection of existing statistics education research on small-group learning to better inform teachers and educators. Springer, Stanne, and Donovan's (1999) meta-analysis review appears to be the only known quantitative review that has focused on the effects of various forms of small-group learning methods at the college level across all STEM classrooms. The researchers included 37 STEM primary studies in their review with only one of these studies conducted in a college statistics classroom. To date, we are not aware of any meta-analytic study that quantitatively examined the effectiveness of different small-group learning methods in comparison to lecture-based instruction in college statistics classes.

In general, small-group learning methods are defined as being an umbrella for various forms of inductive and active student-centered instructional methods that empower the learners in small groups to work collaboratively and cooperatively with other members of the group in a team-based environment using effective communication and social skills (Cartney 2006; Fink 2004; Springer, et al. 1999). The following are definitions of the various forms of small-group learning methods that have been commonly used in statistics classrooms based on the present meta-analytic review:

- *Cooperative learning* is defined as a structured, systematic, and teacher-guided small-group instruction strategy in which students work together in small learning groups to maximize their own and each other's common learning goals (Johnson and Johnson 1989; Garfield 1993; Slavin 1995). The cooperative learning method is often guided by the following five major principles: (a) positive interdependence among the members of the group through adoption of different teacher-assigned roles that support the group's goal to complete a specific task, (b) peer social interactions in the classroom, (c) classroom activities structured by the teacher to establish individual accountability and personal responsibility, (d) development of interpersonal and small-group dynamic process skills, and (e) self-assessment of group functioning (Johnson and Johnson 2009; Johnson, Johnson, and Smith 1998; Johnson, Johnson, and Stanne 2000).

- *Collaborative learning*, in contrast to cooperative learning, is an unstructured form of small-group learning that incorporates a wide range of formal and informal instructional methods in which students interactively work together in small groups toward a common goal (Roseth, Garfield, and Ben-Zvi 2008; Springer, et al. 1999). Springer, et al. (1999) described the collaborative learning method as "relatively unstructured process through which participants negotiate goals, define problems, develop procedures, and produce socially constructed knowledge in small groups." (p. 24).

- *Inquiry-based learning* is a small-group instructional method for seeking information and knowledge in which students work in teams to solve a problem or inquiry through exploring, developing and asking relevant questions, investigating, making discoveries, presenting the results of the discoveries to other students in the classroom, and writing a scientific report (Chiappetta 1997; National Research Council 1996). In inquiry-based learning environments students actively explore and experience a specific scientific

problem and inquiry before they learn the vocabulary, concepts, and scientific content of the inquiry (Chiappetta 1997).

The present meta-analytic study focuses on the comparison of the achievement of undergraduate students who experienced various forms of the small-group learning methods in their college statistics classes to their counterparts who experienced the traditional lecture-based instruction. The main objectives of this meta-analysis are to determine the following: (a) how much research has been conducted on the use of various forms of small-group learning methods in comparison to lecture-based instruction in college statistics courses; (b) what different forms of small-group learning methods have been implemented and evaluated in college statistics courses, and (c) how effective is each form of the implemented and evaluated small-group method in maximizing student achievement in statistics classrooms.

## 2. Meta-Analysis Methodology

Meta-analysis is a quantitative statistical method for synthesizing and integrating the reported descriptive statistics from multiple relevant published and unpublished primary studies that address and test the same conceptual research question and hypothesis (Glass 1976; Hedges and Olkin 1985). Multilevel meta-analysis (mixed-effects or random-effects regression) methods (Kalaian and Kasim 2008; Raudenbush and Bryk 2002) were used in the present study to integrate and statistically model the accumulated quantitative literature on the effectiveness of various forms of active small-group learning methods in comparison to lecture-based instruction on students' achievement in undergraduate college statistics classrooms. The mixed-effects method for meta-analysis includes the fixed-effects estimates (e.g., regression coefficients of the predictors) and the random-effects estimates (e.g., standard deviations of the random errors of the between studies model).

In this study, we took five major steps to conduct the meta-analytic review. The first step consisted of the identification of the published primary studies that had focused on the effectiveness of various forms of small-group learning methods in comparison to traditional lecture-based instruction in undergraduate college statistics classrooms. We used extensive library search procedures to identify the published primary studies (articles published in peer-reviewed journals) and unpublished primary studies (dissertations, theses, conference proceedings, and reports). Library searches were conducted by searching the (1) electronic databases, such as Education Resources Information Center (ERIC), ProQuest Digital Dissertations and Theses, JSTOR, ABI, and PsycINFO; (2) web using standard search engines, such as Google and Google Scholar; (3) print and online electronic statistical journals such as "Journal of Statistics Education", "Teaching Statistics", "The American Statistician", and "Statistics Education Research Journal." Additionally, we examined the references of previously conducted (a) Meta-analytic reviews of Science, Technology, Engineering, and Mathematics (STEM) education literature; (b) narrative reviews of statistics studies; and (c) statistics primary studies to identify other potential relevant statistics primary studies that could be included in the present quantitative meta-analytic review. The keywords used to conduct the search included: "cooperative learning," "collaborative learning," "problem-based learning," small-group learning," "peer learning," "team-based learning" as the key learning pedagogies, which were combined with "statistics" or "research methods" subject matter descriptors. Also, these two

keywords were combined with sample descriptors including "college", "undergraduate students", and "college students".

