International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: http://www.tandfonline.com/loi/tsed20

Teaching High School Biology Students to Coordinate Text and Diagrams: Relations with Transfer, Effort, and Spatial Skill

Bradley W. Bergey, Jennifer G. Cromley & Nora S. Newcombe

To cite this article: Bradley W. Bergey, Jennifer G. Cromley & Nora S. Newcombe (2015) Teaching High School Biology Students to Coordinate Text and Diagrams: Relations with Transfer, Effort, and Spatial Skill, International Journal of Science Education, 37:15, 2476-2502, DOI: 10.1080/09500693.2015.1082672

To link to this article: http://dx.doi.org/10.1080/09500693.2015.1082672

Published online: 13 Oct 2015.

International Journal of Science Education, 2015
Vol. 37, No. 15, 2476–2502, http://dx.doi.org/10.1080/09500693.2015.1082672

Routledge
Taylor & Francis Group

Teaching High School Biology Students to Coordinate Text and Diagrams: Relations with Transfer, Effort, and Spatial Skill

Bradley W. Bergey[a*], Jennifer G. Cromley[b] and Nora S. Newcombe[c]
[a]*Department of Psychology and Neuroscience, Life Science Centre, Dalhousie University, Halifax, Canada;* [b]*Cognitive Science of Teaching and Learning, University of Illinois at Urbana-Champaign, Champaign, IL, USA;* [c]*Department of Psychology, Temple University, Philadelphia, PA, USA*

There is growing evidence that targeted instruction can improve diagram comprehension, yet one of the skills identified in the diagram comprehension literature—coordinating multiple representations—has rarely been directly taught to students and tested as a classroom intervention. We created a Coordinating Multiple Representation (CMR) intervention that was an addition to an intervention focused on Conventions of Diagrams (COD) and tested their joint effects on diagram comprehension for near transfer (uninstructed biology diagrams), far transfer (uninstructed geology diagrams), and content learning (biology knowledge). The comparison group received instruction using a previously validated intervention that focused exclusively on COD. Participants were 9th–10th grade biology students ($N = 158$ from two schools), whose classes were randomly assigned to COD alone or COD + CMR conditions and studied with a pretest–posttest experimental design. Both groups showed significant growth in biology knowledge ($d = .30$–$.53$, for COD and COD + CMR, respectively) and biology diagram comprehension ($d = .28$–$.57$). Neither group showed far transfer. Analyses of student work products during the interventions suggest that gains were not simply due to the passage of time, because student effort was correlated with gains in both treatment groups. Directions for improving future CMR interventions are discussed.

Keywords: *Biology education; Visual media; Learning outcomes*

*Corresponding author. Department of Psychology and Neuroscience, Life Science Centre, Dalhousie University, 1355 Oxford Street, P.O. Box 15000, Halifax, Canada NS B3H 4R2. Email: bradley.bergey@dal.ca

There is growing evidence that students have difficulty understanding scientific diagrams (Magner, Schwonke, Aleven, Popescu, & Renkl, 2014; Wu, Lin, & Hsu, 2013) and that targeted instruction can improve diagram comprehension (Cromley et al., 2013a, 2013b; Leopold & Leutner, 2012; Slof, Erkens, Kirschner, & Helms-Lorenz, 2013). However, comprehending isolated representations is often insufficient for understanding a topic or for completing a task. Scientific information from multiple representations must frequently be combined, compared, integrated, related, coordinated, or translated, all of which involve mapping information from one representation onto information from another representation in order to fully comprehend the depicted phenomenon. Therefore, one of the key skills underlying diagram comprehension identified in the literature is coordinating multiple representations (Treagust & Tsui, 2013). Coordinating multiple representations refers to the awareness of when and how to switch between representations when multiple related representations are presented. A facility with coordinating multiple representations is a hallmark of expertise in domains such as biology (Tsui & Treagust, 2013), chemistry (Kozma, Chin, Russell, & Marx, 2000), geology (Kastens, Agrawal, & Liben, 2009), and physics (Chi, Feltovich, & Glaser, 1981), and is critical for much learning in science, technology, engineering, and math domains (Ainsworth, 2006). However, interventions that support the ability to coordinate multiple representations have been tested almost exclusively in laboratory conditions. Intervention research on coordinating multiple representations in ecologically valid classroom settings may support diagram comprehension and topic learning, but this hypothesis needs to be tested empirically.

Training to Improve Diagram Comprehension

A few studies have tested whether instruction can improve diagram comprehension, and have largely found training to be helpful. Laboratory research has demonstrated that adding prompts (e.g. Berthold & Renkl, 2009) or hyperlinks (e.g. Bartholome & Bromme, 2009) to existing diagrams can improve diagram comprehension. In applied studies, researchers have instructed middle and high school students about the conventional features of diagrams (e.g. captions, arrows, and color coding; Bergey, Cromley, Kirchgessner, & Newcombe, 2015; Cromley et al., 2013b), prompted self-explanation with diagrams (Ainsworth & Iacovides, 2005; Cromley et al., 2013a), and used student-constructed drawings as a way to foster diagram comprehension (Leopold & Leutner, 2012; Schmeck, Mayer, Opfermann, Pfeiffer, & Leutner, 2014; Van Meter, Aleksic, Schwartz, & Garner, 2006). One important finding from this small literature is that student effort while learning is important for benefiting from such interventions (Bergey et al., 2015; Cromley et al., 2013a).

Coordinating Multiple Representations and the Design, Functions, and Tasks (DeFT) Model

Coordinating multiple representations is a challenging yet critical skill for the comprehension of illustrated scientific text. Coordinating multiple representations encompasses a set of skills used when a learner moves back and forth between two or more representations of a single phenomenon and attempts to understand both representations together (Ainsworth, 2006). For example, chemists must coordinate multiple representations when they are working with Nuclear Magnetic Resonance spectroscopy, and they coordinate spectroscopy output, chemical diagrams, formulae, and graphs (Kozma et al., 2000). Coordinating multiple representations is a complex skill, in that the two or more representations often have both overlapping and unique information and also may have overlapping and unique cognitive affordances. For example, while a textual representation can direct attention to very specific features of a situation, a diagram is explicit about spatial relations in a way that text is not. As with reading comprehension, the tasks of coordinating multiple representations depend on characteristics of stimuli, tasks, and learners, all of which show complex interactions that yield comprehension (RAND Reading Study Group, 2002). Coordinating multiple representations is particularly challenging because in addition to mastering each representation individually, coordinating representations requires linking partially overlapping representations and sequencing the coordination process.

One common task demanded in reading middle and high school science textbooks is coordinating representations presented in text with representations presented in diagrams and pictures. Textbooks signal readers to coordinate multiple representations through both explicit messages (e.g. when running text refers the reader to a specific figure) and implicit messages (e.g. spatial arrangements of figures and text). Nevertheless, research has found that students often do not pay attention to diagrams while reading in their textbooks (Cromley, Snyder-Hogan, & Luciw-Dubas, 2010), suggesting that students tend to not respond to these messages.

Coordinating multiple representations is a key process in Ainsworth's (2006) Design, Function, and Tasks (DeFT) model of learning with representations. The act of coordinating multiple representations, according to Ainsworth, can be hindered or facilitated by the mutual interaction of *Design, Function, and Task* characteristics. *Design* within the DeFT framework refers to aspects of each representation and each set of representations, such as the extent to which two or more representations present unique and overlapping information and the sequence or spatial arrangement of representations. *Functions* refer to the role of the representations within learning material, such as to present different perspectives on the same phenomenon or to prevent the formation of misconceptions by constraining interpretations. *Tasks* refer to what learners are asked to do with the representations, including the level of task challenge and the role of individual differences, including familiarity with each representation type, topic knowledge, working memory capacity, and spatial skills (Gegenfurtner, Lehtinen, & Säljö, 2011; Khooshabeh & Hegarty, 2010).

In the context of learning biology, coordinating multiple representations is an essential task, and one that often involves coordinating representations that vary in several dimensions (Treagust & Tsui, 2013). Visual representations in biology include a range of types, such as photographs, drawings, maps, diagrams, graphs, tables, equations, and text. These types of visual representations vary in their level of abstraction, from realistic deceptions of natural phenomena to highly abstract symbolic representations (Roth & Pozzer-Ardenghi, 2013; Tsui & Treagust, 2013). These differences in representations affect student learning. For example, differences in the level of abstraction and type of representations of molecules have been associated with differential learning outcomes (Ferk, Vrtacnik, Blejec, & Gril, 2003; Treagust, Chittleborough, & Mamiala, 2003). In addition, visualizations in biology represent phenomena at various grain sizes, from the macroscopic to the sub-microscopic and symbolic (Griffard, 2013; Tsui & Treagust, 2013). Further, biological representations depict a range of types of knowledge, denoting information about biological concepts, processes, and structures. For example, diagrams in biology textbooks commonly represent complex biological actors and events, depicting multiple steps or changes over time, presented in three dimensions, using multiple levels of organization, and with a variety of graphic and alphabetic symbols (Griffard, 2013). These complexities buttress calls for pre-service and in-service biology teacher training to help teachers develop visual literacy and understand how they can assist their students in learning from biological visualizations (Eilam, 2012).

