Learning Disabilities Research & Practice, *26*(4), 188–203

Science Instruction for Students with Learning Disabilities: A Meta-Analysis

William J. Therrien, Jonte C. Taylor, John L. Hosp, Erica R. Kaldenberg, Jay Gorsh
The University of Iowa

Although science has received much attention as a political and educational initiative, students with learning disabilities (LD) perform significantly lower than their nondisabled peers. This meta-analysis evaluates the effectiveness of instructional strategies in science for students with LD. Twelve studies were examined, summarized, and grouped according to the type of strategy implemented. Effect sizes (*ES*) were calculated for each study. Across all studies, a mean *ES* of .78 was obtained, indicating a moderate positive effect on students with LD science achievement. Findings also align with past reviews of inquiry-based instruction for students with special needs, indicating that students with LD need structure within an inquiry science approach in order to be successful. Additionally, results suggest that mnemonic instruction is highly effective at increasing learning disabled students' acquisition and retention of science facts.

Recent reports lament the poor performance of U.S. students in the area of science. The importance of science education was highlighted in the president's "State of the Union" speech in January 2011. In his speech, the president connected the need "to advance science and math education" to the "nation's long-term economic prosperity" (White House, Office of the Press Secretary, 2011). Educators have also expressed urgency for increased emphasis on science learning as a means of economic, social, and industrial advancement (Turner, 2008; Villanueva, 2010).

Recently administered assessments indicate that the concern for science achievement is warranted. The Program for International Student Assessment (PISA), administered in 2009, provides the support for the argument that U.S. students are lagging behind students in other "developed nations" (Fleischman, Hopstock, Pelczar, & Shelley, 2010). Fifteen Organizations for Economic Cooperation and Development (OECD) countries scored higher than the U.S. and only 29 percent of U.S. students scored at or above level four in science proficiency. Students scoring at level 4 proficiency "can select and integrate explanations from different disciplines of science or technology and link those explanations directly to aspects of life situations" (p. 25). Therefore, these results indicate that over 70 percent of U.S. students are unable to connect science to real life situations or integrate science content across disciplines.

Domestic assessments have also indicated a need for increased science achievement. The National Assessment of Educational Progress (NAEP) science scores for 2009 indicated that only one-third of students in fourth grade scored at or above proficient in science and less than one-third of students in 8th and 12th grades scored at or above proficient (National Center for Education Statistics, 2011).

Requests for reprints should be sent to William J. Therrien. Electronic mail may be sent to bill-therrien@uiowa.edu .

Equally poor performance in science has also been reported for students with special needs including students with learning disabilities (LD) (Anderman, 1998; Grigg, Lauko, & Brockway, 2006; Steele, 2004). In fact, according to the NAEP, fourth, eighth, and 12th grade students with disabilities scored significantly lower than students without disabilities. There are several potential reasons why students with LD have difficulty in science. Language disabilities in reading and writing are likely the primary factors that negatively impact a student with LD's science achievement (Parmar, Deluca, & Janczak, 1994; Shepard & Adjogah, 1994; Steele, 2004). Additionally, because science generally involves the use of mathematics, difficulties in this area may also contribute to the limited science achievement for students with LD (Olson & Platt, 2004). Difficulties with academic-related tasks are not the sole problems that students with LD have when it comes to learning science content. There are also behavioral issues that may hinder academic success in science. Steele (2004) noted that students with LD have difficulties behaviorally in the specific areas of sustained attention, attitude toward science, and social skills—all of which predict success in the general education setting.

In order to evaluate the science instruction that students with LD receive, the general education classroom must be examined as the vast majority of students with LD are instructed in general education classrooms (Cawley, Hayden, Cade, & Baker-Kroczynski, 2002). Along with core general education instruction, supplemental programming must also be examined because students with LD are often provided with a variety of additional supports in science (Holahan, McFarland, & Piccillo, 1994).

Traditional general education science instruction can best be described as being textbook and/or lecture driven (Scruggs & Mastropieri, 2007; Steele, 2004). The heavy emphasis on language in instruction is a barrier to learning for many students, including students with LD (Parmar, Duluca, & Janczak, 1994). Due to these student difficulties, all the major professional science organizations now strongly advocate for the use of inquiry-based and hands-on instruction (American Association for the Advancement of Science, 1997; Donovan & Bransford, 2005; National Science Teachers Association Board of Directors, 2004).

Scruggs and Mastropieri (2007) explain inquiry-based instruction as the construction of content knowledge that emphasizes the depth of learning and comprehension of relevant concepts. Inquiry-based instruction helps students focus on the big ideas in science and develop the ability to reason scientifically (Palincsar, Magnusson, Cutter, & Vincent, 2002). Students are encouraged to construct their own science knowledge using questions, observations, and investigations as a means of discovery (Maroney, Finson, Beaver, & Jensen, 2003). Scruggs, Mastropieri, and Boon (1998) noted that while the shift is for more inquiry-based and hands-on activities for teaching students with LD in science, there is still room for varying levels of structure to maximize student success. Instead of pure inquiry-based instruction that is completely student directed with knowledge only constructed via discovery, Cawley (1994) and Scruggs and Mastropieri (1994) suggest that the practice of guided inquiry should be used for students with LD. Guided instruction incorporates the use of inquiry-based instruction and direct instruction (Cawley, 1994). In a review of their own research, Scruggs and Mastropieri (2007) expressed with caution that while students with disabilities can benefit from discovery learning, "instruction needed to be particularly intensive and structured" (p. 65).

Along with core inquiry instruction, students with LD often receive additional supplemental instruction to improve their achievement in science (Holahan, McFarland, & Piccillo, 1994). There have been a number of researchers who have provided evidence regarding the use of supplemental supports in science instruction for students with LD. Some identified effective supplemental supports include the use of a structured teaching format and mnemonics (King-Sears, Mercer, & Sindelar, 1992; Scruggs & Mastropieri, 2007); the incorporation of universal design principles (Cawley, Foley, & Miller, 2003); multiple presentation formats (Steele, 2004); and cooperative learning groups (Scruggs & Mastropieri, 2007; Scruggs, Mastropieri, Bakken, & Brigham, 1993). Munk, Bruckert, Call, Stoehrmann, & Randandt, (1998) suggested the use of multiple supports or accommodations dependent upon the needs of the student and the demands of the lesson.

Literature reviews on teaching students with disabilities in science were conducted in 1992 and 1998, respectively. Mastropieri and Scruggs (1992) were able to identify and evaluate (as positive or negative findings) studies that examined instructional methods and accommodations that included the following: instructional variables, text adaptations, mnemonic strategies, and other elaborative techniques. A number of positions were asserted in the conclusion of the literature review: (a) students with disabilities can benefit from good science instruction, (b) when appropriate, this instruction is mandated to occur in general education classrooms, (c) accommodations for students with disabilities in science education classes may be compatible with science instruction, and (d) applying these accommodations may be overwhelming to general education teachers (Mastropieri & Scruggs, 1992).