In the second step of the methodology, we identified and introduced a set of five rigorous inclusion criteria to determine whether an identified statistics primary study was qualified to be included in the meta-analytic review. The primary studies were required to meet the following criteria:

1. Used a two-group experimental (randomized pre-post research design), quasi-experimental (pre-post research design), or comparative (post-only research design). The focus of the studies was on comparing the achievement of a group of students in undergraduate statistics courses who were instructed using one of the various forms of small-group learning methods to their counterparts who were instructed using traditional lecture-based instruction. Based on this criterion, four primary studies comparing small-group learning to individualized learning methods in statistics courses were excluded from this review. Primary studies with one group, pre-post research designs were also excluded.

2. Conducted in undergraduate college classrooms. Primary studies that have been conducted in graduate college classrooms including classrooms in medical education were excluded.

3. Used student academic achievement in a statistics course as the dependent (outcome) variable. Primary studies that measured students' attitudes towards statistics were excluded.

4. Utilized one of the forms of small-group learning methods in the college statistics classroom(s). Primary studies that focused on subject matter other than statistics (e.g., biology, chemistry, physics, technology, or mathematics) were excluded from the present review.

5. Reported at least one of the following quantitative statistics and/or statistical test values: (a) The necessary descriptive and sufficient statistics such as the means, variances, and sample sizes; (b) Effect sizes measuring the effectiveness of using one of the various forms of small-group methods in the statistics classrooms; or (c) t-test values, F-test values, or p-values to be able to calculate the effect sizes.

Within the framework of these established criteria, we were able to identify nine primary studies that were qualified to be included in the present meta-analytic review. The total number of students in the nine primary studies was 926. A total of 448 students were in statistics classrooms with small-group learning instruction and a total of 478 students were in statistics classrooms with lecture-based instruction. All the identified primary studies were articles published in peer-reviewed journals. We attempted to reduce publication bias by extensively searching for dissertations and theses to be included in this review. Unfortunately, we were not able to identify any dissertations and theses that focused on comparing one of the various forms of small-group learning methods to the lecture-based instruction in college statistics classrooms. We can only

assume that the relative "infancy" of statistics education may be the reason that only a few statistics education programs are in existence nationwide and abroad.

The third meta-analytic step focused on coding the features of the primary studies by developing and using a coding sheet constructed to achieve the following goals:

(a) Extract the summary statistics (e.g., means and standard deviations of the achievement scores) for each of the two comparison groups in this review (the group who were instructed using small-group learning methods and the groups who were taught using lecture-based instruction). The outcome (dependent) variable was the reported achievement scores of the exams, quizzes, and/or assignments in each of the primary studies. These coded summary statistics were used to calculate the standardized mean-difference effect sizes.

(b) Code the features of the primary studies (e.g., research design, publication year, sample characteristics, contextual study features, instructional duration, and achievement test type). These coded study characteristics were used as moderator variables (predictors) in the multilevel regression models to examine the effects and possible influences of these variables on the calculated standardized mean-difference effect sizes.

The first author and a graduate student independently coded the characteristics of each primary study. Discrepancies between the assigned codes were resolved after reviewing and discussing the studies to reach consensus. Inter-rater reliability, which is a measure of the percentages of the coding agreements between the two raters on the coded variables before resolution of coding differences, was calculated for each of the coded variables. Inter-rater percent agreement ranged from 78% to 100%. Twenty-one of the 26 coded variables (e.g., publication year, publication type, institution type, course period, use of computers, classroom setting, location, method of placing students in small groups) had inter-rater agreement of 100%. Five coded variables had inter-coder percent agreement between 67% and 89%. For example, coding instructional duration had 67% inter-rater percent agreement; coding the majority of the students in the statistics classroom as being first-year college students or not had 78% agreement; coding the number of students assigned to the small groups in statistics classrooms had 89% agreement.

The fourth step of the meta-analysis process consisted of estimating and calculating a standardized mean-difference index (effect-size) and its variance using the reported summary statistics from each of the collected primary studies.  To ensure the independence of the effect sizes for each primary study, the following procedures were followed. First, if a primary study reported summary statistics for multiple independent samples of students (e.g., college grade levels) then multiple effect sizes were calculated from such a study. For example, the primary study conducted by Keeler & Steinhorst (1994) reported summary descriptive results of two independent samples of college students.  Therefore, these samples and their effect sizes were kept distinct for the statistical analyses in this review (see Table 1 in the results section). Second, if multiple standardized mean-difference effect sizes were calculated based on the summary descriptive statistics of the multiple subcategories of the outcome measure (e.g., assignment grades, mid-term exam, final exam) for the same students per statistics primary study, then the dependent effect sizes from such multiple dependent measures were averaged and a single effect size was reported and included in this review.  The primary studies conducted by Borresen