Prior Training Studies on Coordinating Multiple Representations

Several studies, many conducted with undergraduate populations in laboratory settings, have demonstrated that manipulations can effectively support learners' ability to coordinate multiple representations. For example, Bartholomé and Bromme (2009) provided undergraduate psychology students who had low knowledge about a botany topic with either hyperlinks or numbering to connect text segments with the relevant parts of an illustration. Using a 2 × 2 design, they also asked some students to follow a series of steps in learning from the text (e.g. read, summarize, and elaborate), which included specific instructions to link the text and the diagram (prompting). Overall, the numbering intervention was associated with the largest benefits for knowledge, while prompting showed few effects. Bodemer and Faust (2006) examined psychology students learning about a science topic (geothermal heating) using a computer-based drag-and-drop approach to foster coordinating multiple representations skills, and found better performance in the coordination of multiple representations using numbering compared to drag-and-drop scaffolds, but only for low-knowledge learners. In both laboratory studies, participants showed better performance when coordinating multiple representations skills were supported—though the effects of specific supports differed and interacted with individual characteristics (i.e. prior knowledge).

In a series of training studies with middle school students, Ploetzner and colleagues tested a complex CMR intervention for text and diagram comprehension (Schlag & Ploetzner, 2011) and comprehension of animations (Kombartzky, Ploetzner, Schlag, & Metz, 2010; Ploetzner & Schlag, 2013). For both animation and text + diagrams, 6th-grade students were asked to get an overview, make sketches with captions (for animations), underline text and circle important corresponding parts of diagrams, label the diagrams, and write a summary of the scientific topic. Across all three studies, students who received the complex intervention significantly outscored control groups on conceptual knowledge and transfer measures, and all but one sample also outscored control groups on factual knowledge. Thus, there is preliminary evidence that coordinating multiple representations skills can be trained, and that training can result in better understanding of diagrams. However, training of coordinating multiple representations has not been studied in science class settings, across multiple learning sessions, with students' actual learning materials, and has only rarely included school-aged children.

Coordinating Multiple Representations and Spatial Abilities

Spatial abilities are implicated in science learning (Wai, Lubinski, & Benbow, 2009) in general. For example, spatial visualization—'the ability to mentally manipulate, rotate, twist, or invert pictorially presented stimulus objects' (McGee, 1979, p. 893)—has been associated with problem-solving in scientific disciplines, such as chemistry (Carter, Larussa, & Bodner, 1987), earth science, (Black, 2005; Sanchez & Wiley, 2014), engineering (Alias, Black & Grey, 2002), and physics (Pallrand & Seeber, 1984). Learning from diagrams is a particular aspect of science learning that has been found to be correlated with spatial skills (Hegarty & Kriz, 2008; Höffler, 2010). Higher spatial skills have sometimes been found to be associated with differential learning from scaffolds designed to improve learning from visual representations (Bergey et al., 2015; Cromley et al., 2013a, 2013b; Hinze et al., 2013; Höffler & Leutner, 2011; Lee & Shin, 2011, 2012; Nguyen, Nelson, & Wilson, 2012), although not always (Imhof, Scheiter, & Gerjets, 2011; Ploetzner & Schlag, 2013). Based on these findings, we analyzed whether effects of each treatment (described below) are the same for students who score low and high on a test of spatial visualization (Mental Rotations Test [MRT]; Peters et al., 1995).

Current Study

In the present research, we add to this small literature by examining the effects of an intervention that aims to improve diagram comprehension in high school science classes through strengthening CMR skills. We use Ainsworth's (2006) DeFT model of learning with representations to inform this intervention. With regard to *Design*, our CMR intervention emphasized information redundancy and sequencing of multiple representations by directly instructing students on when and how to shift back and forth between text and diagrams. With regard to *Function*, our CMR intervention called attention to visual features of representations in order to foster a relational understanding between text and diagrams. With regard to *Tasks,* our CMR intervention focused the learning activity on linking representations in text and diagrams, which is rare in classroom instruction (Ainsworth, 2006). The instructional focus on coordinating multiple representations was in addition to instruction on conventional features of diagrams (e.g. arrows, captions, and color coding), which were deemed to be necessary for coordinating meaning across representations.

We compared effects of an instructional method aimed to develop skills in coordinating multiple representations and understanding conventional features of diagrams (COD + CMR) with that of a simple instructional method aimed at developing an understanding of conventions of diagrams (COD) alone. Each intervention is described in the method section below. Our choice of the COD intervention as the comparison condition was based on prior results showing that the COD treatment is more beneficial than a no-treatment control (Cromley et al., 2013b). Both the school and parents felt it was more ethical to compare this new treatment to one known to be effective rather than to a business as usual condition. This type of research design is common in medical research, where the control group receives a treatment known to be effective and the intervention group receives a newly developed treatment (active concurrent control trial, see, e.g. European Medicines Agency, 2001). Our design is also typical for classroom intervention research (US Department of Education, 2012). In order to account for the possibility that growth in outcome measures is due to factors other than the treatments solely to the passage of time, we coded student work products during the intervention to create an index of their effort while learning (described below).

In the present research, we developed an intervention aimed at fostering CMR skills (in addition to understanding COD) and compared the joint effects to an intervention focused on understanding COD alone. Based on the literature reviewed above, we posed the following research questions using two samples of high school biology students:

(1) Does an intervention focused on Coordinating Multiple Representations along with Conventions of Diagrams (COD + CMR condition) lead to higher scores than an intervention focused only on COD (COD condition) on growth in biology content knowledge, biology diagram comprehension (near transfer), and geology diagram comprehension (far transfer)?
(2) Are shifts in scores on biology knowledge, biology diagram comprehension (near transfer), and geology diagram comprehension (far transfer) related to student effort during the treatment, as measured by answers to questions during the treatment itself?
(3) Does the effect size of either intervention depend on students' MRT scores?

Method

Student Participants

School 1. Participants were the entire population of tenth-grade students ($N = 67$) from a K-12 school for high-achieving low-income students in a large city in the mid-Atlantic region of the USA. They were studied in fall 2009 in four intact biology classes taught by the same teacher; this teacher in the previous year had delivered the COD intervention to a different cohort in the same school. Students' mean age was 15.19 years (SD = .50), they were 56% female and 80% African-American, 3% Asian, 7% Hispanic, and 10% identifying as mixed race. As a proxy for socioeconomic status, students self-reported parental education: 39% percent of mothers had received a bachelor's degree or higher, as had 12% of fathers. Due to absences on testing days, we have complete data on 28 students in the COD condition across two classes and 31 in the COD + CMR condition across two classes.

School 2. Participants were 115 biology students from a large, relatively low-achieving public high school in a small city in the mid-Atlantic region of the USA. Students came from 12 biology classes taught by 5 teachers in fall 2009, and classes were randomly assigned to treatment within teacher. Treatments were balanced across teachers, such that all teachers taught at least one section in each condition. Most students were in 9th grade, but some 10th graders were also enrolled in the class. The mean age of participants was 14.5 (SD = .81). Students were 51% female, and 65% were White, 19% African-American, 13% Hispanic, and 3% multiple races or other races. With regard to socioeconomic status, for 79% of participants neither parent had graduated from college. Due to absences on testing days, we have complete data on 46 students in the COD condition across 6 classes and 53 students in the COD + CMR condition across 6 classes.

Teacher Training

School 1. In the year prior to the current study, the teacher in School 1 had participated in a 2-hr workshop on implementing the COD intervention, had implemented the intervention for 32 school days, received feedback on his implementation from the research team, and over the summer had coded student workbook entries. Building on this deep familiarity with the COD condition, we provided a Teacher Edition of the COD + CMR materials, met and discussed any questions or concerns he had about the new intervention, and made several changes based on his comments on the workbooks. The Teacher Edition included ideas for scaffolding student learning that were interspersed between the student workbook pages, together with a copy of the instructional fidelity checklist (see Cromley et al., 2013b for details).