Scruggs, Mastropieri, & Boon (1998) conducted a literature review of studies regarding teaching students with disabilities science content. Their review highlighted research that focused on learning characteristics, intervention studies, and inclusion research. The authors presented a number of conclusions that students with disabilities can be successful in either self-contained or inclusive classrooms, students benefit from structured activity-based instruction, and teaching students with disabilities requires training and support. Ultimately, the authors contend that students with disabilities can benefit from constructivist approaches (i.e., discovery learning or inquiry-based) with the appropriate structured supports.

The current meta-analysis was conducted in order to build and expand on the previous literature reviews. Specifically, this analysis builds on the previous literature reviews in the following ways. First, instead of focusing on all students with disabilities, this review addresses only students with LD. Second, the review includes articles published since the previous reviews were conducted. Third, meta-analytic techniques are used to allow for the aggregation and comparison of results within and across studies.

The overall purpose of the meta-analysis was to examine classroom-based instruction for students with LD in the content area of science. We specifically were interested in the following research questions: (1) What classroom-based instructional methods are effective at increasing the science achievement of students with LD? (2) What student characteristics are necessary to benefit from the intervention(s) summarized in the meta-analyses? (3) What dependent measure(s) are represented by the effect size(s) reported? (4) What study-specific variables (if any) impact the effect size magnitude reported?

METHOD

Literature Search

In order to locate the articles, we first formed eligibility requirements for the studies that would be considered for the review. Articles must (a) have been published in peer-reviewed journals in 1980 or later (b) have been focused on classroom-based interventions[1] (e.g., class-based interventions were defined as interventions that were not geared solely toward acquiring knowledge via reading), (c) have been experimental or quasi-experimental in nature, and (d) have been included school-aged (i.e., grades K-12) participants with LD as subjects.

Next, we located the studies in three ways. First, we collected all articles that met our criteria that were reviewed in the previous two literature reviews (Mastropieri & Scruggs, 1992; Scruggs, Mastropieri, & Boon, 1998) on science education and students with LD. Second, both Educational Resources Information Center (ERIC) and PsycINFO were searched from 1995 to 2010 using the following search terms: *science and learning disabilities; science and exceptional children; and science and special education.* Last, we conducted a hand search of the following journals from 2000 to 2010: *Journal of Special Education, Journal of Learning Disabilities, Learning Disabilities Research and Practice, Exceptional Children*, *Journal of Research in Science Teaching*, *International Journal of Science Education*, and *Science Education*. In total the search produced 26 articles that met our initial criteria.

Next, we reviewed the articles in depth to ensure that they met our inclusion criteria and to determine if they were amenable to meta-analysis techniques. Through this process, we excluded 15 articles for the following reasons: Five articles were excluded because the primary focus of the study was not on science instruction (Bakken & Mastropieri, 1997; Espin & Deno, 1993; Kooy, 1992; Marino, Coyne, & Dunn, 2010; Nelson, Smith, & Dodd 1992). Eight articles were not analyzed because they did not meet initial requirements (Anderman, 1998; Carlisle & Chang, 1996; Cawley, Hayden, Cade, & Baker-Kroczynski, 2002; Gaddy, Bakken, & Fulk, 2008; McFarland & Shepard, 1995; Parmar, Duluca, & Janczak, 1994; Scruggs & Mastropieri, 1994; and Shepard & Adjogah, 1994). One article was excluded because it summarized three studies already included in the review (Mastropieri, Scruggs, & Magnusen, 1999). One article was excluded because it was a single subject design, and therefore, not amenable to meta-analysis methodology (MacDougall, Schnur, Berger, & Vernon, 1981). After this process was complete, a total of 11 articles remained (Bay, Staver, Bryan, & Hale, 1992; Dalton, Morocco, Tivnan, & Rawson Mead, 1997; King-Sears, Mercer, & Sindelar, 1992; Mastropieri, Scruggs, & Levin, 1985; Mastropieri, Scruggs, & Levin, 1986; Mastropieri et al., 1998, Mastropieri et al., 2006; McCleery & Tindal, 1999; Scruggs, Mastropieri, Bakken, & Brigham, 1993; Scruggs, Mastropieri, Levin, & Gaffney, 1985; and Scruggs, Mastropieri, & Sullivan, 1994).

Coding

Articles were initially coded for eight criteria important to the analysis. *Random assignment* indicated whether or not the subjects were randomly assigned to treatment and control groups (0 = no, 1 = yes). *LD broken out* indicated whether or not the students with LD results were analyzed separately (0 = no, 1 = yes). *Grade level* indicated the grade of students included in the study (0 = 3rd–6th grade, 1 = 7th grade+). No studies were found that included students in kindergarten, first or second grade. *IQ reported* indicated whether or not students' intelligence quotient (IQ) score was reported (0 = no, 1 = yes). *Type of intervention* indicated whether the intervention was structured inquiry, supplemental mnemonics, or another supplemental technique (0 = other supplemental, 1 = mnemonic, and 2 = inquiry). *Duration* indicated the number of sessions in the intervention. Although coded as a continuous variable, the data clearly broke into two groups (0 = 4 or fewer, 1 = 12 or more). *Source of measure* indicated whether or not the dependent measure was experimenter generated or not. All dependent measures but one measure from one study was experimenter constructed. Therefore, this variable was removed from further analyses, but is addressed further in the Discussion. *Type of dependent measure* indicated whether the outcome measure was administered immediately after the intervention, after a delay, or was a measure of skill/knowledge generalization. Because of the lack of many delay and generalization measures across studies, these measures were combined into a delay/generalization category (0 = immediate, 1 = delayed/generalization).

The studies ($n = 12$)[2] in the final sample were concurrently coded by three raters. All disagreements were reconciled during the coding sessions with reconciled codes used in all subsequent analyses.

Effect Size Calculations

The effects within each study were coded as the standardized mean difference between treatment and control groups. Initially, all studies were coded using various formulae depending on the information provided in the article. The standardized mean difference was calculated in one of three ways. If means and standard deviations (*SD*s) were provided for both treatment (t) and control (c) groups, the standard formula was used subtracting the control mean from the treatment mean and dividing by the pooled *SD*.

$$ES_1 = \frac{\bar{X}_t - \bar{X}_c}{sd_p}$$

where the pooled *SD* is calculated using the following formula:

$$SD_p = \sqrt{\frac{SD^2(N-1) - \frac{(\bar{X}_t^2 + X_c^2 - 2\bar{X}_t\bar{X}_c)(n_t n_c)}{n_t + n_c}}{N-1}}$$

If means and/or *SDs* were not provided but a *t* value was, the standardized mean difference was derived using the formula for calculating effect size from independent *t*-test results:

$$ES_2 = t\sqrt{\frac{n_t + n_c}{n_t n_c}}$$

If means and/or *SDs* were not provided but an F-value was, the standardized mean difference was derived using the formula for calculating effect size from analysis of variance (ANOVA) results:

$$ES_2 = \sqrt{\frac{F(n_t + n_c)}{n_t n_c}}$$

If analysis of covariance (ANCOVA) was used and mean square error and pretest–posttest correlation were reported, the standardized mean difference was derived using the typical equation (ES_1 above) with the following formula for calculating pooled *SD*:[3]

$$SD_p = \sqrt{\left(\frac{MS_{error}}{1 - r^2}\right)\left(\frac{df_{error} - 1}{df_{error} - 2}\right)}$$