(1990), Giraud (1997), and Potthast (1999) are examples of such cases included in the present review. As a result, ten independent effect sizes were estimated and calculated from the nine primary studies. For each of the nine primary studies, the standardized mean-difference index (effect-size) was calculated by taking the difference between the means of the achievement scores of the students who had been instructed by one of the various forms of small-group methods and those who had been instructed using lecture-based instruction and then dividing the difference between the means by the two groups' pooled standard deviations (Hedges and Olkin 1985; Kalaian and Kasim 2008). A positive effect-size indicates a favorable achievement outcome for the students who had been instructed using one of the various forms of small-group learning methods. The average effect-size was also determined by weighting each effect-size by the inverse of its variance. This weighting procedure allows primary studies with larger samples to have a greater representation in the overall weighted average effect-size than studies with smaller sample sizes.

In the fifth and final step of the review, we applied multilevel modeling methods for meta-analysis using Hierarchical Linear Modeling (HLM*6*) software (Raudenbush, Bryk, Cheong, and Congdon 2004) to analyze the present meta-analytic data. The analysis was conducted in two separate stages. In the first stage, the unconditional multilevel model is used (the coded predictors were not included in the multilevel model), the calculated effect sizes from all the identified primary studies were integrated to estimate and test (a) the overall weighted average effect-size (fixed-effects) and (b) the variance of the effect sizes (random-effects). In general, this unconditional model helps the meta-analyst to estimate and examine the heterogeneity in the effect sizes across the primary studies in order to assess the need for modeling the heterogeneity in the subsequent second stage conditional between-studies models. In the second stage (conditional multilevel model), the coded predictor variables were included in the between-studies model and used to model the variations among the true effect sizes from the first stage as a function of study characteristics and random errors. Generally, this conditional model helps the meta-analyst to explain some of the variations among the effect sizes (Kalaian and Kasim 2008; Raudenbush and Bryk 2002).

## 3. Results

The results of the present review show that the actual number of primary studies reporting the effectiveness of various forms of small-group learning methods that utilized two-group research designs in undergraduate college level statistics classrooms was limited in the statistics education literature. Because of the limited number of primary studies in the review, it was not possible to conduct some important multilevel analyses that could have provided further clarification of the relative strength of the relationships and interactions between important study characteristics and student achievement in college statistics classrooms (e.g., the inclusion of more than one predictor variable at a time in the multilevel conditional models).

Our results are organized into two sections. Section 3.1 describes the characteristics of the primary studies in the review. Section 3.2 reports the results of the multilevel analyses, which were applied to the calculated effect sizes and the coded characteristics of the statistics primary studies.

## 3.1  Description of the Statistics Primary Studies

Table 1 lists the 10 effect sizes extracted from the nine statistics primary studies included in the present meta-analytic review. Table 1 also lists the sample sizes of the two groups of students (the group who has been instructed using one of the various forms of small-group learning methods and the group that has been instructed with traditional lecture-based instruction). In addition, the duration of statistics instruction in hours is listed in Table 1.

The nine primary studies included in this meta-analysis (see Table 1) were published between 1990 and 2009. The 10 independent effect sizes ranged from -0.07 to 0.82. Out of the 10 effect sizes, eight were positive and favored small-group learning, while only two effect sizes were negative and close to zero (-0.07, -0.02).

**Table 1.** Characteristics and Effect Sizes of the Primary Studies in the Meta-Analytic Review

| Author (Publication Year) | Small Group (n) | Lecture Group (n) | Duration (hours) | Effect-size |
|---|---|---|---|---|
| **Borresen (1990)** | 100 | 100 | 45 | 0.71 |
| **Giles, et al. (2006)** | 61 | 82 | 3.00 | -0.07 |
| **Giraud (1997)** | 44 | 48 | 37.5 | 0.58 |
| **Keeler & Steinhorst (1994)** | 41 | 51 | 45 | 0.66 |
| **Keeler & Steinhorst (1994)** | 35 | 51 | 45 | 0.82 |
| **Kvam (2000)** | 15 | 23 | 30 | -0.02 |
| **Metz (2008)** | 16 | 44 | 9.96 | 0.29 |
| **Pariseau & Kezim (2007)** | 68 | 43 | 56 | 0.30 |
| **Potthast (1999)** | 30 | 21 | 37.5 | 0.65 |
| **Ragasa (2008)** | 38 | 15 | 12 | 0.81 |