School 2. Five teachers in School 2 participated in a 2-hr professional development workshop at the school to prepare them to implement the intervention. During the workshop, we asked teachers to imagine that they were students learning from the workbooks we had created as we modeled the instruction and scaffolding that we wanted teachers to provide (e.g. defining a vocabulary term, providing a hint). Teachers received a Teacher Edition of both COD workbooks and COD + CMR workbooks, each of which included ideas for scaffolding interspersed between the student pages, a copy of our instructional fidelity checklist, and index cards on which to write suggestions for improving each activity and worksheet as we presented it. Teachers worked in groups to generate worksheet-specific scaffolds. Teachers' feedback was incorporated in the final version of the intervention workbooks. The materials were well received; teachers stated that they believed the interventions were likely to be effective and felt prepared to implement the intervention.

Fidelity of Implementation

School 1. A member of the research team was present at 60% of the class meetings to observe instruction using a fidelity-of-implementation checklist. Fidelity-of-implementation scores were nonsignificantly different across the two conditions. For both conditions, subscale scores showed high fidelity for initial assignment of practice problems and scaffolding student problem-solving, but lower fidelity for explaining the usefulness of the strategy and demonstrating the strategy.

School 2. A member of the research team was present at 26% of class meetings to observe using the same fidelity instrument described above. There were no significant differences in fidelity between the COD condition and the COD + CMR condition, and patterns for fidelity were the same as in School 1.

Instructional Materials—Coordinating Multiple Representations (COD + CMR) Workbook

School 1. The 26-page COD + CMR instructional workbook was built around three chapters of the biology textbook used in the classes (Johnson & Raven, 2005). The chapters focused on chemical properties and reactions; water, carbon, and nitrogen cycles, and ecosystems. The workbook consisted of scanned textbook images of almost every visual representation in the relevant chapters, together with explanations of COD (i.e. *Diagram Decoding Tips* described below). In Table 1, we describe the frequency of the characteristics of these representations. Using Treagust and Tsui's (2013) taxonomy, we describe visual representations in terms of four levels or grain sizes: macroscopic, microscopic, molecular, and symbolic. We also describe visual representations in terms of six types: photograph, diagram/drawing, table, map, graph, or equation; according to Treagust and Tsui, these types of representations correspond with increasing levels of abstraction as one moves from photographs to equations. In School 1, representations most commonly illustrated biological phenomena at the macroscopic and molecular levels, with many representations including both; common types of representation include photographs, diagrams, and drawings, with many figures including more than one type (e.g. a photograph of a mouse overlaid with a drawing of a fat molecule). We also characterize the representations in terms of three types of knowledge: *biological concepts*, which we define as the presence of two or more biology terms used to convey biological meaning (de Jong & Ferguson-Hessler, 1996); *biological processes*, which we define as an illustration of temporal relationships among biological events (e.g. nitrogen cycle or the transfer of energy through an ecosystem), and *biological structures*, which we define as an illustration of spatial relationship between components of a biological or chemical entity

Table 1. Characteristics of representations used in interventions by school

Representation characteristic	School 1 (%)	School 2 (%)
Level of representation		
Macroscopic	80.8	56.3
Microscopic	3.8	9.4
Molecular	53.8	37.5
Symbolic	7.7	37.5
More than one level	38.5	37.5
Type of representation		
Photograph	65.4	17.6
Diagram/drawing	69.2	73.5
Table	3.8	38.2
Map	3.8	0.0
Graph	0.0	0.0
Equation	0.0	2.9
More than one type	42.3	29.4
Type of knowledge represented		
Concepts	100.0	100.0
Processes	46.2	64.7
Structures	73.1	50.0
Features of representations		
Caption	100.0	85.3
Labels	84.6	82.4
Enlargement	30.8	8.8
Color	65.4	38.2
Arrows	57.7	32.4
Symbols/abbreviations	38.5	44.1
More than one feature	92.3	79.4

or system (e.g. when parts of a molecule are labeled with a symbol or color). In School 1, all representations depicted biological concepts, most depicted biological structures, and nearly half depicted biological processes. We also describe the frequency of diagrammatic conventions (i.e. captions, labels, enlargements, symbolic color, arrows, and symbols or abbreviations). In School 1, frequently used conventions included captions, labels, symbolic color, and arrows.

Workbooks also included a series of targeted questions designed to show students how to coordinate information in text with information in visual representations. For example, these questions asked students to read captions, locate the named parts and processes in the drawing, and draw a line between them; students were instructed to stop reading when they reached figure references (such as 'see figure 12-2'), find the relevant figure, and draw a line connecting them; students were also asked to find descriptive language in running text (e.g. 'the two strands separate') and to find that process depicted in the diagram.

Each question was designed at one of three levels of difficulty: Level 1 questions required the reader to use one piece of information and make no inference, such as 'According to the diagram, what can $H_3O^+$and OH^- be transformed into?' (32% of questions). Level 2 questions required the reader to link, match, or coordinate two pieces of information but did not require inference, such as 'Locate and circle the word that describes the *shape* of a glucose molecule. Draw a line to the diagram that shows that shape' (51% of questions). Level 3 questions required the reader to link, match, or coordinate more than two pieces of information or required interpretation, application of knowledge, or judgment, such as 'Describe in as much detail as possible what you see in the zoom-in [enlargement] above, using your own words. Be sure to find two examples each of adhesion and cohesion' (17% of questions). The workbook was reviewed by the classroom teacher for scientific accuracy.

The unique focus of tasks in the COD + CMR workbook was on coordinating multiple representations; these tasks were presented in addition to questions on COD. Instruction on COD was deemed necessary for the process of coordinating multiple representations, given that understanding diagrammatic conventions is sometimes lacking among high school students (Cromley et al., 2013b). Therefore, the COD + CMR workbooks contained the same questions that appeared in the COD workbook (described below) with additional tasks requiring the coordination of multiple representations; questions requiring the coordination of multiple representations constituted 40% (School 1) and 35% (School 2) of total points in the COD + CMR workbook.

School 2. The COD + CMR and COD workbooks (described below) was built around three chapters of the textbook from this school (Johnson & Raven, 2006); the chapters focused on basic genetics, the structure of DNA, and the process of protein synthesis. In School 2, representations most frequently illustrated biological phenomena at macroscopic, molecular and symbolic levels, with many representations including more than one level. Common types of representations include diagrams or drawings and tables, with many figures including more than one type. All representations depicted biological concepts, most depicted biological processes, and half depicted biological structures. Frequently used conventions included captions and labels (see Table 1). Otherwise, the number of pages, question types and distribution of questions, and diagram decoding tips (see below) were similar to School 1.

Instructional Materials—COD Workbook

School 1. The 26-page COD instructional workbook included the same diagrams used in the CMR-fostering workbook. These diagrams were accompanied by *Diagram Decoding Tips* that explained how to interpret conventions commonly used in biology diagrams: abbreviations, arrows, captions, symbolic color, enlargement, naming and explanatory labels, and symbols. Since visual representations were drawn from the same textbook, they generally had a consistent style and applied diagrammatic conventions in a similar way, consistent with recommendations in prior research (e.g. Cheng & Gilbert, 2015). These consistencies included how figures were labeled and were referred to in text, the presence of captions and their relationship to the figure, how symbolic color was used (e.g. oxygen molecules were depicted in red across representations), and the general stylistic esthetics of line drawings. However, representations within each school covered a broad range of topics and included a wide range of representation levels, types, knowledge, and conventions, as illustrated in Table 1. Where relevant, diagram decoding tips highlighted how features of representations (e.g. use of color, and arrows) may change from one representation to another. For example, one *Diagram Decoding Tip* stated, 'It is important to be aware of how *color* is used in photographs. Photographs often capture real color to show you what things look like in real life. You should be aware, however, that the colors that are photographed may not be the colors of the objects as they naturally appear. Scientists often dye objects, especially very small objects, so they can be better viewed under a microscope' (see Cromley et al., 2013b, for details). The questions were designed to represent the same difficulty levels as in the COD + CMR workbook, described above.

Measures

We used existing measures of diagram comprehension (biology and geology), background knowledge (biology), and the MRT.

Biology Diagram Comprehension. In our prior research, we had developed a 25-item measure of diagram comprehension in biology (Cromley et al., 2013a, 2013b). This is a measure of near-transfer, as the scale measures the ability to identify the main idea of visual representations similar to those in the students' own textbook, but ones that were not taught in the intervention (see Cromley et al., 2013b for sample items). Multiple-choice questions tap literal and inferential comprehension. For example, one question presented a cladogram with a color key, and asked students to identify what this color key indicated about features of one animal shown in the diagram, followed by 4 multiple-choice options. Previously obtained Cronbach's alpha reliability was .70, and discriminant validity is supported by a correlation of .57 with biology knowledge, both with 143 high school students (Cromley et al., 2013b).

