British Journal of Educational Psychology (2015), 85, 59–74

The British Psychological Society

www.wileyonlinelibrary.com

Using diagrams versus text for spaced restudy: Effects on learning in 10th grade biology classes

Bradley W. Bergey[1]*, Jennifer G. Cromley[2], Mandy L. Kirchgessner[3] and Nora S. Newcombe[3]

[1]Dalhousie University, Halifax, Nova Scotia, Canada
[2]University of Illinois at Urbana-Champaign, Illinois, USA
[3]Temple University, Philadelphia, Pennsylvania, USA

Background and Aim. Spaced restudy has been typically tested with written learning materials, but restudy with visual representations in actual classrooms is under-researched. We compared the effects of two spaced restudy interventions: A Diagram-Based Restudy (DBR) warm-up condition and a business-as-usual Text-Based Restudy (TBR) warm-up condition.

Sample. One hundred and twenty-eight consented high school students in 15 classes.

Methods. Students completed daily warm-ups over a 4-week period. Students were randomly assigned to conditions within classrooms. Warm-ups were independently completed at the start of class meetings and consisted of questions about content covered 1–10 days prior to each warm-up. Students received feedback on their answers each week. A series of ANOVAs and ANCOVAs was conducted.

Results and Conclusions. Results showed equal and significant growth from pre- to post-test for both conditions (d = .31–.67) on three outcomes: Biology knowledge, biology diagram comprehension (near transfer), and geology diagram comprehension (far transfer). ANCOVA results suggested that the magnitude of this increase was linked to the number of questions attempted during the intervention. For the DBR condition only, there were interactions with content knowledge on diagram comprehension gain scores and interactions with spatial scores on biology knowledge gain scores. Students with lower biology knowledge and lower Paper Folding Test scores were disadvantaged in the DBR condition, whereas the TBR condition was equitable across all levels of knowledge and spatial ability.

Principles of cognitive psychology can sometimes – but not always – be successfully implemented in applied classroom settings. Laboratory principles do not always hold when translated to the complex world of the classroom, where variability in motivation and time on task are the rules rather than the exception (Hulleman & Cordray, 2009). Two areas where cognitive psychologists have been actively conducting laboratory studies that have shown substantial promise for classroom applications are improving diagram comprehension (e.g., Eitel, Scheiter, Schüler, Nyström, & Holmqvist, 2013) and restudy of previously learned material using delay schedules (e.g., Carpenter, Cepeda, Rohrer, Kang, & Pashler, 2012). Bringing these two lines of research together suggests the question of

*Correspondence should be addressed to Bradley W. Bergey, Department of Psychology and Neuroscience, Dalhousie University, Life Science Centre, 1355 Oxford Street, P.O. Box 15000, Halifax, NS B3H 4R2, Canada (email: bradley.bergey@dal.ca).

DOI:10.1111/bjep.12062

whether diagrams can be used for restudy in science classrooms as effectively as conventional Text-Based Restudy (TBR). We therefore compared Diagram-Based Restudy (DBR) to TBR in 10th grade biology classrooms and looked at effects on biology knowledge, biology diagram comprehension (near transfer), and geology diagram comprehension (far transfer).

Diagram comprehension

In science classes, students see hundreds of visual representations such as line diagrams, photographs, and maps (Dimopoulos, Koulaidis, & Sklaveniti, 2005). Despite the ubiquity of diagrams in science, students often fail to understand these visual representations (e.g., Hegarty & Kozhevnikov, 1999; Kozhevnikov, Motes, & Hegarty, 2007; Wainer, 1992). For example, students do not always follow the path of a diagram correctly (Kozhevnikov *et al.*, 2007) or form an accurate mental model of the represented system or object in the diagram (Bodemer, Ploetzner, Bruchmüller, & Häcker, 2005; Kriz & Hegarty, 2007). When reading a diagram, individuals sometimes fail to identify conceptual relations between multiple representations or components of representations or do not infer necessary information that is not explicitly represented in complex diagrams (Bertin, 1983; Pinker, 1990).