Figure 1 displays the funnel plot of the scatter points of the 10 effect sizes for the nine primary studies along with their total sample sizes to assess publication bias of this review. In the absence of publication bias the funnel plot typically has a symmetrical inverted funnel shape with the vertical line drawn through the average effect size and the points distributed on either side of the line that creates an inverted funnel shape (Begg 1994). For illustrative purposes, a typical funnel plot is also displayed in Figure 1. As is shown in Figure 1, many more primary studies with significant positive effects favoring small-group learning had been published in the statistics literature than studies with negative effects. A visual examination of the funnel plot of the data

points indicates the possibility of publication bias because of the asymmetry of the scatter points around the mean of the effect sizes (in our case 0.47). We caution the reader to draw any conclusions about publication bias because of the small number of the primary studies in order to give a clear and valid resemblance to a typical funnel plot.
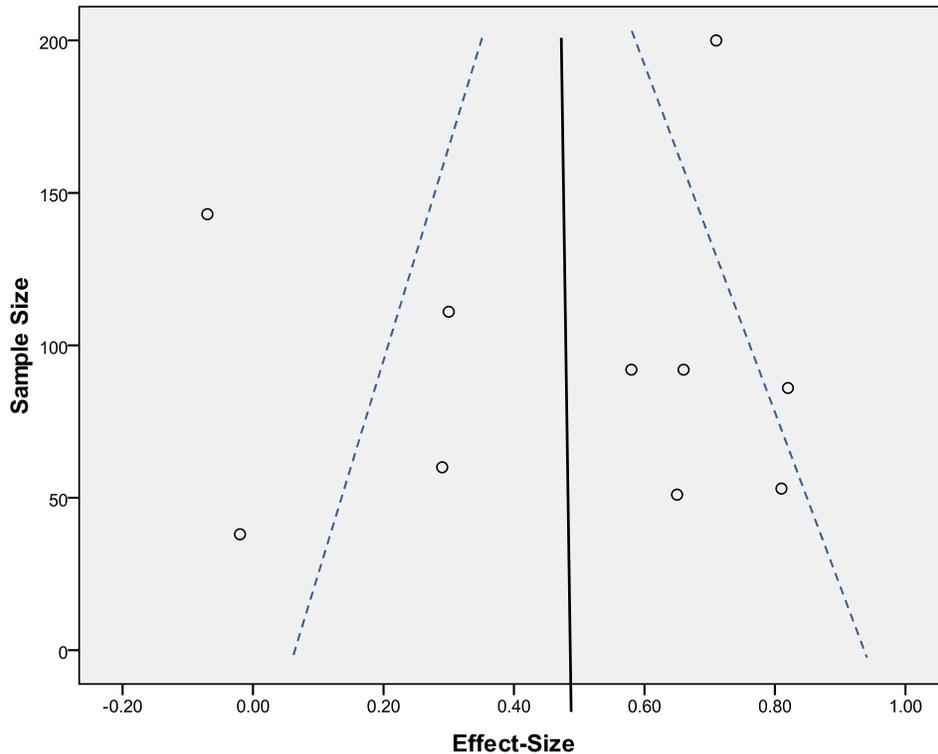


**Figure1**. Funnel Plot of the Statistics Achievement Effect sizes Versus Sample Sizes. Dashed line represents a typical funnel plot.

Table 2 summarizes the major characteristics of the nine primary studies. Four (44%) of the primary studies were published in 2001 or later with a weighted average effect size of 0.20. The remaining five studies (56%) were published in 2000 or earlier with much larger weighted average effect-size of 0.64.

A review of Table 2, shows that far more primary studies (6 studies, 67%) were non-experimental comparative studies (two-group post-only design) with a much higher weighted average effect-size (0.58) than the combined one randomized experimental study and the two quasi-experimental studies with a weighted average effect-size of 0.16 (the one experimental study is combined with the two quasi-experimental studies into one category). The lack of experimental and quasi-experimental studies is something expected in educational research because of the difficulty of conducting randomized experimental research in college settings.

Cooperative learning methods had been implemented in four primary studies (44%) with a weighted average effect-size of 0.60. Three studies (33%) used collaborative learning methods in

the statistics classrooms with a weighted average effect-size of 0.59. Two studies (22%) implemented inquiry-based learning methods with a weighted average effect-size close to zero (0.02). An average effect-size of zero indicates that there are no differences between inquiry-based learning methods and the traditional lecture-based instruction in promoting higher achievement scores.

One-third of the primary studies (3 studies) had much shorter instructional duration (12 hours or less) with a weighted average effect-size of 0.16. The other two-thirds of the studies (6 studies) had instructional durations of 30 hours or more with a much higher weighted average effect-size of 0.58. None of the primary studies in this review had instructional duration between 13 and 29 hours.

Only three (33%) primary studies used computers and statistical software in the classrooms with a weighted average effect-size of 0.40. The other six (67%) studies did not use computers and statistical software for instruction with a weighted average effect-size of 0.49.

In eight of the nine primary studies (89%) the investigators of the primary studies were also the instructors of the statistics classrooms with a weighted average effect-size of 0.58. In only one study (11%) the investigator of the primary study was not the instructor of the statistics classroom with a weighted average effect-size of -0.07.

The majority of the studies reported that the placement of students into small groups was accomplished through self-selection (6 studies, 67%) with a weighted average effect-size of 0.47. Two primary studies (22%) randomly placed students into small groups with a weighted average effect-size of 0.60. In only one primary study the placement of students into small groups was done by assigning students into groups based on their abilities (e.g., Grade Point Average scores) with an effect-size of -0.02. In seven primary studies (78%) the size of the small groups was four students or less.