Next, all standardized mean differences were scaled to Hedges *g* to account for the small sample sizes, which can cause overestimation of effects when using the typical formula (Hedges, 1981). Hedges g is the effect size used in all subsequent analyses and reported throughout:

$$g = \left(1 - \frac{3}{4N - 9}\right) ES$$

Standard errors *(SE)* and weights (w) were also calculated in order to use them to weight individual effect sizes to more accurately model the effects in the analysis (Lipsey & Wilson, 2001):

$$SE = \sqrt{\frac{n_t + n_c}{n_t n_c} + \frac{g^2}{2(n_t + n_c)}}$$

$$w = \frac{1}{SE^2} = \frac{2n_t n_c(n_t + n_c)}{2(n_t + n_c)^2 + n_t n_c(g)^2}$$

Analysis

A test of homogeneity was conducted on the group of effect sizes in order to determine if there was significant variability in *ES* distribution, which would indicate that other variables (e.g., descriptive study characteristics) may be used to explain the portions of the variance (Lipsey & Wilson, 2001). For the homogeneity analysis, one effect size per study was included to control for the dependence introduced when a single study contributes more than one effect. The single effect size per study was determined using a two-step process. First, when a study included more than one effect size per dependent measure category (e.g., a study had two or more immediate dependent measures); the lowest effect size was used. This is a more conservative estimate of overall effect size than using the mean or selecting one at random. When a study contributed more than one effect size across dependent measure categories (i.e., a study had an immediate and a delayed/generalization-dependent measure), the dependent measures were averaged. This was done to ensure that results from both types of dependent measures contributed to overall effect size calculations. The results of the homogeneity test indicated that the variability of effect sizes was significantly ($Q = 51.82$, $p = .00$) larger than what would be expected by sampling error alone, therefore making an examination of study characteristics appropriate.

Typically, in meta-analytic research, the next step would be to conduct an analysis of meditating variables to explain the significant variability using Hedges (1982) analog to the ANOVA or Hedges and Olkin's (1985) modified weighted regression. However, neither of these options was possible because the ANOVA analog would be equivalent to running multiple t-tests on small data sets and there were not enough studies contributing effect sizes to conduct the regression. Instead, the approach we elected to use is more akin to a Best Evidence Synthesis (Slavin, 1986): the research questions were addressed through descriptive examination of weighted means and confidence intervals, provided that each category contained at least two effect sizes.

RESULTS

Results are presented in four sections. First, dependent measures are examined. Second, types of interventions utilized are explored. Third, student characteristics are analyzed, and fourth, study specific variables are examined. See Table 1 for detailed information on all studies including in the review. See Table 2 for a listing of mean effect size, *SE*, and confidence intervals for categories analyzed.

Dependent Measures

A total of 17 *ES*[4] were calculated across dependent measure categories (immediate and delayed/generalization). The majority of dependent measures were immediate assessments of science knowledge or skill acquisition ($n = 11$). The mean *ES* across immediate dependent measures was .769 ($SE = .091$). The remaining calculated *ESs* ($n = 6$) were delayed measures of science knowledge/skill acquisition or generalization measures. The mean *ES* across delayed/generalization measures was .435 ($SE = .111$). With the exception of one measure (i.e., a high stakes science assessment given at the end of the year) implemented by Mastropieri et al. (2006), all other dependent measures were experimenter generated.

Interventions

Interventions were broken down into three types: structured inquiry, supplemental mnemonic instruction, and supplemental nonmnemonic instruction. A total of four studies examined the impact of structured inquiry. As a group, the treatment intervention consisted of the following core components: structured hands-on experiments, student collaboration to make and share predictions, teacher formative feedback, and student writing and/or drawing to summarize the findings. The comparison group for all but one study was a textbook focused on direct instruction condition. The Dalton and colleagues (1997) study was different as the experimental condition was structured inquiry, while the comparison group was an unstructured inquiry approach. The mean *ES* for structured inquiry studies was .727 ($SE = .161$).

A total of four studies examined the effect of supplemental mnemonic instruction. All studies implemented treatments focusing on the keyword mnemonic strategy. King-Sears, Mercer, & Sindelar (1992) used only the keyword mnemonic strategy, whereas Mastropieri, Scruggs, & Levin (1985), Mastropieri, Scruggs, & Levin (1986), and Scruggs, Mastropieri, Levin, & Gaffney (1985) used pegwords in addition to keywords. In the keyword mnemonic strategy, the keyword refers to a chosen word to represent a given vocabulary word. The keyword mnemonic strategy pairs the keyword with an illustration demonstrating how the keyword is related to the vocabulary word. The keyword and illustration are used as a prompt for students to redevelop the correct definition of the represented vocabulary word (Atkinson, 1975). Additionally, the keyword technique has been paired with the pegwords (Mastropieri, Scruggs, & Levin, 1985; Mastropieri, Scruggs, & Levin, 1986; and Scruggs, Mastropieri, Levin, &

TABLE 1
Effect Sizes and Characteristics of Qualifying Science Studies

Article	*Participants*	*Sample Grades*	*Assignment*	*Intervention Description*	*Dependent Variables*	*Effect Size*
Bay, Staver, Bryan, & Hale (1992)	107	4–6	Randomly assigned to conditions	Instruction on displacement consisted of four sessions of 40–60 minutes each. Both conditions followed the same sequence. In the direct instruction condition (control condition), (a) students were provided an advanced organizer, (b) teacher demonstrated/ modeled concepts, and (c) guided/ independent practice provided an opportunity to complete worksheets covering core science concepts. In the discovery teaching condition, students (a) used hands-on material, (b) discussed relationships among concepts, (c) conducted experiments, and (d) made and tested predictions.	Immediate post-test 20 questions (true and false, multiple choice, and matching)	1.168
	10 with LD/6 in control				Delayed post-test Same as immediate post-test administered 2 weeks later	1.493
	Specific IQ information not provided				Generalization test Two weeks later Conducted another experiment	2.238
Dalton, Morocco, Tivnan, & Rawson Mead (1997)	172	4	Teachers assigned nonrandomly to conditions	Instruction on circuits/electricity was conducted in 12 sessions of 50 minutes each. In the activities-based approach (control condition), students (a) used hands-on science material, (b) completed science data sheets, (c) worked collaboratively with peers, and (d) followed guidelines/procedures	Two immediate post-tests	