Why are these representations so difficult for students to comprehend? There are several sources of difficulty, more cumulative than mutually exclusive. First, despite the importance of diagrams for comprehension of scientific text, students frequently fail to attend to diagrams when reading (Cromley, Snyder-Hogan, & Luciw-Dubas, 2010a). Second, students may not understand the conventions used in the representations. For example, novices may not be able to rely on arrows to fully comprehend the information in diagrams, even when steps in the diagram can be inferred solely from the presence of arrows (Hegarty, Kriz, & Cate, 2003; Kriz & Hegarty, 2007). Third, students may fail to notice important details or may be distracted by unimportant features (Canham & Hegarty, 2010; Schwonke, Berthold, & Renkl, 2009). Fourth, students can notice features in visual representations (e.g., that a diagram shows gas molecules moving slowly inside a container) but can fail to integrate these features into a coherent mental model (e.g., that temperature and movement of molecules are directly related to each other; Madden, Jones, & Rahm, 2011). These difficulties may also be related to characteristics of the diagram, inadequate classroom instruction, students' individual differences, or all of these factors at once.

Topic knowledge and spatial abilities as predictors of diagram comprehension

Prior research has identified prior knowledge and spatial abilities as individual differences that relate to diagram comprehension. As with reading comprehension, there is a robust relationship between knowledge about the topic depicted in the diagram and comprehension of a diagram (Butcher, 2006; Chi, Feltovich, & Glaser, 1981; Seufert & Brünken, 2006). Because diagrams represent information spatially, researchers have also found relations between various kinds of spatial abilities and diagram comprehension (Hegarty *et al.*, 2003; Hegarty & Sims, 1994; Huk & Steinke, 2007; Kozhevnikov *et al.*, 2007; Seufert & Brünken, 2006; Stieff, 2007; for a review see Höffler, 2010). Compared with participants with low spatial ability, high spatial ability participants both understand diagrams better and also learn more from instruction in diagram comprehension (Hegarty *et al.*, 2003).

Restudy as a learning technique

Restudy refers to the act of reviewing previously learned or instructed material. Restudy has often been compared to testing-for-learning methods (also known as retrieval or spaced testing; Carrier & Pashler, 1992; Carpenter & DeLosh, 2005; Roediger & Karpicke, 2006). In the current study, however, *restudy* is conceptualized as a common instructional practice in secondary education in which classroom instruction begins with a warm-up exercise that consists of activities designed to review and reinforce previously instructed material.

Restudy consistently shows advantages over no restudy, and larger benefits are generally shown for longer gaps between initial study and restudy (e.g., for spaced restudy vs. a 'cramming' session, see Arnold & McDermott, 2013; Carpenter *et al.*, 2012). Furthermore, findings from laboratory and classroom studies suggest that information becomes more easily retrieved when practice occurs on a schedule that does *not* follow the same sequence as learning (Congleton & Rajaram, 2012; McDaniel & Masson, 1985; Roediger, Agarwal, McDaniel, & McDermott, 2011). This spacing effect has been typically tested with written learning materials, but is likely to be effective with visual representations (Carpenter & Kelly, 2012; Carpenter & Pashler, 2007; Rohrer, Taylor, & Sholar, 2010). Consistent with this conclusion, we designed the sequence of restudy materials to deviate from the initial learning sequence.

Although spaced restudy is a robust method for increasing learning across laboratory and classroom contexts, the literature is limited in several ways. While spaced restudy has been tested in a variety of subject domains, most studies have been conducted with undergraduate psychology students in the laboratory (for exceptions, see Carpenter, Pashler, & Cepeda, 2009; Toppino, Kasserman, & Mracek, 1991). Only a small portion of spaced restudy investigations have been conducted in intact classrooms, and even fewer have investigated restudy with key concepts from the course material students are studying (for examples, see Carpenter *et al.*, 2009; Cranney, Ahn, McKinnon, Morris, & Watts, 2009). We know of no studies that have examined the use of spaced restudy of visual representations, despite indications in the literature that such an intervention may be effective.

To advance the literature in these areas, we sought to prompt students to examine diagrams in their textbooks as one form of restudy of previously instructed biology content and compared it to a business-as-usual TBR condition. We compared two different restudy methods – one using DBR questions and one using conventional TBR questions – and tested effects on biology knowledge (near transfer) and on biology diagram comprehension (near transfer) and geology diagram comprehension (far transfer). Based on the literature reviewed above, we implemented spaced restudy (using either diagrams or text) of key concepts from students' own curriculum, which were embedded in the regular scope and sequence of their 10th grade biology classrooms, to answer the following questions.