Seventy-eight percent of the statistics primary studies were conducted at universities and colleges in the United States. The remaining 22% of the studies (two studies) were conducted in countries outside the U.S. (one primary study in each of Canada and Philippines). The results show that the primary studies conducted in the U.S. had much higher weighted average effect-size (0.56) than the studies conducted outside the U.S. (weighted average effect-size of 0.13).

**Table 2.** A summary of statistics study's characteristics and average effect sizes

| Study Characteristics | # Effect Sizes (%) | # Studies (%) | Average Effect-Size | Weighted Average Effect-Size |
|---|---|---|---|---|
| **Publication Year** | | | | |
| 2000 or Earlier | 6 (60) | 5 (55.6) | 0.57 | 0.64 |
| 2001 or Later | 4 (40) | 4 (44.4) | 0.33 | 0.20 |
| **Research Design** | | | | |
| Non-Experimental | 7 (70) | 6 (66.7) | 0.53 | 0.58 |
| Experimental & Quasi-Experimental | 3 (30) | 3 (33.3) | 0.34 | 0.16 |
| **Small-Group Learning Method** | | | | |
| Cooperative | 5 (50) | 4 (44.4) | 0.54 | 0.60 |
| Collaborative | 3 (30) | 3 (33.3) | 0.60 | 0.59 |
| Inquiry-Based | 2 (20) | 2 (22.2) | 0.11 | 0.02 |
| **Duration of Instruction** | | | | |
| 12 Hours or Less | 3 (30) | 3 (33.3) | 0.34 | 0.16 |
| 13 – 29 Hours | 0 (0) | 0 (0) | 0.00 | 0.00 |
| 30 Hours or More | 7 (70) | 6 (66.7) | 0.53 | 0.58 |
| **Use of Computers in Instruction** | | | | |
| Yes | 3 (30) | 3 (33.3) | 0.47 | 0.40 |
| No | 7 (70) | 6 (66.7) | 0.47 | 0.49 |
| **Classroom Instructor** | | | | |
| Investigator | 9 (90) | 8 (88.9) | 0.53 | 0.58 |
| Other | 1 (10) | 1 (11.1) | -0.07 | -0.07 |
| **Type of Statistics Exam** | | | | |
| Standardized Test | 0 (0) | 0 (0) | 0.00 | 0.00 |
| Teacher Made Test | 10(100) | 9 (100) | 0.47 | 0.47 |
| **Placement into Small Groups** | | | | |
| Random | 2 (20) | 2 (22.2) | 0.61 | 0.60 |
| Ability Grouping | 1 (10) | 1 (11.1) | -0.02 | -0.02 |
| Self-Selected | 7(70) | 6 (66.7) | 0.50 | 0.47 |
| **Small-Group Size** | | | | |
| 3 Students or Less | 5 (50) | 4 (44.40) | 0.57 | 0.44 |
| 4 Students | 3 (30) | 3 (33.30) | 0.43 | 0.52 |
| More than 4 | 2 (20) | 2 (22.20) | 0.28 | 0.40 |
| **Class Level** | | | | |
| First Year | 3 (30) | 2 (20) | 0.59 | 0.64 |
| Not First Year | 7 (70) | 7 (80) | 0.42 | 0.42 |
| **Study Location** | | | | |
| U.S. | 8 (80) | 7 (77.80) | 0.50 | 0.56 |
| Outside the U.S. | 2 (20) | 2 (22.20) | 0.37 | 0.13 |

## 3.2 Results of the Multilevel Analyses

As indicated earlier in Section 2, multilevel modeling for meta-analysis is typically conducted in two stages (unconditional and conditional multilevel modeling). The results are obtained by fitting the multilevel model (Kalaian and Kasim 2008; Raudenbush and Bryk 2002) via hierarchical linear modeling using HLM*6* software (Raudenbush, et al. 2004). Table 3 lists the results of fitting the unconditional (no predictors in the multilevel model) and simple conditional models (one predictor in each of the multilevel models) to the meta-analytic data. Simple conditional modeling was used in this study because of the limited number of primary studies.

**Table 3.** Unconditional and Conditional Multilevel Regression Analysis Results

| Models & Parameters | Fixed-Effects | | | Random-Effects | | |
|---|---|---|---|---|---|---|
| | Coefficient | t | p-value | S.D. | Chi-Square | p-value |
| **Unconditional Model (No Predictors in the Model)** | | | | | | |
| - Intercept | 0.47 | 4.41 | 0.002 | 0.25 | 21.10 | 0.01 |
| **Simple Conditional Models (One Predictor in each Model)** | | | | | | |
| - Use of Computers | 0.05 | 0.20 | 0.84 | 0.27 | 20.87 | 0.01 |
| - Instructional Duration | 0.01 | 1.79 | 0.11 | 0.18 | 13.43 | 0.10 |
| - Research Design | -0.33 | -1.60 | 0.15 | 0.18 | 14.01 | 0.08 |
| - Study Location | 0.31 | 1.31 | 0.23 | 0.20 | 14.97 | 0.06 |
| - Publication Year | -0.40 | -2.43 | 0.04 | 0.12 | 11.62 | 0.17 |

The results (see Table 3) show that the overall weighted average of the 10 effect sizes from the nine statistics primary studies is 0.47 across the three small-group learning methods (cooperative, collaborative, and inquiry-based learning methods). The average effect size of 0.47 may indicate that student achievement in a statistics course is increased from the 50th percentile to the 69th percentile for those instructed using cooperative, collaborative, or inquiry small-group learning methods. In other words, with achievement scores on a scale of 0 to 100, students' achievement scores in statistics may be improved by 19 points if cooperative, collaborative, or inquiry-based learning methods are used in the college statistics classrooms compared to achievement scores of 50 for their counterparts who were instructed using the lecture-based approach.

