

A Meta-Analysis of Teacher and Student-Centered Practices and Processes in Undergraduate Science Education

Robert M. Bernard
Concordia University

Eugene Borokhovski
Concordia University

Brian Mihov
Concordia University

Richard F. Schmid
Concordia University

This meta-analysis investigates the effects of four instructional dimensions rated on a scale from more Teacher-centered (T-C) to more Student-centered (S-C) plus several coded moderator variables on the achievement of undergraduate students in science education courses. More student-centered conditions served as the 'treatment' while more teacher-centered conditions were considered the 'control.' Hedges' g , operationalized as the adjusted standardized differences between treatment and control means, served as the outcome measure. The weighted average difference between groups was $\bar{g} = 0.34$, $k = 140$ (random effects analysis), indicating an overall difference in favor of student-centered instruction. Out of four rated dimensions (Pacing, Teacher's Role, Flexibility, and Adaptation) only Flexibility was significant in meta-regression as a negative predictor of effect size. Two demographic variables (i.e., class size & subject matter), and one instructional moderator variables (i.e., technology use) were also significant when added to Flexibility, producing a model that accounted for 36% of total variation in effect size.

Keywords: student-centered instruction, teacher-centered instruction, learning achievement, meta-analysis

INTRODUCTION AND OBJECTIVES

Perspectives: T-C and S-C Education

To provide students with optimal opportunities to learn and to apply knowledge, the appropriate educational environments need to be present. Two particular learning environments, teacher-centered (T-C) and student-centered (S-C), have enjoyed extensive research (e.g., Kirschner, Sweller & Clark, 2006; Tobias & Duffy, 2009) and application across undergraduate science education, as well as in many other grade levels and content areas. Additionally, a decades-spanning dichotomy has existed between these two

approaches, with a large body of literature characterized by “either/or” (i.e., either this, or that, not both) comparisons.

Teacher-Centered Education (T-C)

Although there has been considerable social experimentation with S-C learning, beginning in the 1970s (e.g., *Summer Hill*, British Infant Schools) until the present, T-C classrooms still represent the predominant pattern, especially in undergraduate science education. According to this orientation, the teacher sets the objectives, plans lessons, teaches students through direct instructional methods (e.g., lecture), assigns readings, provides guidance, evaluates student progress, and awards grades. This has often been referred to as traditional classroom instruction. Observational studies of classroom instruction, such as the research conducted by Rosenshine and Stevens (1986), explored correlations between teacher behaviour and student achievement outcomes, and identified instructional behaviors and patterns that ran parallel to those described via terms such as “direct instruction” (Rosenshine, 1976), “explicit teaching” (Stanovich, 1980), and “systematic teaching” (Morrison, 1926). Specifically, Rosenshine and Stevens (1986) described and grouped their results into six sequential teaching functions reflecting a direct instructional approach: 1) daily review; 2) presenting new material; 3) guiding student practice; 4) providing feedback and corrections; 5) conducting independent practice; and 6) weekly and monthly reviews. Throughout the T-C instructional process, students’ learning experience flows through the instructional conditions and parameters designed, implemented, and monitored by the teacher.

In 1964, a massive experiment called *Project Follow Through* was initiated to test the efficacy of a range of instructional strategies, from so-called Direct Instruction to Open Education. After years of testing and millions of dollars spent, only one really striking finding emerged: direct instruction advantaged students in terms of achievement and affect, outperforming other models by as much as $1.5sd$. While a great deal of controversy surrounds the conduct and findings of *Project Follow Through*, its results set a tone, particularly in mathematics, science and language teaching, that is still influential today (e.g., Klahr, 2009).

Student-Centered Education (S-C)

Constructivist ideas shift the focus of teaching/learning away from direct instruction and teacher dominance, by placing knowledge construction and many other processes by students at the center of the educational enterprise. Jonassen (1991) notes that many educators have applied constructivism to the development of learning environments, and suggests a set of S-C instructional design principles, namely: 1) real-world relevant learning; 2) realistic approaches to solving real-world problems; 3) conceptual interrelatedness and multiple representations or perspectives on the content; 4) goals and objectives are negotiated, not imposed; 5) evaluations that serve as a self-analysis tools; 6) tools and environments that help students interpret multiple perspectives upon the world; and 7) learning that is internally controlled and mediated by the student (Jonassen, 1991). These principles form part of the basis for the “instructional dimensions” that will be examined in this project.

Contrasted with the results of the *Project Follow Through* are more recent meta-analyses (e.g., Rosen & Salomon, 2007) that suggest an advantage for S-C instruction (i.e., constructivist-oriented pedagogy). Rosen and Salomon found a learning difference of $+0.50sd$ for S-C teaching methods and “constructivist-friendly” measures of achievement. Another meta-analysis by Freeman et al. (2014) examined university-level science education along with other STEM subjects, finding an increase in learning over traditional lecturing of $\bar{g} = 0.47sd$ ($k = 158$) and that the odds ratio for failing was 1.95 lower than traditional lecturing. Such a large discrepancy between findings for S-C and T-C instruction suggests that questions of the relative effectiveness of these instructional methods, in general, are still unsettled.

Studying Combinations of T-C and S-C Practices

In spite of the considerable research and conceptual literature on both sides of the S-C and T-C question, the fact is that few classrooms are all of one and none of the other. Gersten et al. (National Mathematics Advisory Panel, 2008) observed, after conducting a systematic review of T-C and S-C mathematics teaching practices: “(We) found no examples of studies in which students were teaching themselves or each

other without any teacher guidance; nor did the Task Group find studies in which teachers conveyed ... content directly to students without any attention to their understanding or response. The fact that these terms, in practice, are neither clearly nor uniformly defined, nor are they true opposites, complicates the challenge of providing a review and synthesis of the literature..." (p. 12). Facing these apparent difficulties in finding a clear demarcation between T-C and S-C instructional conditions, Gersten et al. decided not to synthesize, but to instead describe the studies separately.

This meta-analysis explores undergraduate-level chemistry, physics, biology, geology and psychology content areas and their relationship with T-C and S-C learning. Neither of these has been researched extensively through meta-analysis. In an exception, Aiello and Wolfle (1980) conducted a 30-study meta-analysis on individualized instruction in science compared with a traditional lecture method. Individualized instruction was operationalized by separating studies into particular methods (i.e., audio-tutorial, computer-assisted, personalized system, programmed, and a combination). Therefore, Aiello and Wolfle's meta-analysis was more centered on which method or combinations of instructional methods were superior.

The Current Research Study

This study is an extension of research that applied a conceptual model similar to that of Bernard, Borokhovski, Schmid, Waddington, and Pickup (2019). A literature-derived set of four instructional dimensions (*Pacing, Teacher's Role, Flexibility, and Adaptation*), taken from the 11 instructional dimensions described by Bernard & Borokhovski (2013), was coded on a five-point scale ranging from predominantly T-C (the baseline score of '1') to predominantly S-C (theoretical maximum of '5'). From true and quasi-experimental designs, researchers extracted effect sizes for achievement outcomes ($k = 365$) reflecting comparisons where more S-C condition(s) were designated the treatment and more T-C condition(s) the control. An average weighted effect size of $\bar{g} = 0.44sd$ was found, indicating that classroom studies with more S-C learning produced modestly better achievement results than did studies with less S-C learning. Additionally, meta-regression analysis resulted in two of the four dimensions (*Pacing* and *Teacher's Role*) being significant predictors of ES. Interestingly, *Pacing* was a negative predictor of ES, while *Teacher's Role* was positive.

In the current study, these same four instructional dimensions (i.e., *Pacing, Teacher's Role, Flexibility, and Adaptation*) are used as ES predictors, as well as several additional coded study features: research design, class size, subject matter, and various aspects of technology use to address the following research questions: 1) What are the overall effects of SC in science instruction in higher education classrooms? 2) What do individual dimensions contribute? 3) What is the influence of selected moderator variables? and 4) Is there an overall model which involves dimensions and moderator variables?

METHOD

Literature Search Strategy

Comprehensive literature searches during the relevant search period (i.e., 1960-2012), designed and conducted by a professional Information Specialist, were intended to identify and retrieve primary empirical quantitative studies from the mainstream journals, theses and dissertations, and gray literatures relevant to the major research question. Key terms that were used in search strategies, with some adaptations to account for varying terminology across fields and databases, included: ("teacher centered," "student centered" "learner centered," "inquiry," "problem based," "experiential," "discovery") AND (college, university, postsecondary)," AND ("learn," "achievement*").