Geology Diagram Comprehension. We used our previously developed far-transfer measure of diagram comprehension in geology (Cromley et al., 2013a, 2013b). This 10-item multiple-choice measure drew on images from a high school Earth Science textbook not used in either school. Like the near-transfer measure, 4-option multiple-choice questions tapped the main idea of each image and included both literal and inferential items. For example, one question asked students to identify the geological process involved in the movement of tectonic plates from a map and diagram of Japan. Previously obtained Cronbach's alpha reliability was .79, and discriminant validity is supported by a correlation of .31 with geology knowledge, both with 143 high school students (Cromley et al., 2013b).

Mental Rotations Test. We used the first 12 items of the *Mental Rotations Test-A* (MRT-A; Peters et al., 1995) to measure spatial skills. Participants matched a 3-D target figure with two of four presented figures in order to identify rotated versions of the target. The first 12 items of the measure show high reliability (Cronbach's alpha = .91 with N = 157 undergraduate students; Voyer et al., 2006). Concurrent discriminant validity is supported by correlations of the full scale with Paper Folding of r = .49 with N = 80 undergraduate students (Lequerica, Rapport, Axelrod, Telmet, & Whitman, 2002), and in our prior research, the half-scale was significantly correlated with biology diagram comprehension scores (r = .40) but not geology diagram scores (r = .20; Cromley et al., 2013b).

Diagram-Aligned Biology Background Knowledge. Our previously developed 25-item measure used multiple-choice items to assess topic-specific background knowledge (Cromley et al., 2013a, 2013b). Specifically, the biology background knowledge questions were designed to tap prerequisite knowledge for the biology diagram comprehension measure; this alignment should best enable us to detect a relationship between knowledge and diagram comprehension if there is one. For example, one question asked about purposes of blood in an animal's body, followed by four multiple-choice options. This was aligned with a biology diagram that asked about blood flow in a turtle's body. Previously reported reliability was .80 with 143 high school students; see above for validity evidence (Cromley et al., 2013b).

Diagram-Aligned Geology Background Knowledge. Our previously developed 10-item measure likewise used multiple-choice items to assess topic-specific geology background knowledge (Cromley et al., 2013a, 2013b). As with the biology measures, this was the prerequisite knowledge for the geology diagram comprehension measure, so that we would have the most power to detect a relationship between knowledge and diagram comprehension. For example, one question asked about what happens when cooler and hotter rocks meet, followed by four multiple-choice options. This item was aligned with a geology diagram item that asked about the effects of convection currents below the Earth's crust on movement of tectonic plates in the crust. Previously reported Cronbach's alpha reliability was .74 with 143 high school students; see above for evidence supporting validity (Cromley et al., 2013b).

Effort Scores for Workbook Answers. As students participated in the intervention, they wrote their answers in the workbooks. Each of the student answers was coded for student level of effort, taking into account the difficulty level of each question. For example, a student who summarized information in response to a question that required inference would receive a score of 2 (out of the total of 3 points possible for an inferential question). A student who drew an inference in his or her answer to the same question would receive a score of 3, even if the inference was not completely correct. We also awarded partial credit for multipart answers. In prior research, we had found that scoring for correctness and scoring for level of effort were very highly correlated ($r = .97$), so we scored only for level of effort. In the COD + CMR workbooks, we tallied workbook scores separately for questions that asked only about COD (we refer to these Conventions of Diagrams Effort score, or *COD Effort* scores for short) and for questions that required coordinating multiple representations (we refer to these scores as *CMR Effort* scores). These two scores were highly correlated ($r = .90$ in School 1 and $r = .80$ in School 2). For all analyses, these raw scores were converted to percentages.

Schools 1 and 2. For School 2, one coder scored all 17,583 student workbook answers, and another coder rescored 35% of the answers. Interrater agreement was greater than 99%. The mean COD Effort score was .54 (SD = .18) and the mean CMR Effort score was .40 (SD = .20), supporting our categorization of question difficulty. Analyses of School 2 effort codes indicated that effort scores based on the first half of the workbook exercises yielded effort scores representative of the second half of workbook exercises. To avoid superfluous coding, only the first 18 workbook pages were coded for School 1; a second coder rescored 35% of the answers. Interrater agreement was greater than 98%. For School 1, the mean COD Effort score was .78 (SD = .14) and the mean CMR Effort score was .66 (SD = .21); students were likely to receive more credit for the questions worth 1 point ($M = 85\%$), compared to 2-point questions ($M = 80\%$) or 3-point questions ($M = 61\%$), again supporting our categorization of question difficulty.

Procedure

School 1. After we obtained parent/guardian consent and student assent, we administered the pretests in a single 48-min class session. Classes were randomly assigned to treatments. Over the course of two months, the teacher conducted his lessons as usual, but stopped when he reached a diagram in the textbook and asked students to complete one worksheet that related to that diagram (approximately 5 min per worksheet); this was followed by whole-class discussion of the worksheet (an additional 5 min for a total of approximately 260 min of intervention over 18 class periods). Students were group posttested using the same procedures as for pretesting.

School 2. The same procedures were followed as in School 1.

Preliminary Data Analysis

Intra-class Correlation Coefficients (ICCs) on all pretest variables ranged from .03 to .04 (design effect = 1.05, which is far below the recommended cutoff of a design effect ≤ 1.64; Muthén & Satorra, 1995). This indicates that although students were nested within classrooms, there was extremely low nonindependence of observation. The low ICCs suggest that diagram-related skills were equally underdeveloped across all classrooms and that multilevel modeling was not warranted. For the individual difference × treatment interaction, we used mixed ANCOVA, adding the MRT-A spatial scores as a covariate to the analyses described above. If MRT-A scores advantage or disadvantage students in either treatment, this would be indicated by a significant Time × condition × MRT-A interaction.

Results

Descriptive statistics and intercorrelations among all variables by school are shown in Table 2, and mean scores for each group are shown in Table 3. To evaluate the effects of COD and COD + CMR treatments on biology diagram comprehension, geology diagram comprehension, and biology content knowledge (Research Question 1), we conducted a series of mixed 2 (Between: COD, COD + CMR conditions) × 2 (Between: School 1, School 2) × 2 (Within: Pretest, Posttest) ANOVAs. We followed up significant main effects of time with post hoc repeated-measures and one-tailed *t* tests by group, based on our expectation of positive change over time. To determine whether pre- to posttest shifts were related to student effort (Research Question 2), we conducted a series of ANCOVAs in which effort scores were added as a covariate to the analyses for Research Question 1. Finally, to determine whether effect size of either intervention depended on students' spatial skill (Research Question 3), we conducted a series of ANCOVAs on posttest scores in which spatial scores × treatment interactions were examined.

Research Question 1: Effects of Treatments

Biology Diagram Comprehension. There was a significant main effect of time (F [1, 155] = 21.62, MSE = 5.57, $p < .001$, $\eta_p^2 = .12$). There was no significant main effect of condition, nor were there significant time × condition or time × school interactions (see Table 4 for full results). To investigate the main effect of time, post hoc repeated-measures *t* tests were conducted by group and showed significant growth over time for both the COD + CMR and COD groups in both schools (see Table 5 for full results). The COD + CMR intervention was just as effective as the previously validated COD intervention for biology diagram comprehension.

Geology Diagram Comprehension. There was no significant main effect of time (F [1, 155] = 2.97, MSE = 1.81, $p = .09$, $\eta^2_p = .02$). There was no significant main effect of condition, nor were there significant time × condition or time × school interactions (see Table 4 for full results).