The random-effects test of the unconditional multilevel results also show that these effect sizes are heterogonous (Standard Deviation = 0.25, Chi-square = 21.10, p = 0.01). A significant Chi-square value indicates the existence of a significant amount of variability in the 10 effect sizes. Thus, there is a need to proceed further by fitting the simple conditional multilevel model using the coded variables from the primary studies as predictors to explain some of the variability in the effect sizes.

The results of fitting the simple conditional multilevel model with the publication year of the primary studies as a predictor show that the publication year has a significant negative effect on the achievement effect sizes (beta coefficient = -0.40, t = -2.43, p = 0.04). This significant negative effect of the publication year indicates that the most recently conducted and published primary studies have lower effect-size values than the earlier published studies. The results also show that this exploratory predictor explained most of the variations in the effect sizes, which is evident by the non-significant random-effect test (Standard Deviation = 0.12, Chi-square = 11.62, p = 0.17). Each of the remaining predictor variables were individually tested using the simple conditional multilevel model (e.g., study's research design, instructional duration, study's location, use of computers in the classrooms) and were not statistically significant. This might be due to the limited number of primary studies in this review, which leads to low statistical power of the statistical tests.

## 4. Conclusion

The results of the present meta-analytic review indicate that during the last two decades, three different methods of small-group learning (cooperative learning, collaborative learning, and inquiry-based learning) have been implemented and evaluated in statistics classrooms. When compared with lecture-based instruction, only two of the three small-group learning methods (cooperative and collaborative learning methods) had significant positive impacts on student achievement in college statistics courses with a weighted average effect-size of 0.60. This average effect size of 0.60 means that the statistics achievement can be improved from the 50th percentile for the students who had been taught by the traditional lecture-based instruction to the 73rd percentile for the students who had been instructed using cooperative and collaborative small-group learning methods. In other words, with achievement scores on a scale of 0 to 100, students' achievement scores in statistics could improve by up to 23 points if cooperative or collaborative learning methods are used in the college statistics classrooms in comparison to their counterparts (students instructed using lecture-based approach) with achievement scores of 50. The findings of this study are consistent with and confirm previously reported and published meta-analytic findings in STEM education, which support the effectiveness of small-group learning methods in increasing students' achievement in STEM classrooms at all levels of education (e.g., Johnson, et al. 2000; Springer, et al. 1999).

The results of fitting the conditional multilevel model with the publication year as a predictor show that the publication year of the primary studies has a significant negative effect on the effect sizes. This significant negative average effect-size indicates that the most recently conducted and published primary studies have lower effect-size values than the earlier published studies. In other words, the negative slope (beta coefficient) for publication year indicates that the impacts of small-group learning methods relative to lecture-based instruction are decreasing

over time (two decades period of the publication of the primary studies) in the statistics education literature. One possible explanation for this decline might be that the active learning methods (e.g., classroom discussions, use of technology) incorporating GAISE guidelines (American Statistical Association 2005) were fostered aggressively in addition to lecturing in statistics classrooms during the last decade in contrast to the preferred lecture-centered modality of the earlier decade. The multilevel results also show that this exploratory predictor explained most of the variations in the effect sizes.

The conditional multilevel results also show that instructional duration, research designs of the primary studies, and whether the primary studies were conducted in the United States are not statistically significant. Results indicate that duration of the instruction has positive effects on the effect sizes and the longer the instructional duration, the larger the achievement gaps between the two groups (small-group learning and the lecture-based groups). Results show that non-experimental studies with two-group posttest-only research designs have larger effect sizes than the experimental and quasi-experimental studies. In addition, the primary studies that were conducted in the United States have much higher effect sizes than the studies that were conducted in other counties. One possible explanation for these non-significant results is the small number of the primary studies, which reduces the power of the statistical tests.

As in any quantitative meta-analytic review, the present review was limited by the state of the available statistics education literature. We found that the actual number of primary studies reporting the effectiveness of various forms of small-group learning methods that utilized two-group research designs in undergraduate college level statistics classrooms was limited. Because of the limited number of studies in the present review, it was not possible to conduct some important multilevel analyses that could have provided further clarification of the relative strength of relationships and interactions between important study characteristics and student achievements in college statistics classrooms (e.g., the inclusion of more than one predictor variable in the multilevel conditional model).