TABLE 1
Continued

Article	*Participants*	*Sample Grades*	*Assignment*	*Intervention Description*	*Dependent Variables*	*Effect Size*
				for inquiry aspects of lesson. The Supported Inquiry Science Method (SIS) (a) focused on unifying concepts throughout lessons, (b) emphasized the role of misconceptions, (c) explicitly asked students to share their predictions, and (d) emphasized drawing as a way to represent their thinking.		
	33 with LD/19 in control				8 item paper and pencil test—students used writing and drawing to answer questions on circuits	1.298
	Specific IQ information not provided				8 question diagram test—examined figures to determine if bulb would light. Answered yes or no and provided rationale for answer	1.309
King-Sears, Mercer, & Sindelar (1992)	37	6–8	Teachers assigned at random to conditions	48 science vocabulary words were taught over 4 week period. In the systematic teaching condition (control condition), (a) teacher presented words, (b) students said the word aloud, and (c) teacher prompted them to think of the picture. The imposed keyword condition (treatment 1) used a mnemonic keyword strategy where the teacher provided the mnemonic. The induced keyword condition (treatment 2) used a mnemonic keyword strategy that used a student provided mnemonic after students were instructed on how to make up mnemonics on their own.	Immediate post-test	
	30 with LD				Free recall of definition and matching	1.736

TABLE 1
Continued

Article	Participants	Sample Grades	Assignment	Intervention Description	Dependent Variables	Effect Size
	IQ range of all students 71–120				(Effect size reported for week 4 only and is in favor of imposed keyword over the other conditions.)	
Mastropieri, Scruggs, & Levin (1985)	99	9	Students assigned at random to conditions	Students were instructed on the hardness of 14 minerals individually for 21 minutes (14 minutes spent with the experimenter and 7 minutes spent studying.) In the free-study condition (control condition), (a) teacher presented lesson on minerals for 5 minute lesson and then (b) free studying. In the mnemonic condition, pegword pictures and keyword mnemonics were used.	Immediate post-test	
	All with LD				Recall of mineral hardness	2.266
	Mean IQ 95					
Mastropieri, Scruggs, & Levin (1986)	56	High School	Students assigned at random to conditions	Students were instructed on mineral hardness levels of 14 minerals in groups for 20 minutes. In the direct instruction (control condition), the teacher modeled, prompted, and checked for understanding. The mnemonic instruction condition used a combination of keyword–pegword techniques.	Immediate post-test	
	All with LD				Recall of the mineral hardness level	2.553
	Specific IQ information not provided					
Mastropieri et al. (1998)	75	4	Teacher assigned nonrandomly to conditions	An instructional unit on ecosystems was taught 3 days per week over 7 weeks. The textbook condition (control condition) included: (a) teacher presentation, (b) individual and group work, (c) class discussions,	Two immediate post-tests	

TABLE 1
Continued

Article	Participants	Sample Grades	Assignment	Intervention Description	Dependent Variables	Effect Size
				(d) videotapes, and (d) science textbook and questions. The activities/inclusion condition used (a) curriculum/material that accommodated students with disabilities, (b) small group work, and (c) individual activity booklets to record/draw observations.		
	Five with disabilities in treatment/two with LD				20 point multiple choice test	.613
	Mean IQ of all students with disabilities 79.5				10 item performance test to measure conceptual understanding	2.315
Mastropieri et al. (2006)	213	8	Classes matched and then assigned randomly	The entire study was conducted over 12 weeks covering a unit on scientific investigations. The typical classroom instruction (control condition) used (a) teacher lecture/class notes, (b) laboratory experiments, and (c) worksheets. The peer-assisted learning condition used (a) the same teacher presentation as control and (b) replaced worksheet time with peer-assisted learning with differentiated science activities.	Immediate post-test	
	44 with disabilities/37 students with LD		Each lead teacher taught at least one of each condition		Science content	.331
	IQ average 97.59				Generalization test	
					End of the year high stakes science test	.277
McCleery and Tindal (1999)	57	6	Groups of students assigned nonrandomly	Instruction on the scientific method was conducted in 2–3 sessions of 90 minutes per week for 6 weeks. Hands-on inquiry instruction (control condition) included:	Two immediate post-tests	

TABLE 1
Continued

Article	Participants	Sample Grades	Assignment	Intervention Description	Dependent Variables	Effect Size
				(a) brief introduction of topic, (b) teacher-created experiment, (c) recording data, and (d) additional homework if time permitted over similar content. The pullout instruction treatment included five sessions of 40 minute sessions where (a) explicit examples and nonexamples were taught, (b) rules to guide scientific inquiry were taught, and (c) an outline was provided.		
	14 with LD/4 in control				20 point multiple choice test	1.401
	4 received supplemental instruction				10 item performance test to measure conceptual understanding	2.242
Scruggs, Mastropieri, Bakken, & Brigham (1993)	26	7–8	Cross-over design. Students experienced both conditions.	Instruction on magnetism/electricity and rocks/minerals was conducted in three sessions of 45–60 minutes each. The textbook-based treatment (control condition) included (a) daily review, (b) teacher presentation of material/concepts, (c) guided, and (d) independent practice activities. The activities-based treatment consisted of (a) small-group hands-on activities using Full Option Science System (FOSS) kits, (b) daily review, and (c) instructor provided coaching during activities.	Immediate post-test	
	All with LD				Free recall and then 24 questions that included 8 factual, 8 vocabulary, and 8 application	.444
	Mean IQ 88.81				Delayed post-test	
					1 week later same measure as immediate recall	.469

TABLE 1
Continued

Article	Participants	Sample Grades	Assignment	Intervention Description	Dependent Variables	Effect Size
Scruggs, Mastropieri, Levin, & Gaffney (1985)	56	7–8	Students stratified by grade level and randomly assigned.	Instruction focused on eight minerals and their associated attributes. The time on task was the same (11 minutes). In the free-study condition, students were given a list and were provided time to study. In the mnemonic condition (treatment 1), pegword pictures and keyword mnemonics were used. The direct instruction condition (treatment 2) taught students minerals using a model, prompt, and check procedure.	Immediate post-test	
	All with LD				Recall test after 90 second filler task. (Effect size reported for mnemonic vs. free study only. Mnemonic also significantly higher than direct instruction.)	1.574
	Mean IQ 95					
*Scruggs, Mastropieri, & Sullivan (1994)	36	4–5	Students assigned randomly to one of three conditions	One session of approximately 15 minutes was taught. Each condition spent 40 seconds on each animal fact. In the control condition, teacher (a) taught and repeated 14 facts on animals (e.g., some frogs lay eggs that sink), (b) asked student to remember it, and (c) repeat it. In the experimenter-provided condition (treatment 1), the student was taught the fact as in control condition, but was also provided an explanation (e.g., frog eggs that sink to the bottom are safe, and they will be harder for other animals to find and eat). In the student-generated	Two immediate post-tests	

TABLE 1
Continued

Article	Participants	Sample Grades	Assignment	Intervention Description	Dependent Variables	Effect Size
				explanation (treatment 2), the same procedure as the experimenter condition except the student asked the questioned until they arrived at the explanation on their own.		
	28 with LD Mean IQ 93.3				Factual Recall:	
					-Experimenter provided versus control	.492
					Student generated	
					-Student generated versus control	.672
					Causal explanation:	
					-Experimenter provided versus control	4.582
					-Student generated versus control	4.871
					Two delayed post-tests—1 week later. Factual Recall:	
					-Experimenter provided versus control	.311
					Student generated	
					-Student generated versus control	.669
					Causal explanation:	
					-Experimenter provided versus control	1.531
					-Student generated versus control	1.624

Note: *ES* = effect size; LD = learning disabilities; * = study contributed two unique effect size calculations as there was no overlap between students in condition and *ES* could be calculated for each only against the control condition.