Biology Knowledge. There was a significant main effect of time (F [1, 155] = 20.80, MSE = 3.15, $p < .001$, $\eta^2_p = .12$). There was no significant main effect of condition, nor were there significant time × condition or time × school interactions (see Table 4 for full results). To investigate the main effect of time, post hoc

Table 2. Descriptive statistics on and intercorrelations among study variables for School 1 (below diagonal) and School 2 (above diagonal)

	1	2	3	4	5	6	7	8	9	10	11	*M*	*SD*
1. MRT-A (spatial skills) T1	–	.47	.36	.37	.24	.33	.10	.18	.23	.23	−.03	2.98	2.68
2. Bio Knowledge T1	.35	–	.49	.57	.21	.71	.36	.55	.30	.37	.27	6.80	2.93
3. Geology Knowledge T1	.34	.37	–	.49	.41	.57	.46	.42	.37	.21	.07	4.07	1.75
4. Bio. Diag. Comp. T1	.40	.30	.24	–	.27	.56	.30	.60	.37	.37	.17	10.6	3.76
5. Geo. Diag. Comp. T1	.20	.31	.22	.49	–	.37	.13	.40	.41	.17	.07	2.90	1.63
6. Bio Knowledge T2	.09	.45	.27	.30	.20	–	.50	.60	.56	.42	.31	8.14	3.60
7. Geo Knowledge T2	.15	.15	.43	.31	.27	.44	–	.40	.34	.46	.32	4.30	1.71
8. Bio. Diag. Comp. T2	.25	.10	.10	.46	.42	.18	.40	–	.46	.09	.12	11.85	4.07
9. Geo. Diag. Comp. T2	.18	.20	.01	.26	.17	.18	.05	.12	–	.19	.08	3.39	1.73
10. COD Effort score	.12	−.17	−.04	.21	.04	.16	.27	.02	.12	–	.80	54%	18%
11. CMR Effort score (COD + CMR cond. only)	−.01	−.20	−.10	.29	.26	.13	.35	−.01	.31	.90	–	40%	20%
M	2.59	7.30	5.43	14.22	4.10	8.15	5.83	15.67	4.25	78%	66%		
SD	2.12	2.05	1.54	2.68	1.70	2.65	1.63	3.01	1.46	14%	21%		

Note. Bio = biology; Geo = Geology; Diag. Comp. = Diagram Comprehension; T1 = Pretest; T2 = Posttest; Statistics for School 1 are shown below the diagonal and for School 2 are shown above the diagonal.

Table 3. Descriptive statistics for all measures across conditions and times

Measure (maximum score)	Pretest *M* (*SD*) COD condition	Pretest *M* (*SD*) COD + CMR condition	Posttest *M* (*SD*) COD condition	Posttest *M* (*SD*) COD + CMR condition
MRT-A (12)				
School 1	2.67 (1.73)	2.52 (2.45)		
School 2	2.97 (2.82)	3.00 (2.58)		
Biology knowledge (25)				
School 1	7.10 (1.79)	7.48 (2.28)	8.18 (2.82)	8.13 (2.59)
School 2	7.61 (2.97)	6.51 (2.82)	8.76 (3.72)	7.36 (3.22)
Geology knowledge (10)				
School 1	5.27 (1.41)	5.58 (1.66)	5.86 (1.67)	5.81(1.64)
School 2	4.41 (1.93)	3.96 (1.54)	4.46 (1.83)	4.13 (1.51)
Biology Diagram Comprehension (25)				
School 1	14.07 (2.46)	14.36 (2.90)	15.32 (2.71)	15.90 (3.30)
School 2	11.54 (4.15)	10.02 (3.42)	12.96 (3.81)	10.98 (3.91)
Geology Diagram Comprehension (10)				
School 1	3.73 (1.68)	4.42 (1.68)	4.07 (1.63)	4.39 (1.31)
School 2	3.13 (1.82)	2.72 (1.42)	3.35 (1.90)	3.30 (1.55)
COD Effort score				
School 1			.83 (.07)	.73 (.17)
School 2			.60 (.18)	.51 (.16)
CMR Effort score				
School 1				.66 (.21)
School 2				.40 (.20)

Note. COD Effort score and CMR Effort score refer to the level of inference made on workbook questions focusing on COD and multiple representations, respectively.

repeated-measures *t* tests in each school showed significant growth over time for the COD + CMR condition and one school showed significant growth over time for the COD condition (see Table 5 for full results). The COD + CMR intervention was more consistently effective compared to the previously validated COD intervention for biology knowledge.

Research Question 2: Are Shifts in Scores Related to Effort During Learning?

To determine shifts in scores over time were related to effortful uptake of the intervention rather than the simple passage of time, we added effort score as a covariate to the mixed ANOVA analyses described above to test whether students who engaged in more effort (i.e. obtaining higher workbook scores) during the intervention shifted more from pre- to posttest. Recall that higher scores were awarded for summarizing rather than repeating verbatim, giving two-part rather than one-part answers, and attempting to draw inferences when these were asked for. If the main effect of time is no longer significant after accounting for variance explained by effort, this can

Table 4. Results of mixed ANOVAs

	F	MSE	p	η_p^2	Obs. Power
Biology diagrams					
Time	21.62	5.57	<.001	.12	1.00
School	49.69	19.48	<.001	.24	1.00
Condition	3.34	19.48	.07	.02	.44
Time × School	0.13	5.57	.72	<.01	.07
Time × Condition	0.95	5.57	.68	<.01	.07
Geology diagrams					
Time	2.97	1.81	.09	.02	.40
School	23.06	3.56	<.001	.13	1.00
Condition	<0.01	3.56	.99	<.01	.05
Time × School	0.78	1.81	.38	<.01	.14
Time × Condition	0.33	1.81	.57	<.01	.09
Biology knowledge					
Time	20.80	3.15	<.001	.12	1.00
School	0.19	13.92	.66	<.01	.07
Condition	3.48	13.92	.06	.02	.46
Time × School	0.07	3.15	.79	<.01	.06
Time × Condition	0.31	3.15	.58	<.01	.09

Note. *df* are 1, 155 for all effects.

Table 5. Results of post hoc repeated-measures t tests to follow up on main effects of time

	COD			COD + CMR		
Measure (maximum score)	t	p	d	t	p	d
Biology Diagram Comp. (25)						
School 1	2.97	<.01	.51	2.27	.02	.57
School 2	2.68	<.01	.34	2.01	.03	.28
Biology Knowledge (25)						
School 1	2.10	.05	.53	1.74	.05	.38
School 2	0.82	.21	.12	2.34	.01	.41

Note. COD = Conventions of Diagrams intervention; COD + CMR = Coordinating Multiple Representations intervention; Cohen's d was calculated as (Posttest M—Pretest M)/Pretest SD.

indicate that change over time was due to effortful engagement with the workbooks rather than the simple passage of time or students' shared classroom experiences (recall that conditions were assigned within teachers). The maximum effort percentage score was entered as the covariate; for students in the COD condition, this effort score was based on COD questions; for students in the COD + CMR condition, this effort score was based on COD and CMR questions. We conducted this analysis only for biology diagram and biology knowledge scores since there was no main effect of time for geology diagram comprehension.

For biology diagram comprehension, the main effect of time was no longer significant after accounting for variance associated with effort scores (F [1, 149] = 0.03, MSE =

5.43, $p = .87$, $\eta^2 < .001$). Likewise, for biology topic knowledge, the main effect of time was no longer significant after accounting for variance associated with effort scores (F [1, 149] = 1.36, MSE = 3.07, $p = .25$, $\eta^2 = .01$). The effect size of time for biology diagram comprehension was reduced from $\eta^2 = .12$ in a significant ANOVA to $\eta^2 < .001$ in a non-significant ANOVA, and for biology topic knowledge, it was reduced from $\eta^2 = .12$ in a significant ANOVA to $\eta^2 = .02$ in a nonsignificant ANOVA. In sum, when effort scores were added as a covariate, the main effect of time for growth in biology diagram comprehension and content knowledge was no longer significant. We interpret these results to suggest that engagement with the invention materials themselves was responsible for students' increased scores at posttest compared to pretest.

Given the apparent differences in mean effort scores across schools and conditions, we conducted analyses to explore these differences. Results from a 2 (School: School 1, School 2) × 2 (Condition: COD, COD + CMR condition) ANOVA on COD Effort scores indicated a significant main effect of school (F [1, 163] = 88.12, MSE = 0.03, $p < .001$, $\eta^2 = .35$), with School 1 showing higher COD Effort scores than School 2 (M = .78 vs .54). There was a significant main effect of condition (F [1, 163] = 9.72, MSE = 0.03, $p < .001$, $\eta^2 = .06$), with the COD condition showing higher COD Effort scores than the COD + CMR condition (M = .67 vs .59). There was no significant school by condition interaction (F [1, 163] = 0.55, MSE = 0.03, $p = .46$, $\eta^2 < .01$). For CMR Effort scores, an independent samples t test showed a significant difference between schools (t [88] = 5.72, $p < .001$), with School 1 showing higher CMR Effort scores than School 2 (M = .66 vs .40). Within the COD + CMR condition, COD Effort scores were significantly higher than CMR Effort scores in both School 1 (t [33] = 4.59, $p < .001$) and School 2 (t [55] = 6.68, $p < .001$).