We hope that as more future primary studies in statistics education are conducted, future reviews will provide enhanced explanations adding to the results of the present meta-analytic study. In addition, there is a need for better reporting of small-group instructional processes, activities, and the research results of statistics education including the effectiveness of small-group learning research. For example, we had to exclude a handful of primary studies that did not report either the standard deviations of the two groups and/or the sample sizes, which are necessary descriptive summary statistics for calculating effect sizes.

In conclusion, this meta-analytic study has shed some light on the accumulated literature related to the effectiveness of various forms of small-group learning methods in college statistics classes. We found that the academic achievement of students enrolled in college statistics can be accelerated if they are placed in an environment in which they can actively connect statistics instruction to previously learned mathematical and statistical concepts and materials through collaborative and cooperative scientific inquiry. Based on the findings of this study, we recommend that (a) statistics teachers and educators should consider employing one of the various forms of small-group pedagogies and methods (e.g., cooperative and collaborative) that have been shown to be effective in improving student achievement in statistics, and (b) authors

and publishers of statistics textbooks develop well-designed group activities and statistical materials to be used by statistics teachers to successfully implement small-group learning pedagogies in statistics classrooms. Thus, the findings of this quantitative review have significant policy implications in undergraduate statistics education (e.g., implementing small-group methods across all classrooms in the academic institutions) and it is of great interest for statistics and research methods educators.

## Acknowledgements

## References

American Statistical Association (2005), "*Guidelines for Assessment and Instruction in Statistics Education*," [Online]. Available at http://www.amstat.org/education/gaise/index.cfm.

American Statistical Association (2010), "*GAISE College Report,*" [Online]. Available at http://www.amstat.org/education/gaise/GAISECollege.htm.

Begg, C. B. (1994), "Publication Bias," in H. Cooper and L. V. Hedges (Eds.), *The Handbook of Research Synthesis* (pp. 399-409). NY: Russell Sage Foundation.

Borresen, R. C. (1990), "Success in Introductory Statistics with Small Groups," *College Teaching,* 38(1), 26-29.

Bryce, G. R. (2002), "Undergraduate Statistics Education: An Introduction and Review of Selected Literature," *Journal of Statistics Education* [Online], 10 (2). Available at www.amstat.org/publications/jse/v10n2/bryce.html.

Bryce, G. R. (2005), "Developing Tomorrow's Statisticians," *Journal of Statistics Education* [Online], 13, 1. Available at http://www.amstat.org/publications/jse/v13n1/bryce.html.

Cartney, P. (2006), "The Emotional Impact of Learning in Small Groups: Highlighting the Impact on Student Progression and Retention," *Teaching in Higher Education*, 11 (1), 77-91.

Chiappetta, E. L. (1997), "Inquiry-Based Science. Strategies and Techniques for Encouraging Inquiry in the Classroom," *The Science Teacher*, 64, 22-26.

Cobb, G. (1993), "Reconsidering Statistics Education: A National Science Foundation Conference," *Journal of Statistics Education* [Online], 1, 1. Available at http://www.amstat.org/publications/jse/v1n1/cobb.html.

Cooper, J., and Robinson, P. (1998), "Small Group Instruction in Science, Mathematics, Engineering, and Technology: A Discipline Status Report and Teaching Agenda for the Future," *Journal of College Science Teaching*, 27(6), 383-388.

Cooperstein, S., and Kocevar-Weidinger, E. (2004), "Beyond Active Learning: a Constructivist Approach to Learning," *Reference Services Review,* 32(2), 141-148.

Fink, L. D. (2004), "Beyond Small Groups: Harnessing the Extraordinary Power of Learning Teams," in L. Michaelsen, A. Knight, and D. Fink (Eds.), *Team-Based Learning: A Transformative Use of Small Groups* (pp. 1-36).VA: Stylus Publishing.

Garfield, J. (1993), "Teaching Statistics Using Small-Group Cooperative Learning," *Journal of Statistics Education* [Online], 1, 1, Available at http://www.amstat.org/publications/jse/v1n1/garfield.html.

Garfield, J. (1995), "How Students Learn Statistics," *International Statistical Review,* 63(1), 25-34.

Garfield, J., and Ben-Zvi, D. (2007), "How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics," *International Statistical Review*, 75, 3, 372-396.

Giles, J., Ryan. D. A. J., Belliveau, G., Freitas, E. D., and Casey, R. (2006), "Teaching Style and Learning in a Quantitative Classroom," *Active Learning in Higher Education,* 7(3), 213-225.

Giraud, G. (1997), "Cooperative Learning and Statistics Instruction," *Journal of Statistics Education* [Online], 5, 3. Available at http://www.amstat.org/publications/jse/v5n3/giraud.html.

Glass, G. V. (1976), "Primary, Secondary, and Meta-analysis," *Educational Researcher*, 5, 3-8.

Hedges, L. V., and Olkin, I. (1985), *Statistical Methods for Meta-analysis*. Orlando, FL: Academic Press.

Johnson, D. W., and Johnson, R. T. (1989), *Cooperation and Competition: Theory and Research.* MN: Interaction Book Company.