Research questions

RQ1: How does growth in knowledge compare between the two restudy conditions (Diagram-Based Restudy and Text-Based Restudy) when implemented in high school biology classrooms? The restudy effect suggests that practice with restudying material learned previously will increase memory for that material. As both DBR and TBR conditions ask students to work with the same information presented 1–10 days previously in class, we expect similar effects on biology knowledge for the two conditions.

RQ2: Is diagram comprehension fostered more by the Diagram-Based Restudy condition than the Text-Based Restudy condition? On the one hand, diagram skills are dependent on topic knowledge – which should increase in both conditions – so we expect growth in biology diagram comprehension for both conditions. On the other hand, the DBR condition is focused specifically on instructing diagram skills and directs students to use their textbook diagrams to answer the warm-up questions, so we expect greater gains in both biology (same domain) and geology (transfer domain) diagram comprehension for DBR than for TBR.

RQ3: Are there spatial-by-treatment interactions for diagram comprehension, whereby high Paper Folding Test scores will predict higher scores for the Diagram-Based Restudy condition but not for the Text-Based Restudy condition? Prior research suggests that high-spatial students learn more from various types of diagram instruction, including both manipulations of the text (e.g., providing text–diagram hyperlinks) and classroom instruction in diagram conventions and skills (Höffler, 2010). We therefore expect an interaction between spatial abilities and treatment whereby spatial skills will not be related to learning in the TBR condition but will be related to learning in the DBR condition.

Method

Subjects

Subjects were 128 students from a high school in the mid-Atlantic region of the United States with parental consent, drawn from 15 biology classes. Most students were in ninth grade, but some 10th graders were also enrolled in the class. The mean age of participants was 14.93 (SD = 0.66); they were 58% female, 47% White, 18% African American, 15% Hispanic, and 20% other or multiple races. With regard to socio-economic status, 86% of mothers and 84% of fathers had graduated from high school. Twenty per cent of mothers had received a bachelor's degree or higher, as had 9% of fathers. Due to absences on testing days, we have complete data on 60 in the DBR condition and 57 in the TBR condition.

Materials

The interventions consisted of a series of daily half-page warm-up activities that addressed key concepts presented in the students' biology textbooks (see Figure S1 in Supplementary Materials for examples). Warm-ups provided either a diagram decoding tip (DBR) or a quote about science (business-as-usual TBR), and all warm-ups posed two questions (one factual and one involving inference) that required students to use previously instructed information. Warm-ups asked students to use information covered in class 1–10 days earlier, and the sequence of warm-ups deviated from the order in which topics were taught; the exact sequence was the same across the two conditions. Warm-ups were designed to be completed during the first 5 min of each class period, in line with existing instructional practice at the high school (see 'Business-as-usual Text-Based Restudy' section below). As we describe below, question type, topics, and sequence were matched across conditions.

Diagram-Based Restudy Condition

For the DBR condition, we created a sequence of warm-ups that addressed key concepts presented in diagrams from the students' biology textbooks. Each warm-up in the DBR condition directed students to examine a specified diagram in their textbook to answer the questions. Each warm-up included a diagram decoding tip that explained the use and importance of a relevant diagram convention (e.g., captions, labels, arrows, and colour coding). For example, one tip reminded students that captions help readers identify the main idea expressed in a diagram and therefore are a good place to start when studying a diagram. The warm-ups asked students to answer two questions that required use of a specific diagram from the textbook: One question on each worksheet was a straightforward factual item (e.g., How many flies have red eyes in the F_1 generation [according to the diagram]?), and one question required integrating multiple pieces of information or making an inference (e.g., Can you tell the genotype for eye colour of a male *drosophila* by looking only at the eye colour? Explain your answer.).

Business-as-usual Text-Based Restudy condition

Prior to the intervention, individual teachers varied in the frequency with which they used warm-ups (daily to weekly) and the materials they used (ranging from puzzles and word games not related to instructed topics to questions from the textbook to restudy previously taught information). Thus, for the business-as-usual condition, we standardized the practice of tasking students with completing a text-based warm-up activity at the start of class meetings that related to a previously instructed topic.

For the business-as-usual TBR condition, we created a similar sequence of half page warm-ups addressing the same concepts as the DBR condition. In the TBR condition, questions were adapted from supplementary exercises from a biology textbook published by the same publisher (Holt). Thus, restudy questions were similar to but not the same as the questions in the textbook used by the classes, which could not be used for the TBR condition because these textbook questions were sometimes used for homework and other review assignments. As with the DBR condition, the TBR warm-ups used the same restudy sequence guided by the restudy principle and asked one factual question (e.g., The presence of genes found on the X or Y chromosome is called [blank]) and one inferential question (e.g., How are sex-linked traits different from non-sex-linked traits?). The warm-up directed students to look in specified paragraphs in their textbooks to answer the questions. Instead of diagram decoding tips, the TBR warm-ups included quotations from famous scientists and writers about the nature of science.