We next examined the differences in mean raw effort scores (as opposed to percentage scores) of the two conditions to evaluate whether the COD condition was associated with more total effort than the COD + CMR condition or whether students in the COD + CMR condition demonstrated a comparable amount of effort, but this effort was distributed across more questions. An independent sample t test indicated no significant differences in raw effort scores, t[150] = −.02, $p = .98$; that is, although COD Effort scores (as a percentage of total possible points) were higher than CMR Effort scores (as a percentage), the raw scores were not different from each other, indicating a similar amount of student effort in both conditions. In summary, regardless of school, the COD and COD + CMR conditions prompted a similar amount of effort, and regardless of condition or type of workbook question, students in School 1 demonstrated more effort than did students in School 2.

Research Question 3: Does the Effect Size of Either Intervention Depend on Students' Mental Rotations Test scores?

To determine whether spatial skills were related to the effects of each intervention, we categorized students by MRT-A scores: students above the median ($Mdn = 2$) were coded as high spatial ability relative to the group ($n = 80$) and students at or below the median were coded as low spatial ability ($n = 89$). We then conducted a series of one-way ANCOVAs in which the dependent variable was the posttest score (biology knowledge, biology diagram comprehension, or geology diagram comprehension) with spatial ability and treatment condition as between-subjects variables and pretest knowledge and pretest diagram scores as covariates. For biology diagram comprehension, results indicated no main effect of spatial score, F (1, 153) = 1.52, MSE = 9.32, $p = .22$, $\eta_p^2 = .01$, and a trend for spatial ability by treatment interaction, F (1, 153) = 2.67, MSE = 9.32, $p = .10$, $\eta_p^2 = .02$. That is, in the COD + CMR condition, students with low spatial scores showed greater gains in biology diagram comprehension than did those with high spatial scores. For geology diagram comprehension, results indicated no main effect of spatial score, F (1, 153) = 1.88, MSE = 2.31, $p = .17$, $\eta_p^2 = .01$, and a trend for a spatial ability × treatment interaction, F (1, 153) = 3.23, MSE = 2.31, $p = .07$, $\eta_p^2 = .02$. That is, in the COD condition, students with low spatial scores gained less in geology diagram comprehension than did those with high spatial scores. For biology content knowledge, results indicated no main effect of spatial score, F (1, 153) = 0.24, MSE = 6.07, $p = .63$, $\eta_p^2 = .01$, and no significant treatment x spatial ability interaction, F (1, 153) = 0.03, MSE = 6.07, $p = .86$, $\eta_p^2 = .01$.

Discussion

Coordinating multiple representations has been identified as a key skill for biology learning (Treagust & Tsui, 2013; Tsui & Treagust, 2013), yet it has been under-researched as an intervention, especially in non-laboratory settings, in science classrooms, and with secondary school students. Our study examined the learning effects of two interventions that used teacher-led classroom activities and discussion to support learning from biology diagrams: a COD + CMR intervention, which provided instruction on COD along with prompts to help students connect text and diagrams, and a COD intervention, which provided information about diagrammatic conventions alone. Results indicated that both the COD + CMR and COD interventions were effective in increasing comprehension of biology diagrams and for building basic biology knowledge; the COD + CMR intervention showed more consistent effects on learning and was more effective for students with low spatial ability compared to the COD intervention. We found medium effects of each intervention across two diverse and quite different samples of high school students. As we discuss below, these results extend prior research and highlight how the coordination of multiple representations can be supported in actual science classrooms.

Results from our COD + CMR intervention extend prior laboratory research (e.g. Bartholomé & Bromme, 2009; Bodemer & Faust, 2006) in demonstrating that the coordination of multiple representations can be supported in classroom settings, and that doing so is associated with increased comprehension of biology diagrams and biology content learning. The COD + CMR intervention involved instructing students on relevant diagrammatic conventions and prompting them to connect text and diagrams at figure references, when new vocabulary terms were introduced, when new concepts were shown, and when spatial language was used. Our results indicate that these prompts were associated with both significant improvements in biology diagram scores and biology knowledge scores at both the higher and lower achieving schools (Schools 1 and 2, respectively). Our results are not consistent with Bartholomé and Bromme (2009), who found nonsignificant results for prompting undergraduate students to coordinate multiple representations. This discrepant pattern of results likely reflects the many methodological differences between Bartholomé and Bromme's and our study; respectively, these differences include computer-based vs. workbook-based learning tasks, absence vs. presence of teacher-led discussion, a single learning session vs. distributed instruction over one month; absence vs. presence of instruction on diagrammatic conventions; and presence vs. absence of visual manipulations linking text and diagrams such as number or color highlighting. Alternatively, our findings may reflect developmental differences, suggesting that students with less instruction in using diagrams and at a lower level of schooling can indeed benefit from direct instruction and prompting on the process of coordinating multiple representations.

Our results add to the limited evidence base on helping secondary students coordinate multiple representations in classroom settings. Consistent with prior classroom research (e.g. Kombartzky et al., 2010; Ploetzner & Schlag, 2013; Schlag & Ploetzner, 2011), results from our study demonstrate that students can be taught learning strategies that improve learning from text and visuals. We extend this research by demonstrating that students' ability to coordinate multiple representations can be improved in in-tact classroom settings, over extended periods of time, and via instruction and discussion that are delivered with the help of teachers who received a modest amount of training. We further extend prior research by showing that training high school students to coordinate multiple representations can result in improved diagram comprehension skills that transfer to untaught biology diagrams, while also supporting the learning of biology concepts.

Perhaps most importantly, our results indicate learning effects when students are coordinating multiple representations found in their actual textbooks, where text is accompanied by a diverse set of visual representations. In our study, we found that in a period of one month high school biology students were tasked with learning from a broad range of visuals that represented biological concepts, structures, and processes through a wide range and combination of grain sizes (e.g. macroscopic vs. microscopic), representation types (e.g. tables vs. diagram), and visual features (labels vs. arrows). This variability in representations is likely the norm rather than the exception in secondary biology textbooks, and underscores the need to train students to learn from the varied visual representations with which they are presented in biology (Griffard, 2013; Roth & Pozzer-Ardenghi, 2013; Treagust & Tsui, 2013; Tsui & Treagust, 2013). This diversity in visual representation underscores the value of interventions that train students to apply flexible learning strategies for coordinating texts and diagrams (e.g. Schlag & Ploetzner, 2011). Indeed, the near-transfer effects on biology diagram comprehension may be the result of students working with a diverse set of visual representations and learning how to flexibly coordinate them.

Similar results across the two sites suggest that our approach is likely to support students' diagram comprehension skills even as biology content, characteristics of visual representations, and contextual variables differ. The learning materials at each school covered different biological concepts and used different visual representations. School 1 and School 2 differed in several demographic features (private vs. public, urban vs. suburban, small vs. large school; one vs. five participating teacher, respectively). Students and teachers at both sites were able to complete the exercises and discussion during class time, and teachers were able to deliver the two interventions with good fidelity even when the intervention was new to them. This replication is particularly important given the wide range of visual representations students encounter in biology, from one textbook page to the next, and from one classroom to another.

While effects were generally replicated at both schools, we note that effect sizes for School 1 were generally larger than that for School 2 despite the smaller sample and lower variability in School 1. We further note that the students in School 1 appear to be of higher ability, as shown by significantly higher biology and geology diagram scores (see Table 3), and that the teacher in School 1 was more experienced with delivering diagram interventions. Perhaps students with better diagram skills at pretest learn more from the interventions because they can build on very basic pre-existing skills. On the other hand, the higher pretest scores could simply be a proxy for some other important unmeasured differences between the schools, such as reading comprehension or vocabulary, motivation, some other spatial ability, and so on.

We found that our COD + CMR intervention was as effective as a previously validated intervention that focused on understanding COD. Why might the COD + CMR intervention not benefit students *more* than the simpler COD intervention? One possible explanation is that workbook questions focusing on coordinating multiple representations displaced questions on COD. Students were given the same amount of time for both conditions but had to answer more questions in the COD + CMR condition than in the COD condition. The presence of more questions competed for limited time allocated to the intervention during class as well as student effort. As a result, more time and effort on CMR questions meant less time and effort allocated to COD questions. Our analyses of the workbook effort scores support this explanation. Across both schools, COD Effort scores were significantly lower in the COD + CMR condition than in the COD condition, suggesting that questions on coordinating multiple representations may have displaced effort otherwise devoted to COD questions. This is further supported by the fact that raw effort scores were not significantly different across conditions. Therefore, any potential additional benefits of questions focusing on coordinating multiple representations may have been washed out by reduced benefits of attention to COD.