Gaffney, 1985). The pegwords represent a numerical value associated with the initial vocabulary word. The pictures rhyme with numbers; therefore, the student is prompted by the pegword to remember the corresponding number. In the three studies, the numbers represented the minerals' hardness level. The comparison group for all studies consisted of free study of the target words or direct instruction. The mean *ES* for supplemental mnemonic instruction studies was 1.997 (*SE* = .218).

A total of four studies examined the effect of supplemental nonmnemonic instruction. The intervention aggregated in this category includes peer-assisted learning (Mastropieri et al., 2006), pullout explicit instruction method (i.e., direct teaching of rules to guide scientific inquiry and use of specific examples and nonexamples) on science concepts and the scientific method (detailed instruction on specific material) (McCleery & Tindal, 1999), and teacher-provided or student-generated explanation (e.g., a frog's eggs sink so they cannot be eaten by predators) for science facts (Scruggs and Mastropieri, 1994). The comparison group for these studies was typical classroom instruction. The mean *ES* for supplemental nonmnemonic instruction studies was .422 (*SE* = .123).

Along with the type of intervention, we also examined the impact of intervention length on *ES*. Across studies, the length of intervention ranged from 1 to approximately 60 sessions. Studies ($n = 7$) with a length of four sessions or fewer obtained an average *ES* of 1.128 (*SE* = .15), while studies ($n = 5$) with a length of 12 or more sessions obtained an average *ES* of .59 (*SE* = .111). Due to the large mean *ES* obtained

TABLE 2
Mean Effect Sizes, Standard Errors, and Confidence Intervals for Categories Analyzed

Category Designation	*Mean ES (SD)*	*95% Confidence Interval* *Low*	*High*
Overall	.780	.605	.955
Dependent measures			
Immediate	.769(.091)	.591	.947
Delayed/Generalization	.435(.111)	.218	.653
Intervention			
Inquiry	.727(.161)	.412	1.042
Supplemental mnemonic	1.997(.218)	1.569	2.424
Supplemental nonmnemonic	.422(.123)	.181	.664
Subjects—grade level			
Third through sixth grade	.825(.153)	.524	1.126
7th through 12th grade	.757(.110)	.542	.973
Subjects-–IQ reported for students			
IQ report	.681(.095)	.494	.868
IQ not reported	1.478(.253)	.981	1.974
Study specific characteristic-–random assignment			
Randomly assigned	.787(.109)	.574	1.00
Not randomly assigned	.767(.156)	.460	1.073
Study specific characteristic-–results for students with LD reported separately			
Reported separately	1.346(.16)	1.033	1.660
Not reported separately	.525(.108)	.314	.735

Note: ES = effect size, *SD* = standard error

by supplemental mnemonic instruction studies, we also examined the *ES* difference for intervention length without including these studies. After excluding mnemonic studies, studies ($n = 4$) with a length of four sessions or fewer obtained an average *ES* of .565 (*SE* = .19), while studies ($n = 4$) with a length of 12 or more sessions obtained an average *ES* of .525 (*SE* = .114).

Student Characteristics

No studies included students below third grade. Studies were therefore coded[5] as including students in third through sixth grade or 7th through 12th grade. Studies ($n = 6$) with subjects in third through sixth grade had a mean *ES* of .825 (*SE* = .153); studies ($n = 6$) with subjects in 7th through 12th grade had a mean *ES* of .757 (*SE* = .110). We also examined the *ES* difference without including mnemonic studies. Studies ($n = 6$) including students in third through sixth grade obtained an average *ES* of .825 (*SE* = .153), while studies ($n = 2$) including students in 7th through 12th grade obtained an *ES* of .336 (*SE* = .127).

Besides grade-level information, student demographic reporting across studies was not consistent enough to allow aggregation and comparison. Therefore, studies were coded only as providing or not providing IQ information for students. This analysis was completed because Swanson and Hoskyn (1998) and Swanson (1999) found a significant *ES* difference in favor of studies that did not provide IQ data in large-scale LD meta-analyses. In our analysis, studies ($n = 4$) that did not include any IQ information for their subjects obtained an average *ES* of 1.478 (*SE* = .253), while studies ($n = 8$) that did include IQ information obtained an average *ES* of .681 (*SE* = .095). We also examined the *ES* difference without including mnemonic studies. Studies ($n = 3$) that did not report IQ information obtained an average *ES* of 1.337 (*SE* = .269), while studies ($n = 5$) that did report IQ information obtained an *ES* of .413 (*SE* = .105).

Study-Specific Variables

Two study-specific variables were examined: random assignment and whether the achievement of students with LD was reported separately from the remaining subject population. Studies ($n = 8$) that assigned students or teachers randomly to conditions obtained an average *ES* of .787 (*SE* = .109), while studies ($n = 4$) that did not assign students or teachers randomly obtained an average *ES* of .767 (*SE* = .156). We also examined the *ES* difference without including mnemonic studies. Studies ($n = 4$) that assigned teachers or students at random obtained an average *ES* of .386 (*SE* = .125), while studies ($n = 4$) that not assign randomly obtained an *ES* of .767 (*SE* = .156).

Studies varied in whether they reported students with LD results separately or only in combination with other subjects without LD included in the research. Whenever LD-specific information was available, *ESs* were calculated using only students with LD data. Studies ($n = 6$) that did not report results separately for students with LD obtained an average *ES* of .525 (*SE* = .108), while studies that did report results separately for students with LD obtained an average *ES* of 1.346 (*SE* = .16). The difference was similar after excluding mnemonic studies. Studies ($n = 5$) that did not report results separately for students with LD obtained an average *ES* of .460 (*SE* = .110), while studies ($n = 3$) that did report results separately for students with LD obtained an average *ES* of .810 (*SE* = .212)

DISCUSSION

The purpose of this article was to examine the classroom-based science instruction for students with LD. A total of 12 articles published between 1980 and 2010 were reviewed. Across all studies, a mean *ES* of .78 was obtained indicating a moderate positive effect [small effect *ES* = below .50, medium effect *ES* = .50–.80, and large effect *ES* = above .80 (Cohen, 1988)] on students with LD science achievement. To examine this overall effect in detail, we coded and analyzed studies in the following categories: dependent measures, interventions, student characteristics, and study-specific variables. Below, we describe the results of these analyses.

Dependent Measures

The majority (i.e., 65 percent) of dependent measures were immediate assessments of science knowledge or skill acquisition. Common measures in this category include paper and pencil tests on science concepts and factual recall. The remaining measures (i.e., 35 percent) assessed students' information retention after a delay or measured their achievement on a generalization task. Typically, the delay between intervention and assessment was relatively short (i.e., 1–2 weeks) and the delay measures paralleled the immediate measures (e.g., factual recall) in form and content. Only two studies (Bay, Staver, Bryan, & Hale, 1992; Mastropieri et al., 2006) included a generalization measure. As would be expected, mean student achievement was higher ($ES = .769$) on immediate measures compared to delayed/generalization measures ($ES = .435$).