A variety of databases were searched within the field of education (e.g. CBCA, ERIC, Education Full Text, EditLib) and in related fields (e.g. ABI/Inform Global, Academic Search Complete). In addition, theses and dissertations were searched (e.g., ProQuest Digital Dissertations and Theses) as well as international systematic review organizations (e.g., Campbell Collaboration, EPPI Centre).

Inclusion/Exclusion Criteria

In order to be included in this meta-analysis, primary studies must have: 1) been publicly available (or archived); 2) been conducted in a formal undergraduate-level (i.e., first years of university) post-secondary educational settings, and address any of the following science subject matters: Biology, Chemistry, Physics, Geology or Psychology (Clinical and Experimental); 3) contained legitimate measures of academic achievement (i.e., instructor-made/researcher-made, standardized); 4) contained at least two groups of students receiving different instructional strategies/practices that could be compared as either more or less S-C instruction; 5) included course content and outcome measures that were compatible in the groups that form these comparisons; 6) contained sufficient descriptions of major instructional events that occurred in all instructional conditions; 7) fulfilled the requirements for either experimental (RCTs) or high-quality quasi-experimental design (QEDs); and 8) contained sufficient statistical information for accurate effect size extraction.

Outcome Measures

All types of objective measures of academic achievement were considered. This included both standardized and non-standardized instructor/research-made assessment tools, as well as both cumulative final examination grades, and averages of several performance tasks covering various components of the course/unit content. Self-assessments were excluded, as were attitudinal and behavioral measures.

Types of Interventions

The intervention in question (treatment condition) was considered to be any combination of instructional events that was rated higher in S-C qualities than a comparison (control) condition. As such, the phenomenon being investigated in this meta-analysis is not an intervention in the typical sense as it is used in the experimental literature. In this case, it is a set of instructional practices that have been rated along a continuum from predominantly T-C to predominantly S-C via ratings on the four instructional events (i.e., dimensions) presented earlier: Pacing, Teacher's Role, Flexibility, and Adaptation.

Subsequently, teaching and learning were deconstructed according to the events associated with them. Therefore, a more S-C learning environment is one in which students play a more central role in the conduct of instructional events, and a more T-C learning environment is one where instructional events are dominated by instructors. As a result, any classroom research study, regardless of the intervention being investigated, is eligible for inclusion as long as there is sufficient information provided as to what each participating group did, including the control condition.

Primary Predictor Variables

Two experienced reviewers, working independently, coded every participating group in each study from 1 (mostly T-C) to 5 (mostly S-C) on the same four primary predictors of student achievement (i.e., effect size-defining dimensions) presented earlier. While we recognize the potential difficulties of High Inference Coding (e.g., Cooper, 2017) all reviewers involved in this project had extensive experience gained through working on the coding on a similar systematic review, as well as extensive prior training involving multiple practice runs on studies previously judged to have been accurately and reliably coded. If there were any disagreements in coding, the reviewers and an additional research team member discussed the discrepancy until a consensus was reached, while documenting the initial coding in order to establish and report independent reliability statistics. This procedure of independent coding, followed by joint discussion and appropriate changes to the coding, was employed at all stages of this review including effect size extraction and coding from moderator variables.

Extracting and Synthesizing Effect Sizes

Two reviewers, working independently, screened abstracts, assessed full-text documents, and extracted *d*-type effect sizes from each study included. Cohen's *d* and its standard error were converted to Hedges' *g* with bias-corrected standard errors and were analyzed/synthesized using the random effects model in *Comprehensive Meta-Analysis*, Version 3.3.070 (Borenstein et al. 2014).

Dealing With Dependent Effects Sizes

Within the studies that were accepted for inclusion, some contained multiple experiments that used the same control condition. If not dealt with these studies could have posed a problem of dependency, resulting in a risk of a making a Type II error (i.e., falsely rejecting the null hypothesis). The issue of independence of findings in a meta-analysis has been repeatedly named among the most critical aspects of methodological quality that can affect the trustworthiness and the applied value of a systematic review (e.g., Cheung, 2019; Hedges, Tipton, & Johnson, 2010; Scammacca, Roberts, & Stuebing, 2014).

To ensure unique representation of data collected from each included sample (not a study, as one study may report several independent data sets), the follow approach was followed, where appropriate: 1) out of all outcomes possibly reported in any given study, only data representing a category in question was used for the effect size extraction and in subsequent analyses; 2) if more than one measure addressed the same outcome type within a category, the most representative measure was selected to reflect the treatment effect; and 3) if several measures were equally representative of the outcome category in question, then individual effect sizes were extracted and averaged to reflect the joint effect of a sequence of instructional events. This approach avoids counting samples more than once, and reduces the chances of bias due to dependency.

Coding Instructional Dimensions

The treatment condition was considered to be any combination of instructional events (i.e., sum of the ratings of the dimensions) that is rated higher in S-C qualities than a comparison (control) condition, even if it was not the treatment group designated in the original study. Descriptions of the coding categories and descriptions of the 1-5 rating criteria are shown in Table 1.

To arrive at an overall differential score for each dimension within each study, ratings of control groups were subtracted from ratings of treatment groups leaving a number indicative of the extent to which treatment groups were higher (i.e., positive sign) or lower (i.e., negative sign) than control groups.

TABLE 1
DESCRIPTION AND CODING DESCRIPTIONS OF FOUR
INSTRUCTIONAL/LEARNING DIMENSIONS

Dimension Descriptions

Code Descriptions

Teacher's Role represents a continuum of a teacher's major responsibilities for organizing/delivering instruction/managing classroom activities, etc.

- 1) Teacher almost exclusively lectures and is the main source of content.
 - 2) Provides some guidance, feedback, initiates and supports discussions, etc.
 - 3) Functions as a guide, coach, tutor, provocateur of thinking.
 - 4) Functions as a colleague, partner in learning.
 - 5) Almost exclusively acts as a facilitator of learning.
-

Pacing reflects the degree of student control over the time of instruction/ learning and over the progression through the course content (i.e., pedagogical flexibility – revisiting/selecting/etc.).

- 1) Instruction is highly structured and progresses step-by-step; no flexibility is allowed.
 - 2) Minor degree of either logistical or pedagogical flexibility is available to students.
 - 3) Program/teacher's control over course progression is balanced with that of students.
 - 4) Students have a substantial amount of flexibility in course progression.
 - 5) High degree of flexibility.
-

Flexibility describes the degree of student control over course design, selection, the provision of study materials, and the setting up of learning objectives.

- 1) No involvement of students (most is determined by the teacher or program/curriculum).
 - 2) Student involvement in at least one of the components of course planning is limited.
 - 3) Teachers and students collaborate in the planning, but teacher's role is still dominant.
 - 4) Teachers and students collaborate in the course planning equally.
 - 5) High student involvement.
-

Adaptability describes the degree to which levels or modifications in instructional process is provided to accommodate individual students.

- 1) Learning materials, settings, group formation, activities are predetermined and unchanged throughout the instruction.
 - 2) Minor modifications are allowed to learning materials, or group composition.
 - 3) Elements of either individualized feedback, or role and tasks assignments.
 - 4) Adapting several instructional components to students' individual needs, or interests.
 - 5) High levels of joint adaptability of several components of instruction.
-

Note: Adapted from Bernard et al. (2019).

The rating method used belongs to a type described by Cooper (2017) as high-inference, because a great deal of training and judgement on the part of coders is required to produce ratings that are consistently valid and reliable. The inter-rater agreement rate for the instructional dimensions coding was judged to be high with Cohen's *Kappa* = 0.89.