Measures

Measures were a curriculum-based test of basic biology knowledge, a biology diagram comprehension measure, a test of basic geology knowledge, a geology diagram comprehension measure, and one measure of spatial ability. All of the biology and diagram measures tested for near or far transfer, as none of the measures used stimuli or questions from the students' own textbook or the restudy exercises.

Biology knowledge

A 25-item multiple-choice biology knowledge test was developed which was closely aligned with the biology diagrams measure. Biology knowledge items measured *near* *transfer* of biology content knowledge as these items tapped basic biology concepts that were key to the course as a whole, such as adaptation to the environment, what distinguishes species from each other, and the purposes of blood in the body. Previously obtained Cronbach's alpha reliability with 143 students from prior intervention work in the same high school was .80, and the measure was strongly correlated with biology diagram comprehension (r = .63). For details and examples of the biology knowledge measure as well as the measures of biology diagram comprehension, geology knowledge, and geology comprehension described below, see Cromley, Perez, *et al.*, (2013).

Biology diagram comprehension

A 25-item multiple-choice biology diagram comprehension test was developed which asked students to use captions, labels, arrows, and other conventions to understand the main idea of a diagram, photograph, or table. Biology diagram comprehension items measured *near transfer* of diagram comprehension skills as these items tasked students with comprehending visual representations taken from a 9th grade biology textbook different from the one used in the intervention classrooms. Previously obtained Cronbach's alpha reliability with 143 students from prior intervention work in the same high school was .70, and the measure was moderately correlated with geology diagram comprehension (r = .35).

Geology knowledge

A 10-item multiple-choice geology knowledge test was developed which tapped the basic geology knowledge required to answer the geology diagram comprehension questions. Previously obtained Cronbach's alpha reliability with 143 students from prior intervention work in the same high school was .74, and the measure was moderately correlated with geology diagram comprehension (r = .31).

Geology diagram comprehension

As a measure of far transfer, a 10-item multiple-choice geology diagram comprehension test was developed which asked students to use captions, labels, arrows, and other conventions to understand the main idea of a geology line diagram or photograph. The visual representations were taken from an introductory high school geology textbook not used at the high school. Previously obtained Cronbach's alpha reliability with 143 students from prior intervention work in the same high school was .79 (for validity evidence, see correlations in Table 1).

Spatial ability

We administered the first 10 items of the Paper Folding Test (Ekstrom, French, Harman, & Derman, 1976), a 3-min group-administered paper-and-pencil measure of spatial visualization published by ETS. In this test, participants see a sequence of line drawings of a square sheet of paper with one to three folds made in it and the last drawing has a hole punched in it. Participants are asked to identify which of the five choices would match the hole-punched drawing if the sheet of paper were unfolded. This test has shown good reliability with high-school-aged (naval recruit) samples (Cronbach's alpha, .84 with

$N > 2{,}500$; Ekstrom, French, & Harman, 1979) and good concurrent validity (using the first 10 items) with spatial working memory ($r = .49$ with letter rotation and $r = .38$ with a dot matrix task for 167 undergraduate students; Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001).

Design and procedure

We used a true experimental design, where students were randomly assigned to condition within classes. Subjects were pretested on the two knowledge and two diagram comprehension measures and one spatial ability measure during a single, whole-class 45-min session approximately 2 weeks before the intervention began. Over 4 weeks (13 instructional periods), teachers distributed individually labelled warm-ups to students as they entered the classroom each period. Subjects individually and silently read the warm-up and answered the questions on the sheet. After 5 min, teachers collected the completed sheets and stored them in folders we provided. Once per week we scored and scanned the warm-ups, attached an answer key, and gave this feedback to teachers to return to the students.