Previous LD meta-analyses (Swanson, 1999; Swanson & Hoskyn, 1998) found that mean *ESs* reported for experimentally generated dependent measures tend to be significantly higher than those reported on standardized achievement measures. All but one of the measures utilized in the studies reviewed was experimenter generated. Therefore, although researcher-generated assessments are appropriate for establishing content validity, caution should be used when attempting to generalize results of this analysis to standard science achievement measures as the interventions are unlikely to have as strong of an impact on these distal measures.

Interventions

Interventions were broken down into three categories: structured inquiry, supplemental mnemonics instruction, and supplemental nonmnemonics instruction. Previous literature reviews found that students with disabilities can be successful within an inquiry-based approach (Mastropieri & Scruggs, 1992; Mastropieri et al., 1998). Results from our analysis echo this conclusion for students with LD. Studies that evaluated inquiry instruction for students with LD obtained an overall *ES* of .727, indicating a moderate to large impact on achievement. What this finding means, however, for classroom science instruction for students with LD is hard to determine because there is great variability in approaches-labeled inquiry ranging from a pure discovery approach to a structured teacher-directed approach (Scruggs & Mastropieri, 2007).

Although it is difficult to quantify the level of structure within the inquiry approaches reviewed in this analysis, it is safe to conclude that all provided significantly more structure than is seen in a pure discovery learning approach. Furthermore, two studies reviewed compared levels of structure in inquiry methods and found results in favor of a more structured approach. Dalton and colleagues (1997) compared a general inquiry approach to structured inquiry and reported significantly better achievement for students in the structured inquiry classrooms. While McCleery and Tindal (1999) provided pull-out explicit instruction to supplement an inquiry approach and reported significantly better achievement. Results of our analysis for students with LD therefore align with reviews of inquiry-based instruction for students with special needs, in general, indicating the need for structure within inquiry approaches in order for students with LD to be successful.

Previous literature reviews (Mastropieri & Scruggs, 1992; Mastropieri et al., 1998) found that instructional components/adaptations that are beneficial for students with disabilities within a structured inquiry approach include preteaching, reducing language, and literacy demands, providing hands-on experiences with extra supports, giving formative feedback, providing additional practice, and reviewing key concepts. Our analysis indicates that these components also appear to be effective for students with LD. Inquiry studies reviewed in our analysis provided hands-on experience under the direction of the teacher, focused on overall concepts instead of extraneous facts, provided the formative feedback, allowed the students to demonstrate understanding in a variety of formats (e.g., orally, pictures, and writing), and provided the additional supports for students with LD to ensure that they were appropriately engaged in the lessons.

The majority of studies (8 of 12) examined the impact of supplemental instructional supports/interventions for student with LD in science. Half of these studies evaluated the effectiveness of keyword or pegword and keyword mnemonics on students' acquisition and/or retention of science factual knowledge. The results of these studies were unequivocally positive ($ES = 1.997$), indicating that mnemonic instruction is highly effective at increasing students with LD knowledge of science vocabulary.

Along with mnemonic instruction, other supplemental programs examined included peer-assisted learning with tiered material geared toward students' needs (Mastropieri et al., 2006), supplemental explicit instruction (McCleery & Tindal, 1999), and providing or having students generate an explanation to increase the acquisition of science facts (Mastropieri & Sullivan, 1994). Although the combined *ES* magnitude for this category ($ES = .422$) was smaller than the other categories, all studies reported significant findings with Mastropieri and colleagues (2006) reporting an increase ($ES = .277$) on a distal-dependent measure (i.e., student achievement at the end of the year high stakes science test). Additional research is needed to further explore these supplemental interventions, but all appear to have potential for improving students with LD achievement in science.

In the intervention category, we also examined the length of instruction. The vast majority of studies were short in duration with a mean intervention length of 11.6 sessions and a mode of only 3.5 sessions. Mean *ES* (1.128) was larger for shorter duration studies (i.e., fewer than four sessions) than the mean *ES* (.59) for longer duration studies (greater than 12 sessions). This difference, however, is likely an artifact of the following: (a) all but one of the mnemonic studies with extremely high *ESs* were short in duration, and (b) short duration studies tended to implement dependent measures (e.g., recall of facts taught) that were directly aligned with instruction. Comparison of intervention length after excluding mnemonic instruction indicates no difference ($ES = .565$ for shorter duration studies and $ES = .525$ for longer duration studies).

In order to assess the impact of science intervention on students with LD overall science achievement, future research should include the studies of longer duration. This is especially critical to determine the impact of core science instructional approaches such as structured inquiry as the students with LD receive the vast majority of their science instruction in general education classrooms (Cawley, Hayden, Cade, &

Baker-Kroczynski, 2002) where these approaches are often utilized.

Students

Students in grades 3rd through 12th were participants in the studies examined. When including mnemonic instruction studies in the calculation, *ESs* were equivalent across age ranges (*ES* = .825 for third through sixth grade and *ES* = .757 for 7th through 12th grade). However, when mnemonic instruction was excluded from the analysis, mean *ES* was larger for younger students (*ES* = .825) than older students (*ES* = .336). Therefore, similar to other areas of instruction such as reading (Roberts, Torgesen, Boardman, & Scammacca, 2008), it appears that science achievement is more difficult to impact as students with LD get older.

Besides grade level, student demographic information reported across studies was not consistent enough to allow aggregation and analysis. We did compare whether studies provided information on participants IQs. We did this for two reasons. First, previous LD meta-analyses (Swanson, 1999; Swanson & Hoskyn, 1998) found a significant difference between studies that did and did not provide this information with larger mean *ES* reported for studies that did not report IQ data. Second, Scruggs and Mastropieri (2007) found that IQ was strongly predictive of students' ability to draw relevant inductive conclusions, an essential skill to do well in science instruction, particularly inquiry science instruction. Our findings paralleled the results from previous LD meta-analyses with a higher mean *ES* for studies that did not provide IQ data (1.478) and lower mean *ES* (.681) when IQ data were provided. This difference was also present after excluding mnemonic studies (no IQ data *ES* = 1.337, IQ data reported *ES* = .413). The dramatic difference between studies based on this simple dichotomous variable indicates that student characteristics (e.g., reading achievement, IQ) likely play a significant moderating role in determining *ES* magnitude.

Study-Specific Variables

Two study-specific variables were examined: random assignment and whether achievement for students with LD was reported separately from other students involved in the investigation. No difference was found between studies that did and did not assign teachers/students randomly. There was, however, a sizable difference in favor of studies that provided enough information to calculate *ESs* using only students with LD. The difference was evident both when including and excluding the mnemonic studies from the analysis. This dramatic difference is surprising because it indicates that the students with LD maybe more responsive to instruction than students without LD. Future studies should separate data for students with LD so that *ES* can be calculated for both the student population as a whole and for students with LD separately.