RESULTS

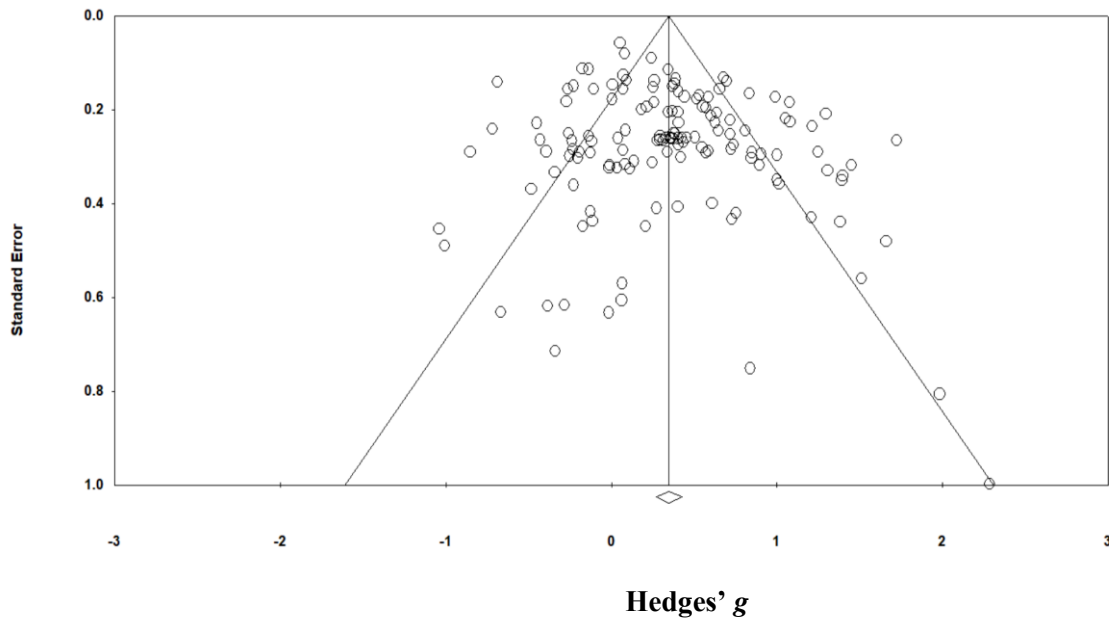
These are the results of searches: 1) total search, 9,759 document abstracts (after 1,285 duplicates removed, 8,474 documents); 2) after abstract screening, 3,749 documents; and 3) final studies after inclusion/exclusion criteria were applied yielded, 101 studies containing 140 individual experiments (effect sizes).

Assessment of the Risks of Bias

Publication Bias

Assessment of publication bias analysis uses a set of graphical and statistical approaches to help assess the adequacy of the corpus of included studies. Observation of the random model funnel plot in conjunction with Duvall and Tweedie's (2000) analytical routine indicated no missing studies from either tail of the distribution. The Funnel Plot (Figure 1) indicates that there was no discernable publication bias on the negative side of the plot (i.e., left of the mean effect size). The Trim and Fill results suggest a similar pattern of inclusiveness.

FIGURE 1
FUNNEL PLOT OF 140 EFFECT SIZES
(HEDGES' G – X-AXIS; STANDARD ERROR – Y-AXIS)



Publication date was subjected to an analysis that involves five options (Table 2): 1960-1979; 1980-1989; 1990-1999; 2000-2009; and 2010-2012. 1960-1979 and 2010-2012, all produced non-significant results ($p = 0.45$ and $p = 0.12$). The years 2000-2009 produced a statistically significant moderate effect size ($\bar{g} = 0.34$) from the largest sample of effect sizes ($k = 80$ out of 140). While 1980-1989 produced the highest effect size ($\bar{g} = 0.65$), it only contained $k = 12$ out of the possible 140 effect sizes. The overall model produced a statistically significant result (Table 2). In spite of these difference, there was no linear effect across publication date, treated as an integer-level variable in meta-regression analysis (i.e., $Q = 0.52$, $df = 1$, $p = .47$).

TABLE 2
PUBLICATION DATE ANALYSIS

Publication Categories	k	\bar{g}	SE	LL	UL	z -value	p -value	Q -Bet.	df	p -value
1960-1979	9	0.07	0.07	-0.11	0.25	0.75	.45			
1980-1989	12	0.65	0.65	0.28	1.01	3.44	.001			
1990-1999	20	0.16	0.16	-0.00	0.32	1.96	.05			
2000-2009	80	0.35	0.49	0.34	0.53	8.95	< .001			
2010-2012	19	0.15	0.12	-0.09	0.39	1.24	.12			
Total between								21.54	4	< .001

Sensitivity Bias Analysis

Sensitivity analysis seeks to determine if effect sizes, especially at the upper and lower ends of the distribution (where higher and lower effect sizes are sometimes paired with anomalously large sample sizes), have any undue influence on the overall random effects outcomes (Borenstein et al., 2009). Table 3 shows the six highest and six lowest effect sizes and their overall influence when they are systematically removed

from the distribution and the effects are recalculated (i.e., one study removed, as described below). The overall procedure was conducted using the software *Comprehensive Meta-Analysis*.

In Table 3, Column 1 is the Study Name and Date of Publication. Column 2 is the actual calculated g for each of the six highest and six lowest effect sizes. Columns 3 through 8 are the means, standard errors, confidence intervals, and the corresponding z -value and p -value when each study is removed and the statistics recalculated. Column 9 is the relative weight that is applied under the random model. Higher weights produce more influence than lower weights.

There appear to be no anomalous results across the 12 studies. This suggests that there is little or no ‘effect size by sample size bias,’ at least at the extremities. This does not mean that there is no bias within the remaining 128 studies, but it is likely that if bias is present in smaller effect size studies, the overall results will not be as affected as it would in these 12 studies. As a result of this analysis, no effect sizes were removed as outliers and no study was Winsorized (i.e., given the value of the next highest or lowest effect).

TABLE 3
SENSITIVITY ANALYSIS (RANDOM EFFECTS)

Study Names	Actual	One Study Removed						Relative Weight
	g	\bar{g}	SE	LL	UL	z	p	
Wozniak2012	2.30	0.35	0.04	0.26	0.42	8.52	.00	0.14
Okebukola1988-2	1.99	0.34	0.04	0.26	0.42	8.51	.00	0.20
Doymus2008	1.73	0.34	0.04	0.26	0.41	8.46	.00	0.73
Okebukola1988-4	1.66	0.34	0.04	0.26	0.41	8.46	.00	0.42
Okebukola1988-2	1.51	0.34	0.04	0.26	0.41	8.49	.00	0.34
Folconer1988	1.45	0.33	0.04	0.26	0.41	8.44	.00	0.64

128 Studies not shown.								
Swanson1990	-0.48	0.35	0.04	0.27	0.43	9.04	.00	0.35
Kapp2011	-0.69	0.35	0.04	0.27	0.42	8.65	.00	0.34
Hulshof2005	-0.72	0.35	0.04	0.27	0.43	8.84	.00	0.35
Reinhardt2012	-0.85	0.35	0.04	0.27	0.43	8.83	.00	0.35
Martin2009-2	-1.01	0.37	0.04	0.27	0.42	8.73	.00	0.35
Martin2009-1	-1.04	0.38	0.04	0.27	0.43	8.76	.00	0.35
Overall ($k = 140$)	0.34	0.34	0.04	0.26	0.42	8.58	0.00	100.00

Research Design Bias Analysis

Both experimental designs, RCTs (i.e., randomized control trials where groups are randomized to conditions), and QEDs (i.e., quasi-experimental designs where group membership is not random) were included in this meta-analysis. By our definition, “high-quality” QEDs are groups not randomized to conditions but possessing some indicator of group equivalence (i.e., pretest). Typically, RCTs produce lower average effect sizes than do QEDs. A significant difference between these groups of studies may indicate that they are too different and should not be combined into an overall average effect size. Table 4 shows the results of this analysis of group differences. They indicate a nearly significant difference between the groups of studies ($p = .06$). The designs, however, were deemed close enough so as to suggest only minimal risk.

TABLE 4
RESEARCH DESIGN BIAS ANALYSIS

Codes	<i>k</i>	\bar{g}	<i>SE</i>	<i>LL</i>	<i>UL</i>	<i>z</i>	<i>p</i>	<i>Q-B</i>	<i>df</i>	<i>p</i>
QEDs	84	0.41	0.05	0.31	0.50	8.10	< .001			
RCTs	56	0.25	0.07	0.12	0.38	3.68	< .001			
Total between								3.54	1	.06

Judging by the outcomes of these analyses, we conclude that the risk of bias in terms of publication inclusion, sensitivity of the collection, and the difference between outcomes from RCTs and QEDs is minimal.

Overall Analysis

Does more S-C instructional approaches to undergraduate-level science learning result in higher achievement outcomes than less S-C (more T-C) approaches? Answering this question (Table 5) involved synthesizing 140 effect sizes, each representing the difference between a designated treatment condition (i.e., predominantly S-C instruction) and a designated control condition (i.e., predominantly T-C instruction). The complete analysis is shown in Table 5. The average weighted of $\bar{g} = 0.34$, $SE = 0.04$, $p < .001$ indicates a modest positive effect for S-C learning conditions. The distribution is significantly heterogeneous under the fixed effect model ($Q_T = 618.13$, $df = 140$, $p = .001$, $I^2 = 0.77$, $t^2 = 0.15$). This is an interesting finding in that it does not unambiguously support either form of instruction to the exclusion of the other, although it does suggest a modest learning advantage for students being given more involvement and freedom in their own learning processes.