Another explanation for the lack of additive effects of CMR beyond COD is that the questions addressing coordinating multiple representations may have led students to make shallow connections between diagrams and text, but not engage in the effortful inferences that are required to comprehend diagrams (as has been seen in some interventions with researcher-added cues in diagrams; e.g. Jian, Wu, & Su, 2014). Within the COD + CMR condition, CMR Effort scores were significantly lower than COD Effort scores answered by the same students, suggesting that either CMR questions were harder than COD questions or they did not lead students to draw inferences to the same extent, or both. One implication of these results for future interventions with COD + CMR is that questions need to not only draw students' eyes from running text to diagram and back, but they also need to prompt the deep inferential activity that is associated with better learning (Cromley et al., 2013a; Hegarty, 2005; Jian et al., 2014). An additional alternative explanation for the similar results across COD and COD + CMR interventions is that once students have mastered the basics of conventions used in diagrams, this may enable them to engage even without prompting in the inferential processes that presumably underlie better diagram comprehension.

Our analyses also show that effort during learning—which falls under Ainsworth's (2006) *Task* category of individual characteristics of learners—makes a difference in the effectiveness of diagram instruction. Prior research has demonstrated that an intervention that supports students' understanding of COD is more effective than a no-treatment control (Cromley et al., 2013b). In the present research, we have shown that a higher level of engagement with the workbooks, such as trying to make the effortful inferences demanded by the more-difficult questions, was associated with better learning from the intervention. The fact that students needed to engage in effortful processing when learning from diagrams is no surprise, given prior laboratory research findings to this effect (Butcher, 2006). The current study shows the importance of effortful processing of diagrams in teacher-led workbook-and-discussion activities in secondary science classroom settings. Furthermore, our findings echo Ainworth's (2006) claim that individual characteristics like effort interact with task demands involved in diagram comprehension. For example, the level of effort was significantly different between conditions, with the presumably easier tasks associated with understanding COD yielding higher student effort than did tasks associated with coordinating multiple representations.

Spatial skill as measured by the Mental Rotations Test (MRT-A)—the second individual difference that we considered—was related to how much some students learned from the interventions. Our results indicated that the COD + CMR intervention was advantageous (in biology diagram comprehension) or equitable (for geology diagram comprehension) for low-spatial students compared to high-spatial students. By contrast, low-spatial students in the COD intervention were disadvantaged for geology diagram comprehension compared to high-spatial students. We speculate that the COD + CMR intervention provided scaffolds in the form of prompts that helped students compensate for low spatial skills. For example, prompting students to find a phrase in text and drawing a line to the corresponding diagram provided a structured task that may have assisted students with low spatial skills in connecting related information in different spatial locations. These scaffolds may be particularly important in the context of geology (e.g. Sanchez & Wiley, 2014), where diagrams often show changes across all three dimensions simultaneously. These findings contribute to limited classroom research on spatial x treatment interactions, though more research is needed to replicate these tentative trends. Our results are consistent with prior studies that have found that spatial abilities influence the effectiveness of interventions to support learning with diagrams (Bergey et al., 2015; Cromley et al., 2013a, 2013b; Hinze et al., 2013; Höffler & Leutner, 2011; Lee & Shin, 2011, 2012; Nguyen, Nelson, & Wilson, 2012), and underscore the importance of examining spatial skills in future intervention studies. Future research might also examine how other spatial skills may be differentially related to different domains and different interventions. For example, picking out patterns from a complex array (figure-ground tasks such as those tapped by the Hidden Figures Test) may be more related to learning with biology diagrams, where cross-sectioning skills are commonly required.

Limitations and Future Research

Our research design compared the effects of CMR instruction in addition to basic COD instruction to the benefits of basic COD instruction alone. The COD intervention—not thought to prompt inferences—and the COD + CMR intervention—which should have prompted inferences—produced very similar results. Despite collecting data from 12 classrooms across six teachers and two schools, the observed power for detecting some interactions of interest, such as a time × condition interaction, was low; therefore, we may not have detected interaction effects that were in fact present. Furthermore, given the research design, the unique effects of each instructional method cannot be fully disentangled. In future, classroom intervention studies researchers might consider gathering more process data from students while they are engaged in the learning tasks, which could illuminate the unique and common effects of different interventions. In the current study, we used effort in workbook questions as a measure of engagement with the intervention, yet we cannot rule out the possibility that gains attributed to engagement with the intervention were not instead driven by other unmeasured differences, such as higher self-efficacy (task-specific self-confidence) for diagram comprehension.

Conclusion

Diagram comprehension is a key competency of good learners and experts in science, and coordinating information from multiple representations is central to this competency. Although laboratory interventions cannot always be translated into successful classroom instructional methods, we were able to instantiate the principles of design, function, and tasks from Ainsworth's (2006) DeFT framework in classroom instruction that can help students build biology knowledge and better understand diagrams in the instructed domain. Results indicated that a teacher-delivered intervention that supports the coordination of multiple representations in classroom instructional materials is as effective as an intervention that fosters understanding COD.

ORCID

Bradley W. Bergey http://orcid.org/0000-0002-0495-7219
Jennifer G. Cromley http://orcid.org/0000-0002-6479-9080
Nora S. Newcombe http://orcid.org/0000-0002-7044-6046

Acknowledgements

We thank the students who participated in the study, their parents, teachers, and school administrators. We also thank John Romano and Dar Sheth for their assistance with scoring and data entry.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This material is based in part upon work supported by the US National Science Foundation Award #0815245. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not represent views of the US National Science Foundation.

References

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, *16*(3), 183–198.

Ainsworth, S. E., & Iacovides, I. (2005, August). *Learning by constructing self-explanation diagrams*. Paper presented at the EARLI conference, Nicosia, Cyprus.

Alias, M., Black, T. R., & Gray, D. E. (2002). Effect of instruction on spatial visualization ability in civil engineering students. *International Education Journal*, *3*(1), 1–12.

Bartholomé, T., & Bromme, R. (2009). Coherence formation when learning from text and pictures: What kind of support for whom? *Journal of Educational Psychology*, *101*(2), 282–293. doi:10.1037/a0014312

Berthold, K., & Renkl, A. (2009). Instructional aids to support a conceptual understanding of multiple representations. *Journal of Educational Psychology*, *101*(1), 70–87.

Bergey, B. W., Cromley, J. G., Kirchgessner, A., & Newcombe, N. S. (2015). Using diagrams versus text for spaced restudy: Effects on learning in 10th grade biology classes. *British Journal of Educational Psychology*, *85*(1), 59–74. doi:10.1111/bjep.12062

Black, A. A. (2005). Spatial ability and earth science conceptual understanding. *Journal of Geoscience Education*, *53*(4), 402–414.

Bodemer, D., & Faust, U. (2006). External and mental referencing of multiple representations. *Computers in Human Behavior*, *22*(1), 27–42.

Butcher, K. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology*, *98*(1), 182–197.

Carter, C. S., LaRussa, M. A., & Bodner, G. M. (1987). A study of two measures of spatial ability as predictors of success in different levels of general chemistry. *Journal of Research in Science Teaching*, *24*(7), 645–657.

Cheng, M. M., & Gilbert, J. K. (2015). Students' visualization of diagrams representing the human circulatory system: The use of spatial isomorphism and representational conventions. *International Journal of Science Education*, *37*(1), 136–161.

Chi, M. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representations of physics problems by experts and novices. *Cognitive Science*, *5*(2), 121–152.

Cromley, J. G., Bergey, B. W., Fitzhugh, S. L., Newcombe, N., Wills, T. W., Shipley, T. F., & Tanaka, J. C. (2013a). Effectiveness of student-constructed diagrams and self-explanation instruction. *Learning and Instruction*, *26*(1), 45–58. doi:10.1016/j.learninstruc.2013.01.003

Cromley, J. G., Perez, A. C., Fitzhugh, S., Newcombe, N., Wills, T. W., & Tanaka, J. C. (2013b). Improving students' diagrammatic reasoning: A classroom intervention study. *Journal of Experimental Education*, *81*(4), 511–537. doi:10.1080/00220973.2012.745465

Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Cognitive activities in complex science text and diagrams. *Contemporary Educational Psychology*, *35*, 59–74. doi:10.1016/j.cedpsych.2009.10.002

De Jong, T., & Ferguson-Hessler, M. G. (1996). Types and qualities of knowledge. *Educational Psychologist*, *31*(2), 105–113.