Limitations

There are three limitations to the conclusions of this analysis. First, although we included studies published during a 30-year window, only 12 studies met our criteria to be included in the analysis making finite comparisons among and between study variables difficult. Second, the majority of dependent measures used in the studies examined were experimentally generated immediate recall measures. This type of measure tends to be more reactive to intervention, and thus, may have resulted in larger effect sizes than would be seen on typical science achievement measures. Third, it was not always possible to calculate *ES* for students with LD separately from the rest of the study population. Although an examination of this variable indicated a significant *ES* difference in favor of studies where *ESs* were calculated only on students with LD, students with LD achievement in the interventions where the populations were not separated may have been masked by the overall means reported.

Implications for Practice

There are two practical implications from this analysis. First, results indicate that students with LD can be successful within inquiry-based instructional approaches, provided that the approach is structured (i.e., not pure discovery learning). Although additional research is needed to discern exactly what instructional components maximize success for students with LD within inquiry instruction, it appears that students with LD likely benefit from (a) a focus on overall concepts/ big ideas, (b) hands-on concrete experiences conducted under the direction of the teacher, (c) formative feedback, (d) behavioral supports to ensure that students are actively and appropriately engaged in the lesson, and (e) additional practice and review of core concepts and vocabulary. Second, the results from this analysis indicate that supplemental instruction (e.g., mnemonics and peer-assisted learning strategies) provided in conjunction with the regular education science curriculum can significantly improve the students with LD science achievement.

CONCLUSION

Overall results from this analysis indicate that classroom-based science instruction (e.g., structured inquiry and supplemental instruction) is effective at increasing students with LD achievement on immediately administered science assessments that are closely aligned to the intervention provided. Keyword and combined pegword/keyword mnemonics are particularly powerful at improving students' acquisition and retention of science facts. Along with proximal measures, future studies should include standardized science achievement assessments and generalization measures so that we may refine our understanding of the effect of these programs and interventions. Furthermore, future studies should expand and refine our understanding of various instructional approaches for students with LD in science. Particular attention should be paid to evaluating the instructional components needed for students with LD to be successful in inquiry-based classrooms.

ACKNOWLEDGMENTS

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grants R305A090094 and R305B10005 to The University of Iowa. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education.

NOTES

1. Only classroom-based interventions were reviewed in this article because the Mason & Hedin's article in this issue addressed science-text-based interventions.
2. Although there were only 11 articles in the review, the study conducted by Scruggs, Mastropieri, and Sullivan (1994) could be analyzed as two studies because it included two unique intervention groups that could be compared to the control.
3. Missing correlations were estimated using a conservative estimate of .7 as higher correlations lead to lower effect sizes and all known correlations were in the .4–.5 range.
4. Unlike for the remaining analyses that only included one effect size per comparison for the dependent measure review, a total of 17 effect sizes were available as we were comparing across dependent measure categories.
5. King-Sears, Mercer, & Sindelar (1992) were coded as a 7th through 12th grade as it included more students in grades 7th and 8th than it did students in sixth grade.

REFERENCES

American Association for the Advancement of Science. (1997). *Resources for science literacy: Professional development*. New York: Oxford University Press.

Anderman, E. M. (1998). The middle school experience: Effects on the math and science achievement of adolescents with LD. *Journal of Learning Disabilities*, *32*, 128–138.

Atkinson, R. C. (1975). Mnemotechnics in second-language learning. *American Psychologist*, *30*, 821–828.

Bakken, J. P., & Mastropieri, M. A. (1997). Reading comprehension of expository science material and students with learning disabilities: A comparison of strategies. *Journal of Special Education*, *31*, 300–324.

Bay, M., Staver, J. R., Bryan, T., & Hale, J. B. (1992). Science instruction for the mildly handicapped: Direct instruction versus discovery teaching. *Journal of Research in Science Teaching*, *29*, 555–570.

Carlisle, J. F., & Chang, V. (1996). Evaluation of academic capabilities in science by students with and without learning disabilities. *Journal of Special Education*, *30*, 18–34.

Cawley, J. F. (1994). Science for students with disabilities. *Remedial and Special Education, 15*, 667–71.

Cawley, J., Hayden, S., Cade, E., & Baker-Kroczynski, S. (2002). Including students with disabilities into the general science classroom. *Exceptional Children*, *68*, 423–435.

Cawley, J. F., Foley, T. E., & Miller, J. (2003). Science and students with mild disabilities: Principles of universal design. *Intervention in School and Clinic*, *38*(3), 160–171.

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

Dalton, B., Morocco, C. C., Tivnan, T., & Rawson Mead, P. L. (1997). Supported inquiry science: Teaching for conceptual change in urban and suburban science classrooms. *Journal of Learning Disabilities*, *30*, 670–684.

Donovan, S., & Bransford, J. (2005). *How students learn: History, mathematics, and science in the classroom*. Washington, DC: National Academies Press.

Espin, C. A., & Deno, S. L. (1993). Content-specific and general reading disabilities of secondary-level students: Identification and educational relevance. *Journal of Special Education*, *27*, 321–337.

Fleischman, H. L., Hopstock, P. J., Pelczar, M. P., & Shelley, B. E. (2010). *Highlights from PISA 2009: Performance of U.S. 15-year old students in reading, mathematics, and science literacy in an international context* (NCES 2011-004). Retrieved April 3, 2011, from http://nces.ed.gov/pubs2011/2011004.pdf

Gaddy, S. A., Bakken, J. P., & Fulk, B. M. (2008). The effects of teaching text-structure strategies to postsecondary students with learning disabilities to improve their reading comprehension on expository science text passages. *Journal of Postsecondary Education and Disability*, *20*, 100–119.

Grigg, W., Lauko, M., & Brockway, D. (2006). *The Nation's report card: Science 2005* (NCES 2006 466). Retrieved April 4, 2011, from http://nces.ed.gov/nationsreportcard/pdf/main2005/2006466.pdf

Hedges, L. V. (1981). Distribution theory for glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*, 107–128.

Hedges, L. V. (1982). Fitting categorical models to effect sizes from a series of experiments. *Journal of Educational Statistics*, *7*, 119–137.

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

Holahan, G. G., McFarland, J., & Piccillo, B. A. (1994) Elementary school science for students with disabilities. *Remedial and Special Education*, *15*, 86–93.

King-Sears, M. E., Mercer, C. D., & Sindelar, P. T. (1992). Toward independence with keyword mnemonics: A strategy for science vocabulary instruction. *Remedial and Special Education*, *13*(5), 22–33.

Kooy, T. (1992). The effect of graphic advance organizers on the math and science comprehension with high school special education students. *B.C. Journal of Special Education Students*, *16*, 101–111.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis* (Vol. 49). Thousand Oaks, CA: Sage.

MacDougall, A., Schnur, R., Berger, C, & Vernon, D. (1981). The use of activity-centered science activities to facilitate the mainstreaming of elementary school children with special needs. *Science Education*, *65*, 467–475.

Marino, M. T., Coyne, M., & Dunn, M. (2010). The effect of technology-based altered readability levels on struggling readers' science comprehension. *Journal of Computers in Mathematics and Science Teaching*, *29*, 31–49.