TABLE 5
OVERALL RESULTS

Model	<i>k</i>	\bar{g}	<i>SE</i>	<i>LL</i>	<i>UL</i>	<i>z</i>	<i>p</i>
Random effects	140	0.34	0.02	0.24	0.43	8.70	< .001
Between-group Heterogeneity							
Model	<i>Q</i>	<i>df</i>	<i>p</i>	I^2	Tau^2		
Total Collection	606.84	139	< .001	0.77	0.15		

Analysis of Demographics and Instructional Characteristics

Four categories of demographic moderator variables were coded. These were *Subject Matter* (Table 6-a), *Length of Courses/Sessions* (6-b), *Presence and Absence of Technology* (6-c), and Table 6-d, *Average Class Size*, treated as an ordinal variable and conducted in simple meta-regression.

In the analysis of *Subject Matter*, the results from ‘Psychology versus Chemistry’ is the most likely contributor to the significant difference in effects among the variable levels.

In Table 6-b, short or medium length courses appear to perform better than longer courses, even though the overall effect is not significant ($p = .28$).

The results from the presence or absence of technology shown in Table 6-c is complex and non-intuitive. Normally, in the educational technology literature, the presence of technology in the experimental condition produces somewhat higher average effect sizes compared to the control condition (Schmid, Bernard, Borokhovski, Tamim, Abrami, & Surkes (2014). However, in these results, the presence and absence of technology in both treatment and the control are about equal, and are higher than when technology is present

in both conditions. These results may be indicative of the wide range of technologies used over the long timespan of studies (1960s to 2000s) represented in this meta-analysis.

In Table 6-d, average class size is almost, but not quite a significant predictor of achievement results.

TABLE 6A-B-C-D
CODED DEMOGRAPHIC CHARACTERISTICS (MODERATORS) OF THE SAMPLE

Categories	<i>k</i>	\bar{g}	<i>SE</i>	<i>LL</i>	<i>UL</i>	<i>z</i>	<i>p</i>	<i>Q-B</i>	<i>df</i>	<i>p</i>
a) Subject Matter (<i>k</i> = 140)										
Biology	47	0.31	0.06	0.20	0.42	5.56	.00			
Chemistry	24	0.58	0.09	0.40	0.76	6.26	.00			
Geology	5	-0.31	0.10	-0.55	0.45	-0.19	.84			
Physics	37	0.45	0.08	0.2	0.61	5.27	.00			
General Science	7	0.22	0.12	0.45	-0.63	1.20	.23			
Psychology	20	0.06	0.09	-0.12	0.24	0.65	.51			
Total between								20.73	5	.001
b) Length of Courses/Sessions (<i>k</i> = 137; Missing = 3)										
Shorter than a semester	83	0.34	0.057	0.28	0.61	5.27	.00			
Semester	46	0.39	0.06	0.27	0.51	6.54	.00			
Longer than a semester	7	0.05	0.20	0.20	0.42	5.56	.00			
Total between								3.73	2	.28
c) Presence (Yes) and Absence (No) of Technology (<i>k</i> = 140)										
Yes * Treat. /Yes ** Cont.	54	0.19	0.06	0.08	0.30	3.34	.00			
Yes Treat. /No Cont.	21	0.49	0.11	0.27	0.71	4.39	.00			
No Treat. /No Cont.	65	0.42	0.06	0.31	0.54	7.12	.00			
Total between								10.30	2	.01
* Treat. = Treatment; ** Cont. = Control. <i>Note:</i> There were <i>k</i> = 0 for category No Treat. /No Cont.										
d) Average Class Size (<i>k</i> = 140)										
Covariate	Coefficient	<i>SE</i>	<i>LL</i>	<i>UL</i>	<i>z</i>	<i>p</i>	VID			
Intercept	0.42	0.05	0.30	0.50	8.02	.00	1.61			
Class Size (avg.)	-0.0009	0.0005	-0.002	0.0001	-1.81	.07	1.00			
Model: $Q = 2.71$, $df = 1$, $p < .10$ Goodness of Fit: $\tau^2 = 0.15$, $\tau = 0.39$, $F^2 = 76.1$ $Q = 583.35$, $df = 1$, $p < .00001$										

Analysis of Instructional Dimensions

Which primary predictor variable(s) or dimension(s) of student achievement predict effect sizes? Results indicated that the overall set of predictors, tested as a group (i.e., *Teacher's Role, Pacing, Adaptation, and Flexibility*), was not significant ($Q = 3.00$, $df = 4$, $p = 0.56$). However, when tested separately in simple regression the dimension of *Flexibility* (Table 7), produced a significant negative result. This result suggests a tendency towards less flexibility in the ways that are suggested in Table 1 for T-C instruction.

TABLE 7
META-REGRESSION – CODED INSTRUCTIONAL DIMENSION: FLEXIBILITY

Covariate	Coefficient	SE	LL	UL	z	p	VID
Intercept	0.39	0.47	0.30	0.48	8.35	.00	1.35
Flexibility	-0.09	0.05	-0.19	0.02	-1.93	.05	1.00

Model: $Q = 3.73, df = 1, p = .05$ Goodness of Fit: $\tau^2 = 0.15, \tau = 0.39, Q = 606.13, df = 138, p < .001$

Combination of Dimensions and Other Coded Variables

The final analysis involves the combination of two interval variables, *Flexibility* and *Class Size*, and two categorical variables (i.e., dummy-coded variables), *Subject Matter* and *Technology*. The complete model (Table 8) was significant (Q -Between [predictors] = 42.24, $df = 8, p < .0001, R^2 = 0.26$ or 26%) and also significantly heterogeneous (Q -Between [studies] = 430.65, $df = 132, p < .0001, I^2 = 69.35\%$, $\tau^2 = 0.01$). Individually, each predictor in the model was significant (some positive and some negative).

Analysis of the dimensions alone yielded only one result – *flexibility*. *Course Length*, one of the coded moderator variables, was not significant in the complete model and thus was excluded.

TABLE 8
META-REGRESSION COMPLETE MODEL – CODED INSTRUCTIONAL DIMENSION FLEXIBILITY PLUS ‘CLASS SIZE,’ ‘SUBJECT MATTER,’ AND ‘TECHNOLOGY USE’

Covariates	Coefficient	SE	LL	UL	z	P	VID
Intercept	0.06	0.10	-0.14	0.25	1.53	.13	8.78
Flexibility	-0.11	0.04	-0.20	-0.03	-2.60	.009	1.123
*Class Size							
Small	-0.10	0.12	-0.35	0.14	-0.82	.41	1.12
Large	-0.04	0.09	-0.22	0.15	-0.38	.71	1.17
Very Large	-0.29	0.95	-0.47	-0.10	-3.03	.003	1.14
**Subject Matter							
Biology	0.26	0.11	0.05	0.47	2.47	.01	2.29
Chemistry	0.49	0.13	0.24	0.75	3.80	.0001	2.16
Geology	-0.18	0.22	-0.60	0.25	-0.82	.42	1.37
Physics	0.38	0.12	0.15	0.61	3.24	.001	2.31
**Technology Use:							
Treat. Yes/Control No	0.28	0.11	0.06	0.49	2.56	.01	1.51
Treat. No/Control Yes	-1.00	0.40	-1.78	-0.21	-2.49	.01	1.09
Treat. No/Control No	0.19	0.08	0.03	0.36	2.13	.02	1.51

Class Size (overall): $Q = 9.38, df = 3, p = .02$.
 Subject Matter (overall): $Q = 23.94, df = 4, p = .0001$.
 Technology Use (overall): $Q = 15.92, df = 3, p = .001$.
 Model: $Q = 58.98, df = 11, p < .000; R^2$ (analog) = 0.36 or 36% Goodness of Fit: $\tau^2 = 0.09, I^2 = 66.49\%, Q = 381.93, df = 128, p < .00$.
 Reference Categories: *‘Average’ **‘Psychology’ ***‘Treat. No/Cont. No’.

Limitations and Potential Biases in the Review Process

The main limitation and potential bias in the overall review process is the use of high-inference coding (Cooper, 2017). As outlined in the Methods section, the designation of treatment and control in each individual study is not directly derived from the specific designations, but rather determined by two

reviewers working independently and basing their decisions on a set of judgments on each of the four instructional dimensions (*Pacing, Teacher's Role, Flexibility, and Adaptation*). In some cases, this required reversing the group designations. It is important to note in considering the accuracy of coding, and as explained in the Methods section, reviewers received extensive training for this task, which included multiple practice runs on studies previously judged to have been accurately and reliably coded. The same two reviewers also had extensive experience from working on the coding for the several other systematic reviews. The inter-rater agreement rate for the dimensions coding was judged to be high (Cohen's $\kappa = .89$).