Eilam, B. (2012b). *Teaching, learning, and visual literacy: The dual role of visual representation in the teaching profession*. New York, NY: Cambridge University Press.

European Medicines Agency. (2001). *ICH 10: Choice of control group and related issues in clinical trails*. International Conference on Harmonization; London UK. 2000. Adopted by CPMP July 2000 (CPMP/ICH/363/96).

Ferk, V., Vrtacnik, M., Blejec, A., & Gril, A. (2003). Students' understanding of molecular structure representations. *International Journal of Science Education*, *25*(10), 1227–1245.

Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review*, *23*(4), 523–552.

Griffard, P. B. (2013). Deconstructing and decoding complex process diagrams in university biology. In D. F. Treagust & C-.Y. Tsui (Eds.), *Multiple representations in biological education* (pp. 164–184). Dordrecht: Springer.

Hegarty, M. (2005). Multimedia learning about physical systems. In M. Hegarty & R. E. Mayer (Eds.), *The Ccmbridge handbook of multimedia learning* (pp. 447–465). New York, NY: Cambridge University Press.

Hegarty, M., & Kriz, S. (2008). Effects of knowledge and spatial ability on learning from animation. In R. Lowe & W. Schnotz (Eds.), *Learning with animation: Research implications for design* (pp. 3–29). Cambridge, MA: Cambridge University Press.

Hinze, S. R., Rapp, D. N., Williamson, V. M., Shultz, M. J., Deslongchamps, G., & Williamson, K. C. (2013). Beyond ball-and-stick: Students' processing of novel STEM visualizations. *Learning and Instruction*, *26*, 12–21.

Höffler, T. N. (2010). Spatial ability: Its influence on learning with visualizations—A metaanalytic review. *Educational Psychology Review*, *22*(3), 245–269. doi:10.1007/s10648-010-9126-7

Höffler, T. N., & Leutner, D. (2011). The role of spatial ability in learning from instructional animations–Evidence for an ability-as-compensator hypothesis. *Computers in Human Behavior*, *27*(1), 209–216.

Imhof, B., Scheiter, K., & Gerjets, P. (2011). Learning about locomotion patterns from visualizations: Effects of presentation format and realism. *Computers & Education*, *57*(3), 1961–1970.

Jian, Y.-C., Wu, C.-J., & Su, J.-H. (2014). Learners' eye movements during construction of mechanical kinematic representations from static diagrams. *Learning and Instruction*, *32*, 51–62.

Johnson, G., & Raven, P. (2005). *Holt biology*. Austin, TX: Holt, Rinehart, & Winston.

Johnson, G., & Raven, P. (2006). *Holt biology*. Austin, TX: Holt, Rinehart, & Winston.

Kastens, K. A., Agrawal, S., & Liben, L. S. (2009). How students and field geologists reason in integrating spatial observations from outcrops to visualize a 3-D geological structure. *International Journal of Science Education*, *31*(3), 365–393.

Khooshabeh, P., & Hegarty, M. (2010). Inferring cross-sections: When internal visualizations are more important than properties of external visualizations. *Human–Computer Interaction*, *25*(2), 119–147.

Kombartzky, U., Ploetzner, R., Schlag, S., & Metz, B. (2010). Developing and evaluating a strategy for learning from animations. *Learning and Instruction*, *20*(5), 424–433.

Kozma, R., Chin, E., Russell, J., & Marx, N. (2000). The roles of representations and tools in the chemistry laboratory and their implications for chemistry learning. *Journal of the Learning Sciences*, *9*(2), 105–143.

Lee, D. Y., & Shin, D.-H. (2011). Effects of spatial ability and richness of motion cue on learning in mechanically complex domain. *Computers in Human Behavior*, *27*(5), 1665–1674.

Lee, D. Y., & Shin, D.-H. (2012). An empirical evaluation of multi-media based learning of a procedural task. *Computers in Human Behavior*, *28*(3), 1072–1081.

Leopold, C., & Leutner, D. (2012). Science text comprehension: Drawing, main idea selection, and summarizing as learning strategies. *Learning and Instruction*, *22*(1), 16–26. doi:10.1016/j.learninstruc.2011.05.005

Lequerica, A., Rapport, L., Axelrod, B. N., Telmet, K., & Whitman, R. D. (2002). Subjective and objective assessment methods of mental imagery control: Construct validation of self-report measures. *Journal of Clinical and Experimental Neuropsychology*, *24*(8), 1103–1116. doi:10.1076/jcen.24.8.1103.8370

Magner, U. I. E., Schwonke, R., Aleven, V., Popescu, O., & Renkl, A. (2014). Triggering situational interest by decorative illustrations both fosters and hinders learning in computer-based learning environments. *Learning and Instruction*, *29*, 141–152. doi:10.1016/j.learninstruc.2012.07.002

McGee, M. G. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin*, *86*(5), 889–918.

Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology*, *25*, 268–316. http://doi.org/10.2307/271070

Nguyen, N., Nelson, A. J., & Wilson, T. D. (2012). Computer visualizations: Factors that influence spatial anatomy comprehension. *Anatomical Sciences Education*, *5*(2), 98–108.

Pallrand, G. J., & Seeber, F. (1984). Spatial ability and achievement in introductory physics. *Journal of Research in Science Teaching*, *21*(5), 507–516.

Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg and Kuse mental rotations test-different versions and factors that affect performance. *Brain and Cognition*, *28*(1), 39–58.

Ploetzner, R., & Schlag, S. (2013). Strategic learning from expository animations: Short-and mid-term effects. *Computers & Education*, *69*, 159–168.

RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND.

Roth, W-M., & Pozzer-Ardenghi, L. (2013). Pictures in biology education. In D. F. Treagust & C.-Y. Tsui (Eds.), *Multiple representations in biological education* (pp. 39–54). Dordrecht: Springer.

Sanchez, C. A., & Wiley, J. (2014). The role of dynamic spatial ability in geoscience text comprehension. *Learning and Instruction*, *31*, 33–45. doi:10.1016/j.learninstruc.2013.12.007

Schlag, S., & Ploetzner, R. (2011). Supporting learning from illustrated texts: Conceptualizing and evaluating a learning strategy. *Instructional Science*, *39*(6), 921–937.

Schmeck, A., Mayer, R. E., Opfermann, M., Pfeiffer, V., & Leutner, D. (2014). Drawing pictures during learning from scientific text: Testing the generative drawing effect and the prognostic drawing effect. *Contemporary Educational Psychology*, *39*(4), 275–286. doi:10.1016/j.cedpsych.2014.07.003

Slof, B., Erkens, G., Kirschner, P. A., & Helms-Lorenz, M. (2013). The effects of inspecting and constructing part-task-specific visualizations on team and individual learning. *Computers & Education*, *60*(1), 221–233. doi:10.1016/j.compedu.2012.07.019

Treagust, D., Chittleborough, G., & Mamiala, T. (2003). The role of submicroscopic and symbolic representations in chemical explanations. *International Journal of Science Education*, *25*(11), 1353–1368.

Treagust, D. F., & Tsui, C-Y. (2013). Conclusion: contributions of multiple representations to biological education. In D. F. Treagust & C.-Y. Tsui (Eds.), *Multiple representations in biological education* (pp. 349–368). Dordrecht: Springer.

Tsui, C.-Y., & Treagust, D. F. (2013). Introduction to multiple representations: Their importance in biology and biological education. In D. F. Treagust & C.-Y. Tsui (Eds.), *Multiple representations in biological education* (pp. 3–18). Dordrecht: Springer.

US Department of Education. (2012). Request for applications, Education research and special education research grant programs. Retrieved from http://ies.ed.gov/funding/pdf/2013_84305A.pdf

Van Meter, P., Aleksic, M., Schwartz, A., & Garner, J. (2006). Learner-generated drawing as a strategy for learning from content area text. *Contemporary Educational Psychology*, *31*(2), 142–166.

Voyer, D. B., Cordero, T., Brake, J., Silbersweig, B., Stern, D., & Imperato-McGinley, E. J. (2006). The relation between computerized and paper-and-pencil mental rotation tasks: A validation study. *Journal of Clinical & Experimental Neuropsychology*, *28*(6), 928–939. doi:10.1080/13803390591004310

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, *101*(4), 817.

Wu, H. K., Lin, Y.-F., & Hsu, Y.-S. (2013). Effects of representation sequences and spatial ability on students' scientific understandings about the mechanism of breathing. *Instructional Science*, *41*(3), 555–573. doi:10.1007/s11251-012-9244-3