Maroney, S. A., Finson, K. D., Beaver, J. B., & Jensen, M. M. (2003). Preparing for successful inquiry in inclusive science classrooms. *Teaching Exceptional Children*, *36*, 18–25.

Mastropieri, M. A., & Scruggs, T. E. (1992). Science for students with disabilities. *Review of Educational Research*, *62*, 377–411.

Mastropieri, M. A., Scruggs, T. E., & Levin, J. R. (1985). Mnemonic strategy instruction with learning disabled adolescents. *Journal of Learning Disabilities*, *18*, 94–100.

Mastropieri, M. A., Scruggs, T. E., & Levin, J. R. (1986). Direct vs. mnemonic instruction: Relative benefits for exceptional learners. *Journal of Special Education*, *20*, 299–308.

Mastropieri, M. A., Scruggs, T. E., & Magnusen, M. (1999). Activities-oriented science instruction for students with disabilities. *Learning Disability Quarterly*, *22*, 240–249.

Mastropieri, M. A., Scruggs, T. E., Mantzicopoulous, P., Sturgeon, A., Goodwin, L., & Chung, S. (1998). "A place where living things affect and depend on each other": Qualitative and quantitative outcomes associated with inclusive science teaching. *Science Education*, *82*, 163–179.

Mastropieri, M. A., Scruggs, T. E., Norland, J. J., Berkeley, S., McDuffie, K., Tornquist, E. H., et al. (2006). Differentiated curriculum enhancement in inclusive middle school science: Effects on classroom and high-stakes tests. *Journal of Special Education*, *40*, 130–137.

McCleery, J. A., & Tindal, G. A. (1999). Teaching the scientific method to at-risk students and students with learning disabilities through concept anchoring and explicit instruction. *Remedial and Special Education*, *20*, 7–18.

McFarland, J., & Shepard, T. (1995). Oral and written compositions of students with learning disabilities in the content area of science. *Learning Disabilities: A Multidisciplinary Journal*, *6*, 13–18.

Munk, D. D., Brucket, J., Call, D. T., Stoehrmann, T., & Randandt, E. (1998). Strategies for enhancing the performance of students with LD in inclusive science class. *Intervention in School and Clinic*, *34*, 73–78.

National Center for Education Statistics. (2011). *The Nation's Report Card: Science 2009 (NCES 2011-451)*. Washington, DC: Institute of Education Sciences, U.S. Department of Education. Retrieved April 1, 2011, from http://nces.ed.gov/nationsreportcard/pdf/main2009/2011451.pdf

National Science Teachers Association Board of Directors. (2004). National science teachers association position statement: Science inquiry. Retrieved April 1, 2011, from National Science Teachers Association: Official Positions website: http://www.nsta.org/about/positions/inquiry.aspx

Nelson, J. R., Smith, D. J., & Dodd, J. M. (1992). The effects of teaching a summary skills strategy to students identified as learning disabled on comprehension of science text. *Education and Treatment of Children*, *15*, 228–243.

Olson, J. L., & Platt, J. C. (2004). *Teaching children and adolescents with special needs*. Saddle River, NJ: Merrill.

Palincsar, A. S., Magnusson, S. J., Cutter, J., & Vincent, M. (2002). Supporting guided-inquiry instruction. *Teaching Exceptional Children*, *34*(3), 88–91.

Parmar, R. S., Deluca, D. B., & Janczak, T. M. (1994). Investigations into the relationship between science and language abilities of students with mild disabilities. *Remedial and Special Education*, *15*, 117–126.

Roberts, G., Torgesen, J., Boardman, A., & Scammacca, N. (2008). Evidence-based strategies for reading instruction of older students with learning disabilities. *Learning Disabilities Research & Practice*, *23*, 63–69.

Scruggs, T. E., & Mastropieri, M. A. (1994). The construction of scientific knowledge by students with mild disabilities. *Journal of Special Education*, *28*, 307–321.

Scruggs, T. E., Mastropieri, M. A., Bakken, J. P., & Brigham, F. J. (1993). Reading versus doing: The relative effects of textbook-based and inquiry-oriented approaches to science learning in special education classrooms. *Journal of Special Education*, *27*, 1–15.

Scruggs, T. E., Mastropieri, M. A., & Boon, R. (1998). Science education for students with disabilities: A review of recent research. *Studies in Science Education*, *32*, 21–44.

Scruggs, T. E., Mastropieri, M. A., Levin, J. R., & Gaffney, J. S. (1985). Facilitating the acquisition of science facts in learning disabled students. *American Educational Research Journal*, *22*, 575–586.

Scruggs, T. E., Mastropieri, M. A., & Sullivan, S. (1994). Promoting relational thinking: Elaborative interrogation for students with mild disabilities. *Exceptional Children*, *60*, 450–457.

Scruggs, T. E., & Mastropieri, M. A. (2007). Science learning in special education: The case for constructed versus instructed learning. *Exceptionality*, *15*, 57–74.

Shepard, T., & Adjogah, S. (1994). Science performance of students with learning disabilities on language-based measures. *Learning Disabilities Research & Practice*, *9*, 219–225.

Slavin, R. E. (1986). Best-evidence synthesis: An alternative to meta-analytic and traditional reviews. *Educational Researcher*, *15*, 5–11. doi: 10.3102/0013189x015009005

Steele, M. (2004). Teaching science to middle school students with learning problems. *Preventing School Failure*, *49*(1), 19–22.

Swanson, H. L. (1999). Instructional components that predict treatment outcomes for students with learning disabilities: Support for a combined strategy and direct instruction model. *Learning Disabilities Research & Practice*, *14*, 129–140.

Swanson, H. L., & Hoskyn, M. (1998). Experimental intervention research on students with learning disabilities: A meta-analysis of treatment outcomes. *Review of Educational Research*, *68*, 277–321.

Turner, S. (2008). School science and its controversies: Or, whatever happened to scientific literacy. *Public Understanding of Science*, *17*(1), 55–72.

Villanueva, M. G. (2010). *Integrated teaching strategies model for improved scientific literacy in second-language learners*. Unpublished doctoral dissertation. Nelson Mandela Metropolitan University, Port Elizabeth, South Africa.

White House, Office of the Press Secretary. (2011, January 25). *Remarks by the president in the state of union address*. Retrieved from http://www.whitehouse.gov/the-press-office/2011/01/25/remarks-president-state-union-address

About the Authors

William J. Therrien is an associate professor and co-director of the Center for Disability Research and Education at the University of Iowa. His current research interests include science and reading instruction for students with learning disabilities.

Jonte C. Taylor is a postdoctoral scholar at the University of Iowa. His research interests include science instruction for students with mild cognitive disabilities and task engagement.

John Hosp is an associate professor and co-director of the Center for Disability Research and Education at the University of Iowa. His research interests include aligning assessment and instructional decision making, curriculum-based measurement, and the disproportionate representation of individuals of color in special education programs.

Erica R. Kaldenberg is a doctoral student at the University of Iowa. Her research interests include science instruction for students with disabilities and the effective use of paraprofessionals in instructional settings.

Jay Gorsh is a doctoral student at the University of Iowa. His research interests include science instruction and co-teaching.