While this form of *high-inference coding* does result in greater risk of bias than the standard treatment/control designation (i.e., *low-inference coding*), for the purposes of this systematic review, it is deemed to be the only way to advance the research literature beyond relatively simple comparisons between 'either this or that' like the standard treatment/control designations that populate the educational research literature.

DISCUSSION

Overall, this systematic review provides evidence that S-C instruction may lead to increases in undergraduate level science students' achievement outcomes. This is seen by the medium sized overall random effects average of $\bar{g} = 0.34$, based on average effect sizes that are structured to run in a T-C (smaller) to S-C (larger) direction. However, the argument is weakened by the fact that a significant linear relationship between degree of S-C and achievement was not found ($\beta = 0.01, SE = 0.01, z = 0.03, p = 0.40$). As a result, the simple argument that increasing the amount of S-C learning in an undergraduate level science course will in turn increase student achievement cannot be easily justified. The relationship between S-C learning and student achievement in this particular instructional setting is more nuanced, and factors such as type and degree of S-C learning, as well as combinations of S-C learning and combinations of S-C and T-C learning, likely have their roles.

When considering instructional settings in which undergraduate science courses take place, explanations for the results of the primary predictor variable analyses can only be inferred. The fact that *Pacing, Teacher's Role, and Adaptability* were all non-significant in both meta-regression and mixed moderator variable analysis, could be indicative of the expectation that the typical classroom structure of undergraduate chemistry, biology, physics, geology, and psychology courses, are simply more T-C based. Courses tend to be larger, particularly in the more introductory levels, and they also tend to be predominantly lecture-based. Courses are often structured around the officially recognized textbooks, and professors who use them are sometimes provided pre-produced lecture slides or notes supplied by the manufacturer of the textbook. As a consequence, there may be a certain hesitance at this level of education that resists smaller class sizes, more student involvement, and deviation from a lecture-based culture. In short, the typical undergraduate-level science course might not always afford significant opportunities for the implementation of more S-C based instructional practices. There is no direct evidence of this in these data, but it is important to note that the majority of the studies in this systematic review did not score higher than +2 on each of the four instructional dimensions, with most scoring a differential score of between 0 and +2. This may suggest that efforts to implement S-C practices, especially in undergraduate science course, has been only minimally successful.

While the operationalization of *Flexibility* in the Methods section also has aspects that would enable it to be rated more on the T-C side of the spectrum in undergraduate level science settings, there are certain aspects of the science subject matters chosen for this meta-analysis that could help explain why *Flexibility*, of the four dimensions, was the only significant (albeit negative) predictor of effect size. *Flexibility* in part reflects the degree of student participation in course design, selection of course materials, and the determination of learning objectives. The laboratory components of science courses such as chemistry, biology, and physics could provide the particular instructional setting that reflects an increased degree of student involvement and control in this aspect. Labs are often a mandatory component of undergraduate level science courses, especially in more introductory courses, and within labs students have options of

more flexibility than students within class settings. Of course, students are still working within the overall parameters of the length of time allotted to finish and submit work, and labs are still within the overall, pre-established course syllabus. However, different students (or students and their partners, if it is paired work) can approach work in different ways depending on their individual preferences and strategies. While a laboratory component could be conducive to increased *Flexibility*, other aspects of a science course setting (i.e., class size, lecture-based, rigid syllabus) could result in that same instructional dimension exemplifying T-C components.

Studies in which both participating groups used technology ($\bar{g} = 0.19$) were significantly outperformed by studies in which either both groups did not use technology ($\bar{g} = 0.40$) or the treatment group did and the control group did not ($\bar{g} = 0.48$). Interestingly, the degree of technology use across degrees of *Flexibility* was only significant ($p = 0.03$) when both groups did not use technology, and there was a negative relationship.

The fact that settings in which only the treatment group used technology were not significantly different than settings in which both groups did not use technology could be due to the type of technology being used. The technology found in typical undergraduate level science instructional settings includes PowerPoint slides with the occasional video during lectures, as well as some form of an online course management system, such as Moodle or Blackboard. A meta-analysis conducted by Schmid, Bernard, Borokhovski, Tamim, Abrami, & Surkes (2014) on the effects of technology use on achievement in post-secondary education looked into the different pedagogical uses of technology, and found that technological applications that simply present information in an alternative form, such as PowerPoint, yielded small average effect sizes, in the range of $0.10 > \bar{g} < 0.20$ (i.e., what would be considered ‘trivial’ in the social sciences). The fact that this particular application of technology does not provide much of an advantage in terms of achievement outcomes could be an explanation as to why no significant differences between instructional settings with no technology and settings with only treatments using it exist. However, the above does not explain why achievement outcomes in instructional settings, in which both groups used technology, is significantly lower than when at least one group (i.e., at least the control group) did not use it. If this particular form of technology use does incur benefits, albeit small, as reported by Schmid et al. (2014), then why does adding technology to the control group significantly diminish effect size so much?

Another angle from which to approach the relationship between technology use and effect size is by considering how the technology supports the selected pedagogy. Does the pedagogical use of technology, mentioned above – alternative forms of presenting information – do more to support S-C learning, or T-C learning? Is this interaction further complicated by the nature of the content and learners’ prior knowledge?

In summary, while the overall model suggests that S-C learning does result in modest student learning achievement improvements ($\bar{g} = 0.34$), the effects are somewhat confounded when considering the instructional dimension of *Flexibility*, as well as when considering the potential impact of laboratory components in science courses such as biology, physics, and in particular chemistry. It is important to remember that the discussion about labs is merely speculation – the specific course content, and syllabi, of the science courses in the studies used for this systematic review were not examined. It is also important to remember that a consideration of how the pedagogy itself in these instructional settings influence student achievement outcomes, as well as how factors such as technology support the selected pedagogy, is warranted in the quest to better understand the relationship between instructional settings and student learning achievement outcomes.

REFERENCES

- Aiello, N.C., & Wolfle, L.M. (1980). *A meta-analysis of individualized instruction in science*. Paper presented at the annual meeting of the American Educational Research Association, Boston, MA. Retrieved from <http://files.eric.ed.gov/fulltext/ED190404.pdf>
- Alcazar, M.T.M., & Fitzgerald, V.L. (2005). An experimental design to study the effectiveness of PBL in higher education, in first year science students at a university in Peru, South America. *College*

- Quarterly*, 8(2). Retrieved from http://collegequarterly.ca/2005-vol08-num02-spring/alcazar_fitzgerald.html
- Altiparmak, M., & Nakiboglu Tezer, M. (2009). Hands on group work paper model for teaching DNA structure, central dogma and recombinant DNA. *US-China Education Review*, 6(1), 19–23.
- Andrews, J.D.W. (1984). Discovery and expository learning compared: Their effects on independent and dependent students. *Journal of Educational Research*, 78(2), 80–89. <http://dx.doi.org/10.1080/00220671.1984.10885578>
- Arburn, T.M., & Bethel, L.J. (1999, March). *Assisting at-risk community college students: Acquisition of critical thinking learning strategies*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Boston, MA.
- Armstrong, N., Chang, S-M., & Brickman, M. (2007). Cooperative learning in industrial-sized biology classes. *CBE - Life Sciences Education*, 6(2), 163–171. <http://dx.doi.org/10.1187/cbe.06-11-0200>
- Atan, H., Sulaiman, F., & Idrus, R.M. (2005). The effectiveness of problem-based learning in the web-based environment for the delivery of an undergraduate physics course. *International Education Journal*, 6(3), 430–437.
- Azevedo, R., & Cromley, J.G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology*, 96(3), 523–535. <http://dx.doi.org/10.1037/0022-0663.96.3.523>
- Azevedo, R., Cromley, J.G., & Seibert, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology*, 29(3), 344–370. <http://dx.doi.org/10.1016/j.cedpsych.2003.09.002>
- Barab, S.A., Scott, B., Siyahhan, S., Goldstone, R., Ingram-Goble, A., Zuiker, S.J., & Warren, S. (2009). Transformational play as a curricular scaffold: Using videogames to support science education. *Journal of Science Education and Technology*, 18(4), 305–320. <http://dx.doi.org/10.1007/s10956-009-9171-5>
- Barak, M., & Dori, Y.J. (2005). Enhancing undergraduate students' chemistry understanding through project-based learning in an IT environment. *Science Education*, 89(1), 117–139. <http://dx.doi.org/10.1002/sce.20027>
- Basili, P.A., & Sanford, J.P. (1991). Conceptual change strategies and cooperative group work in chemistry. *Journal of Research in Science Teaching*, 28(4), 293–304. <http://dx.doi.org/10.1002/tea.3660280403>
- Beam, M. (2010). *A comparison of didactic and inquiry teaching methods in a rural community college earth science course* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3440772)
- Bechtel, L.P. (1963). *Comparative effects of differentiated teaching methods on certain personality characteristics of college students: The effect of the traditional approach to teaching psychology as compared to an interpersonal approach to teaching psychology upon beliefs, attitudes, values and adjustment of college students in a course in general psychology* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 302095386)
- Bernard, R.M., & Borokhovski, E. (2013, April). *Teacher-centered and Student-centered pedagogy: A meta-analysis of classroom practices and processes*. Poster presented at the 2013 meeting of the American Educational Research Association, San Francisco, CA.
- Bernard, R.M., Borokhovski, E., Schmid, R.F., Waddington, D.I., & Pickup, D. (2019). Twenty-first century adaptive teaching and individualized learning operationalized as specific blends of student-centered instructional events: A systematic review and meta-analysis. *Campbell Systematic Reviews*, 15, e1017. <https://doi.org/10.1002/c12.1017>
- Bilgin, I. (2006). The effects of pair problem solving technique incorporating Polya's problem solving strategy on undergraduate students' performance in chemistry. *Journal of Science Education*, 7(2), 101–106.