Johnson, D. W., and Johnson, R. T. (2009), "An Educational Psychology Success Story: Social Interdependence Theory and Cooperative Learning," *Educational Researcher*, 38, 5, 365–379.

Johnson, D., R., Johnson, and Smith, K. (1998), "Cooperative Learning Returns to College: What Evidence is There That it Works?" *Change*, 30, 4, 26–35.

Johnson, D. W., Johnson, R. T., and Stanne, M. B. (2000), "Cooperative Learning Methods: A Meta-Analysis," [Online]. Available at http://www.tablelearning.com/uploads/File/EXHIBIT-B.pdf.

Kalaian, S. A., and Kasim, R. M. (2008), "Applications of Multilevel Models for Meta-Analysis," in A. O'Connell and B. McCoach (Eds.), *Multilevel Analysis of Educational Data* (pp. 315-343). USA: Information Age Publishing, Inc.

Keeler, C. M., and Steinhorst, R. K. (1994), "Cooperative Learning in Statistics," *Teaching Statistics,* 16, 3, 81-84.

Kvam, P. H. (2000), "The Effect of Active Learning Methods on Student Retention in Engineering Statistics,*" The American Statistician,* 54, 2, 136-140.

Metz, A. M. (2008), "Teaching Statistics in Biology: Using Inquiry-Based Learning to Strengthen Understanding of Statistical Analysis in Biology Laboratory Courses," *CBE-Life Sciences Education*, 7, 317-326.

National Council for Teachers of Mathematics (2006), *Curriculum Focal Points for Prekindergarten through Grade 8 Mathematics: A Quest for Coherence,* Reston, VA.

National Research Council (1996), *From Analysis to Action: Undergraduate Education in Science, Mathematics, Engineering, and Technology*. Washington, DC: National Academy Press.

National Science Foundation (1996), *Shaping the Future: New Expectations for Undergraduate Education in Science, Mathematics, Engineering, and Technology*. Washington, DC: Report by the Advisory Committee to the National Science Foundation Directorate for Education and Human Resources.

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

Potthast, M. J. (1999), "Outcomes of Using Small-Group Cooperative Learning Experiences in Introductory Statistics Courses," *College Student Journal,* 33, 1, 1-8.

Ragasa, C. Y. (2008), "A Comparison of Computer-Assisted Instruction and the Traditional Method of Teaching Basic Statistics," *Journal of Statistics Education* [Online], 16, 1. Available at http://www.amstat.org/publications/jse/v16n1/ragasa.html.

Raudenbush, S. W., and Bryk, A. S. (2002), *Hierarchical Linear Models: Applications and Data Analysis Methods.* Sage Publications.

Raudenbush, S. W., Bryk, A. S., Cheong, Y., and Congdon, R. T. (2004), *HLM 6: Hierarchical Linear and Nonlinear Modeling*, Chicago: Scientific Software International.

Roseth, C. J., Garfield, J. B., and Ben-Zvi, D. (2008), "Collaboration in Learning and Teaching Statistics," *Journal of Statistics Education* [Online], 16, 1.  Available at http://www.amstat.org/publications/jse/v16n1/roseth.html.

Slavin, R. E. (1995), *Cooperative Learning* (2nd Ed.). Boston: Allyn & Bacon.

Snee, R. D. (1993), "What's Missing in Statistical Education?" *The American Statistician*, 47, 149-153.

Springer, L., Stanne, M. E., and Donovan, S. S. (1999), "Effects of Small-Group Learning on Undergraduates in Science, Mathematics, Engineering, and Technology: A Meta-Analysis," *Review of Educational Research*, 69, 1, 21-51.

Tishkovskaya, S., and Lancaster, G. A. (2012), "Statistical Education in the 21st Century: a Review of Challenges, Teaching Innovations and Strategies for Reform," *Journal of Statistics Education* [Online] 20, 2. Available at http://www.amstat.org/publications/jse/v20n2/tishkovskaya.pdf.

Vygotsky, L. S. (1978), *Mind in Society: The Development of the Higher Psychological Processes.* Cambridge, MA: Harvard University Press.

Zieffler, A., Garfield, J., Alt, S., Dupuis, D., Holleque, K., and Chang, B. (2008), "What Does Research Suggest About the Teaching and Learning of Introductory Statistics at the College Level? A Review of the Literature," *Journal of Statistics Education* [Online], 16, 2. Available at www.amstat.org/publications/jse/v16n2/zieffler.html.

---

Sema Kalaian, Ph.D.
Professor of Statistics and Research Design
Eastern Michigan University
College of Technology
School of Technology Studies
161 Sill Hall
Ypsilanti, MI 48197
USA
E-Mail: SKalaian@emich.edu

Rafa Kasim, Ph.D.
Statistical and Research Consultant
Consulting Unlimited
Kent, OH 44240
USA
E-Mail: RafaKasim@gmail.com

---

Volume 22 (2014) | Archive | Index | Data Archive | Resources | Editorial Board | Guidelines for Authors | Guidelines for Data Contributors | Guidelines for Readers/Data Users | Home Page | Contact JSE | ASA Publications

20