- Bilgin, I., Senocak, E., & Sozibilir, M. (2009). The effects of problem-based learning instruction on university students' performance of conceptual and quantitative problems in gas concepts. *EURASIA Journal of Mathematics, Science & Technology Education*, 5(2), 153–164.
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley.
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2014). *Comprehensive Meta-analysis Version 3.3.070*. Englewood, NJ: Biostat.
- Borokhovski, E., Bernard, R.M., Tamim, R., & Abrami, P.C. (2009, May). *Establishing the direction of effect in meta-analyses with multiple treatments (and no obvious control condition)*. Paper presented at the Campbell Collaboration Ninth Annual Colloquium, Oslo, Norway, May 18–20, 2009.
- Burron, B., James, M.L., & Ambrosio, A.L. (1993). The effects of cooperative learning in a physical science course for elementary/middle level preservice teachers. *Journal of Research in Science Teaching*, 30(7), 697–707. <http://dx.doi.org/10.1002/tea.3660300708>
- Cacciatore, K.L., & Seviran, H. (2009). Incrementally approaching an inquiry lab curriculum: Can changing a single laboratory experiment improve student performance in general chemistry? *Journal of Chemical Education*, 86(4), 498–505.
- Cahyadi, V. (2004). The effect of interactive engagement teaching on student understanding of Introductory Physics at the Faculty of Engineering, University of Surabaya, Indonesia. *Higher Education Research and Development*, 23(4), 455–464. <http://dx.doi.org/10.1080/0729436042000276468>
- Cahyadi, V. (2007). Case 11: No more ‘plug and chug’. In G. Crosling, L. Thomas & M. Heagney (Eds.), *Improving student retention in higher education: The role of teaching and learning* (pp. 131–138). London: Routledge. Retrieved from
- Caldwell, E.C. (1978). Mastery: The essential in PSI. *Teaching of Psychology*, 5(2), 59–65. http://dx.doi.org/10.1207/s15328023top0502_1
- Cayton, T.G. (1975). *A comparison of three methods of teaching child psychology* Available from ProQuest Dissertations and Theses database. (UMI No. 7506309)
- Cheung, M.W-L. (2019). A guide to conducting a meta-analysis with non-independent effect sizes. *Neuropsychology Review*, 29, 387–396. <https://doi.org/10.1007/s11065-019-09415-6>
- Chou, C-H. (1998). *The effectiveness of using multimedia computer simulations coupled with social constructivist pedagogy in a college introductory physics classroom* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9839055)
- Cooper, H. (2017). *Research Synthesis and Meta-analysis: A step-by-step approach* (5th ed.). Los Angeles, CA: SAGE Publications.
- Corbalan, G., Kester, L., & J.G. van Merriënboer, J. (2009). Dynamic task selection: Effects of feedback and learner control on efficiency and motivation. *Learning and Instruction*, 19(6), 455–465. <http://dx.doi.org/10.1016/j.learninstruc.2008.07.002>
- Cox, A.J., & Junkin, W.F., III. (2002). Enhanced student learning in the introductory physics laboratory. *Physics Education*, 37(1), 37–44. <http://dx.doi.org/10.1088/0031-9120/37/1/305>
- Davies, C.S. (1981). Teaching introductory chemistry: Generality of the PSI approach. *Journal of Chemical Education*, 58(9), 686–689. <http://dx.doi.org/10.1021/ed058p686>
- Demastes, S.S., Setlage, K., Jr., & Good, R. (1995). Students’ conceptions of natural selection and its role in evolution: Cases of replication and comparison. *Journal of Research in Science Teaching*, 32(5), 535–550. <http://dx.doi.org/10.1002/tea.3660320509>
- Dori, Y.J., & Belcher, J. (2005). How does technology-enabled active learning affect undergraduate students’ understanding of electromagnetism concepts? *Journal of the Learning Sciences*, 14(2), 243–279.
- Doymus, K. (2008). Teaching chemical bonding through jigsaw cooperative learning. *Research in Science & Technological Education*, 26(1), 47–57. <http://dx.doi.org/10.1080/02635140701847470>

- Doymus, K., Karacop, A., & Simsek, U. (2010). Effects of jigsaw and animation techniques on students' understanding of concepts and subjects in electrochemistry. *Educational Technology Research and Development*, 58(6), 671–691. <http://dx.doi.org/10.1007/s11423-010-9157-2>
- Duvall, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2), 455–463. <https://org/10.1111/j.0006-341X.2000.00455.x>
- Elshout, J.J., & Veenman, M.V. (1992). Relation between intellectual ability and working method as predictors of learning. *Journal of Educational Research*, 85(3), 134–143. <http://dx.doi.org/10.1080/00220671.1992.9944429>
- Emerson, I. (1988). A comparative evaluation of computer based and non-computer based instructional strategies. *Journal of Computers in Mathematics and Science Teaching*, 8(1), 46–49.
- Evans, K.L. (2008). *Learning stoichiometry: A comparison of text and multimedia instructional formats* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3284554)
- Ezrailson, C.M. (2004). *EMIT: Explicit modeling of interactive-engagement techniques for physics graduate teaching assistants and the impact on instruction and student performance in calculus-based physics* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3157053)
- Falconer, K., Wyckoff, S., Joshua, M., & Sawada, D. (2001, April). *Effect of reformed courses in physics and physical science on student conceptual understanding*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.
- Feldon, D.F., Timmerman, B.C., Stowe, K.A., & Showman, R. (2010). Translating expertise into effective instruction: The impacts of cognitive task analysis (CTA) on lab report quality and student retention in the biological sciences. *Journal of Research in Science Teaching*, 47(10), 1165–1185. <http://dx.doi.org/10.1002/tea.20382>
- Franklin, G., Griffin, R., & Perry, N. (1994). Effects of cooperative tutoring on academic performance. *Journal of Educational Technology Systems*, 23(1), 13–25. <http://dx.doi.org/10.2190/YEQT-UVF0-E1Q2-95RV>
- Freeman, S., Eddy, S., McDonough, M., Smith, M., Okoroafor, N., Jordt, H., & Winderoth, M.P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America*, 111(23), 8410–8415. Retrieved from <https://www.jstor.org/stable/23776432>
- Friedel, C., Irani, T., Rudd, R., Gallo, M., Eckhardt, E., & Ricketts, J. (2008). Overtly teaching critical thinking and inquiry-based learning: A comparison of two undergraduate biotechnology classes. *Journal of Agricultural Education*, 49(1), 72–84.
- Gersten, R., Chard, D., Jayanthi, M., Baker, S., Morphy, P., & Flojo, J. (2008). *Mathematics instruction for students with learning disabilities or difficulty learning mathematics: A synthesis of the intervention research*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.
- Gifford, V.D., & Vicks, J. (1982). A comparison of the personalized system of instruction and a conventional biology course on the achievement of junior college freshmen. *Journal of Research in Science Teaching*, 19(8), 659–664. <http://dx.doi.org/10.1002/tea.3660190805>
- Grossman, P., Stewart, T., Jaspers, M., & Chapman, B. (2007). Integrating web-delivered problem-based learning scenarios to the curriculum. *Active Learning in Higher Education*, 8(2), 139–153. <http://dx.doi.org/10.1177/1469787407077986>
- Hall, D.A., & McCurdy, D.W. (1990). A comparison of a biological sciences curriculum study (bscs) laboratory and a traditional laboratory on student achievement at two private liberal arts colleges. *Journal of Research in Science Teaching*, 27(7), 625–636. <http://dx.doi.org/10.1002/tea.3660270703>
- Hedges, L.V., Tipton, E., & Johnson, M.C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, 1(1), 39–65. <https://doi.org/10.1002/jrsm.5>

- Hill, L.A. (1999). *The implementation of learning strategies in the design of an on-line medical course supplement* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9926696)
- Hulshof, C.D., Eysink, T.H.S., Loyens, S., & de Jong, T. (2005). ZAPs: Using interactive programs for learning psychology. *Interactive Learning Environments*, 13(1–2), 39–53. <http://dx.doi.org/10.1080/10494820500224079>
- Ibrahim, H.A. (2001). *Examining the impact of the Guided Constructivist teaching method on students' misconceptions about concepts of Newtonian physics* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3029041)
- Jonassen, D.H. (1991). Evaluating constructivist learning. *Educational Technology*, 31, 2833.
- Jones, N.W. (1980). A comparison between lecture and self-paced general geology. *Journal of Geological Education*, 28(5), 235–237. <http://dx.doi.org/10.5408/0022-1368-28.5.235>
- Kapp, J.L., Slater, T.F., Slater, S.J., Lyons, D.J., Manhart, K., Wehunt, M.D., & Richardson, R.M. (2011). Impact of redesigning a large-lecture introductory earth science course to increase student achievement and streamline faculty workload. *Journal of College Teaching and Learning*, 8(4), 23–35.
- Kirschner, P.A., Sweller, J., & Clark, R.E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41, 75–86. https://doi.org/10.1207/s15326985ep4102_1
- Klahr, D. (2009). “To everything there is a season, and a time to every purpose under the heavens”: What about direct Instruction? In S. Tobias & T.M. Duffy (Eds.), *Constructivist Theory Applied to Instruction: Success or Failure?* (pp. 291–310). New York, NY: Routledge.
- Knight, J.K., & Wood, W.B. (2005). Teaching more by lecturing less. *Cell Biology Education*, 4(4), 298–310. <http://dx.doi.org/10.1187/05-06-0082>
- Koenig, K.M., Endorf, R.J., & Braun, G.A. (2007). Effectiveness of different tutorial recitation teaching methods and its implications for ta training. *Physical Review Special Topics - Physics Education Research*, 3(1), 010104. <http://dx.doi.org/10.1103/PhysRevSTPER.3.010104>
- Kremer, J.F., & Dietzen, L.L. (1991). Two approaches to teaching accurate empathy to undergraduates: Teacher-intensive and self-directed. *Journal of College Student Development*, 32(1), 69–75.
- Lee, Y-J. (2010). Effects of instructional preparation strategies on problem solving in a web-based learning environment. *Journal of Educational Computing Research*, 42(4), 385–406.
- LeTexier, K. (2009). *Storytelling as an active learning strategy in introduction to psychology courses* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3320287)
- Levinson, A.J., Weaver, B., Garside, S., McGinn, H., & Norman, G.R. (2007). Virtual reality and brain anatomy: A randomised trial of e-learning instructional designs. *Medical Education*, 41(5), 495–501. <http://dx.doi.org/10.1111/j.1365-2929.2006.02694.x>
- Lewis, S.E., & Lewis, J.E. (2005). Departing from lectures: An evaluation of a peer-led guided inquiry alternative. *Journal of Chemical Education*, 82(1), 135. <http://dx.doi.org/10.1021/ed082p135>
- Liang, L.L., & Gabel, D.L. (2005). Effectiveness of a constructivist approach to science instruction for prospective elementary teachers. *International Journal of Science Education*, 27(10), 1143–1162. <http://dx.doi.org/10.1080/09500690500069442>
- Lord, T., & Orkwiszewski, T. (2006). Moving from didactic to inquiry-based instruction in a science laboratory. *American Biology Teacher*, 68(6), 342–345. [http://dx.doi.org/10.1662/0002-7685\(2006\)68\[342:DTIII\]2.0.CO;2](http://dx.doi.org/10.1662/0002-7685(2006)68[342:DTIII]2.0.CO;2)
- Martin, M. (2009). *The effect of active techniques combined with didactic lecture on student achievement* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 1467105)

- Martin, T., Rivale, S.D., & Diller, K.R. (2007). Comparison of student learning in challenge-based and traditional instruction in biomedical engineering. *Annals of Biomedical Engineering*, 35(8), 1312–1323. <http://dx.doi.org/10.1007/s10439-007-9297-7>
- Mathew, E. (2008). *Learning physics: A comparative analysis between instructional design methods* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3307901)
- McKee, E., Williamson, V.M., & Ruebush, L.E. (2007). Effects of a demonstration laboratory on student learning. *Journal of Science Education and Technology*, 16(5), 395–400. <http://dx.doi.org/10.1007/s10956-007-9064-4>
- McLaren, I.A.M., & Webber, D. (2009). Writing right: Enhancing student engagement and performance in an ecology course. *International Journal of Environmental and Science Education*, 4(4), 365–380.
- Moreno, R. (2009). Constructing knowledge with an agent-based instructional program: A comparison of cooperative and individual meaning making. *Learning and Instruction*, 19(5), 433–444. <http://dx.doi.org/10.1016/j.learninstruc.2009.02.018>
- Morgil, I., Penn, J.H., Secken, N., & Oskay, O.O. (2006). Introduction to complexation and masking within a computer-enriched module for analytical chemistry. *Turkish Online Journal of Distance Education*, 7(2). Retrieved from http://tojde.anadolu.edu.tr/makale_goster.php?id=261
- Morris, E.K., Surber, C.F., & Bijou, S.W. (1978). Self- versus instructor-pacing: Achievement, evaluations, and retention. *Journal of Educational Psychology*, 70(2), 224–230. <http://dx.doi.org/10.1037/0022-0663.70.2.224>
- Morrison, H.C. (1926). *The practice of teaching in the secondary school*. Chicago, IL: University of Chicago Press.
- Muller, D.A., Sharma, M.D., & Reimann, P. (2008). Raising cognitive load with linear multimedia to promote conceptual change. *Science Education*, 92(2), 278–296. <http://dx.doi.org/10.1002/sce.20244>
- Munyofu, M. (2008). *Effects of varied enhancement strategies (chunking, feedback, gaming) in complementing animated instruction in facilitating different types of learning objectives* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3414357)
- National Mathematics Advisory Panel. (2008). *Foundations for Success: The Final Report of the National Mathematics Advisory Panel*. Washington, DC: U.S. Department of Education.
- Nokes, T.J., & Ohlsson, S. (2005). Comparing multiple paths to mastery: What is learned? *Cognitive Science*, 29(5), 769–796. http://dx.doi.org/10.1207/s15516709cog0000_32
- Nugent, G., Kunz, G., Levy, R., Harwood, D., & Carlson, D. (2008). The impact of a field-based, inquiry-focused model of instruction on preservice teachers' science learning and attitudes. *Electronic Journal of Science Education*, 12(2), 94–111.
- Nugent, G., Toland, M.D., Levy, R., Kunz, G., Harwood, D., Green, D., & Kitts, K. (2012). The impact of an inquiry-based geoscience field course on pre-service teachers. *Journal of Science Teacher Education*, 23(5), 503–529. <http://dx.doi.org/10.1007/s10972-012-9283-2>
- Okebukola, P.A., & Jegede, O.J. (1988). Cognitive preference and learning mode as determinants of meaningful learning through concept mapping. *Science Education*, 72(4), 489–500. <http://dx.doi.org/10.1002/sce.3730720408>
- Olajide, J.O., & Adeoye, F.A. (2010). The effect of analogical reasoning and extended wait time on achievement in biology. *International Journal of Educational Sciences*, 2(1), 21–28. Retrieved from <http://www.krepublishers.com/02-Journals/IJES/IJES-02-0-000-10-Web/IJES-02-1-000-10-Abst-PDF/IJES-02-1-21-10-027-Olajide-J-O/IJES-02-1-21-10-027-Olajide-J-O-Tt.pdf>
- Olson, K.V. (1962). An experimental evaluation of a student-centered method and a teacher-centered method of biological science instruction for general education of college students. *Science Education*, 46(4), 367–373. <http://dx.doi.org/10.1002/sce.3730460414>
- Orwin, R.G. (1983). A fail-safe *N* for effect size in meta-analysis. *Journal of Educational Statistics*, 8(2), 157–159.

- Perry, J., Meir, E., Herron, J.C., Maruca, S., & Stal, D. (2008). Evaluating two approaches to helping college students understand evolutionary trees through diagramming tasks. *CBE - Life Sciences Education*, 7(2), 193–201. <http://dx.doi.org/10.1187/cbe.07-01-0007>
- Proske, A., Narciss, S., & McNamara, D.S. (2012). Computer-based scaffolding to facilitate students' development of expertise in academic writing. *Journal of Research in Reading*, 35(2), 136–152. <http://dx.doi.org/10.1111/j.1467-9817.2010.01450.x>
- Quitadamo, I.J., & Kurtz, M.J. (2007). Learning to improve: Using writing to increase critical thinking performance in general education biology. *CBE - Life Sciences Education*, 6(2), 140–154. <http://dx.doi.org/10.1187/cbe.06-11-0203>
- Quitadamo, I.J., Faiola, C.L., Johnson, J.E., & Kurtz, M.J. (2008). Community-based inquiry improves critical thinking in general education biology. *CBE - Life Sciences Education*, 7(3), 327–337. <http://dx.doi.org/10.1187/cbe.07-11-0097>
- Reinhardt, C.H., & Rosen, E.N. (2012). How much structuring is beneficial with regard to examination scores? A prospective study of three forms of active learning. *Advances in Physiology Education*, 36(3), 207–212. <http://dx.doi.org/10.1152/advan.00108.2011>
- Rosen, Y., & Salomon, G. (2007). The differential learning achievements of constructivist technology-intensive learning environments as compared with traditional ones: A meta-analysis. *Journal of Educational Computing Research*, 36, 1–14.
- Rosenshine, B. (1976). Classroom instruction. In N.L. Gage (Ed.), *The psychology of teaching methods* (75th NSSE Yearbook) (pp. 335–371). Chicago, IL: University of Chicago Press.
- Rosenshine, B., & Stevens, R. (1986). Teaching functions. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 376–391). New York: Macmillan.
- Ruiter, W.W. (1971). *An analysis of the effectiveness of a learner-centered teaching system compared to that of a conventional teaching of basic electricity to university students* Available from ProQuest Dissertations and Theses database. (UMI No. 7125072)
- Scammacca, N., Roberts, G., & Stuebing, K.K. (2014). Meta-analysis with complex research designs dealing with dependence from multiple measures and multiple group comparisons. *Review of Educational Research*, 84(3), 328–364. <https://doi.org/10.3102/0034654313500826>
- Schmid, R.F., Bernard, R.M., Borokhovski, E., Tamim, R.M., Abrami, P.C., Surkes, M.A., . . . Woods, J. (2014). The effects of technology use in postsecondary education: A meta-analysis of classroom applications. *Computers & Education*, 72, 271–291. <http://dx.doi.org/10.1016/j.compedu.2013.11.002>
- Selcuk, G.S. (2010). The effects of problem-based learning on pre-service teachers' achievement, approaches and attitudes towards learning physics. *International Journal of the Physical Sciences*, 5(6), 711–723.
- Senocak, E., Taskesenligil, Y., & Sozbilir, M. (2007). A study on teaching gases to prospective primary science teachers through problem-based learning. *Research in Science Education*, 37(3), 279–290. <http://dx.doi.org/10.1007/s11165-006-9026-5>
- Sligh, D.F. (2005). Assessment of the use of the jigsaw method and active learning in non-majors, introductory biology. *Bioscene: Journal of College Biology Teaching*, 31(4), 4–10.
- Spivey, N.R. (1995). *Reciprocal teaching of lecture comprehension and comprehension monitoring skills in college students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9503727)
- Stanovich, K.E. (1980). Effects of explicit teaching and peer tutoring on the reading achievement of learning disabled and low-performing students in regular classrooms. *Reading Research Quarterly*, 16, 32–71.
- Stark, R., Puhl, T., & Krause, U-M. (2009). Improving scientific argumentation skills by a problem-based learning environment: Effects of an elaboration tool and relevance of student characteristics. *Evaluation & Research in Education*, 22(1), 51–68. <http://dx.doi.org/10.1080/09500790903082362>

- Stiller, K.D., Freitag, A., Zinnbauer, P., & Freitag, C. (2009). How pacing of multimedia instructions can influence modality effects: A case of superiority of visual texts. *Australasian Journal of Educational Technology*, 25(2), 184–203. <http://dx.doi.org/10.14742/ajet.1149>
- Stout, L.J. (1978). *A comparison of four different pacing strategies of Personalized System of Instruction and a traditional lecture format* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7808600)
- Strawitz, B.M., & Malone, M.R. (1987). Preservice teachers' acquisition and retention of integrated science process skills: A comparison of teacher-directed and self-instructional strategies. *Journal of Research in Science Teaching*, 24(1), 53–60. <http://dx.doi.org/10.1007/BF02357174>
- Struyven, K., Dochy, F., & Janssens, S. (2010). “Teach as you preach”: The effects of student-centred versus lecture-based teaching on student teachers' approaches to teaching. *European Journal of Teacher Education*, 33(1), 43–64. <http://dx.doi.org/10.1080/02619760903457818>
- Sturges, D., Maurer, T.W., & Cole, O. (2009). Understanding protein synthesis: A role-play approach in large undergraduate human anatomy and physiology classes. *Advances in Physiology Education*, 33(2), 103–110. <http://dx.doi.org/10.1152/advan.00004.2009>
- Suits, J.P. (2004). Assessing investigative skill development in inquiry-based and traditional college science laboratory courses. *School Science and Mathematics*, 104(6), 248–257. <http://dx.doi.org/10.1111/j.1949-8594.2004.tb17996.x>
- Tarhan, L., & Sesen, B.A. (2012). Jigsaw cooperative learning: Acid-base theories. *Chemistry Education Research and Practice*, 13(3), 307–313. <http://dx.doi.org/10.1039/C2RP90004A>
- Tobias, S., & Duffy, T.M. (Eds.). (2009). *Constructivist instruction: Success or failure*. Routledge: New York, NY.
- van den Boom, G., Paas, F., & van Merriënboer, J.J.G. (2007). Effects of elicited reflections combined with tutor or peer feedback on self-regulated learning and learning outcomes. *Learning and Instruction*, 17(5), 532–548. <http://dx.doi.org/10.1016/j.learninstruc.2007.09.003>
- Veenman, M.V.J., & Elshout, J.J. (1994). Differential effects of instructional support on learning in simulation environments. *Instructional Science*, 22(5), 363–383. <http://dx.doi.org/10.1007/BF00891961>
- Vreven, D., & McFadden, S. (2007). An empirical assessment of cooperative groups in large, time-compressed, introductory courses. *Innovative Higher Education*, 32(2), 85–92. <http://dx.doi.org/10.1007/s10755-007-9040-1>
- Walker, J.D., Cotner, S.H., Baepler, P.M., & Decker, M.D. (2008). A delicate balance: Integrating active learning into a large lecture course. *CBE - Life Sciences Education*, 7(4), 361–367. <http://dx.doi.org/10.1187/cbe.08-02-0004>
- Walker, J.P., Sampson, V., Grooms, J., Anderson, B., & Zimmerman, C.O. (2012). Argument-driven inquiry in undergraduate chemistry labs: The impact on students' conceptual understanding, argument skills, and attitudes toward science. *Journal of College Science Teaching*, 41(4), 74–81.
- Wittwer, J., Nuckles, M., Landmann, N., & Renkl, A. (2010). Can tutors be supported in giving effective explanations? *Journal of Educational Psychology*, 102(1), 74–89. <http://dx.doi.org/10.1037/a0016727>
- Wozniak, B.M. (2012). *Effect of process-oriented guided-inquiry learning on non-major biology students' understanding of biological classification* (Master's thesis). Available from ProQuest Dissertations and Theses database. (UMI No. 1517251)
- Wright, J.M. (2008). The comparative effects of constructivist versus traditional teaching methods on the environmental literacy of postsecondary nonscience majors. *Bulletin of Science, Technology & Society*, 28(4), 324–337. <http://dx.doi.org/10.1177/0270467608319638>