Data analyses

Our primary data analyses use repeated-measures ANOVA with treatment group as the between-subjects factor and pre- and post-test scores as the within-subjects factor. Students in the two conditions did not differ in their pre-intervention biology background knowledge score, $t(109) = 1.08$, $p = .28$, biology diagram score, $t(114) = 0.45$, $p = .66$, geology background knowledge score, $t(111) = 0.22$, $p = .83$, or geology diagram score, $t(112) = 1.03$, $p = .31$. Additional analyses use ANCOVA to control for number of warm-ups attempted or quality of warm-up responses. The number of warm-ups completed controls for the extent to which students participated in the treatment, as absence or refusal to answer a question led to a smaller number of warm-ups attempted and hence less exposure to the treatment. Quality of warm-up answers controls for the extent to which students fully and accurately responded to warm-up prompts, as partial or inaccurate responses may indicate less successful learning or rehearsal of the targeted material and hence less substantive exposure to the treatment. If participation in the treatment is the reason for increasing scores from pre- to post-test, both sets of ANCOVAs should show no significant main effect of time, as variance due to time is captured by the covariate (number of warm-ups attempted or quality of warm-up responses). In all cases, the analyses met the ANCOVA assumption of equal slopes for the covariate. Due to small sample sizes and limited power, we assessed interactions with spatial ability by comparing slopes in regressions for individual differences between conditions.

Coding for quality of student answers

Two high school teachers from a Science & Math Education doctoral programme coded all student answers for quality of the answer using a previously developed rubric (Cromley, Bergey, *et al.*, 2013). The coders were trained on student work products from a prior study, and then, each coder scored every one of the 3556 answers; they disagreed on only 47 scores, for an agreement of 99%. Disagreements were resolved by discussion between the two raters.

Results

Descriptive statistics on and intercorrelations among all variables are reported in Table 1. The mean number of daily warm-ups attempted was 9.88 (SD = 2.27) of the 13 assigned by teachers, with no significant difference between groups, $t(118) = 0.450$, $p = .65$.

Main effects

Mean pre- and post-test scores by treatment group are shown in Figure 1.

Biology knowledge

A repeated-measures ANOVA on biology knowledge scores from pre- to post-test showed a significant effect of time, $F(1, 104) = 24.52$, $MSE = 2.32$, $p < .001$, a non-significant effect of treatment, $F(1, 104) = 2.03$, $MSE = 7.85$, $p = .16$, and a non-significant time × treatment interaction, $F(1, 104) = 0.62$, $MSE = 2.32$, $p = .43$. Both groups grew equally and significantly over time, d = .51 and .46, respectively, for TBR and DBR.

Biology diagrams

A repeated-measures ANOVA on biology diagram scores from pre- to post-test showed a significant effect of time, $F(1, 109) = 16.81$, $MSE = 2.80$, $p < .001$, a non-significant effect of treatment, $F(1, 109) = 2.54$, $MSE = 4.77$, $p = .11$, and a non-significant time × treatment interaction, $F(1, 109) = 2.01$, $MSE = 2.80$, $p = .16$. Subjects in both conditions grew equally and significantly over time, d = .31 and .63, respectively, for TBR and DBR.

Geology diagrams

A repeated-measures ANOVA on geology diagram scores from pre- to post-test showed a significant effect of time, $F(1, 107) = 23.62$, $MSE = 1.83$, $p < .001$, a non-significant effect of treatment, $F(1, 107) = 1.67$, $MSE = 2.97$, $p = .20$, and a non-significant time × treatment interaction, $F(1, 107) < 0.01$, $MSE = 1.83$, $p = .99$. Subjects in both conditions grew equally and significantly over time, d = .67 and .59, respectively, for TBR and DBR.

Table 1. Intercorrelations among variables

Variable	1	2	3	4	5	6	7	8
1. Biology knowledge (Pre)	–							
2. Geology knowledge (Pre)	.39*							
3. Biology diagrams (Pre)	.25*	.22*						
4. Geology diagrams (Pre)	.15	.27*	.26*					
5. Biology knowledge (Post)	.55*	.34*	.18*	.18*				
6. Biology diagrams (Post)	.15	.24*	.26*	.28*	.33*			
7. Geology diagrams (Post)	.22*	.29*	.10	.25*	.32*	.20*		
8. Paper Folding Task	.35*	.27*	.21*	.08	.28*	.15	.22*	–
M	5.41	4.05	5.59	2.57	6.34	6.46	3.50	4.42
SD	2.18	1.49	1.93	1.43	2.37	1.96	1.63	2.01
Cronbach's α	.88	.84	.88	.82	.94	.96	.88	.86

Note. Pre, before intervention; post, after intervention. *$p < .05$.

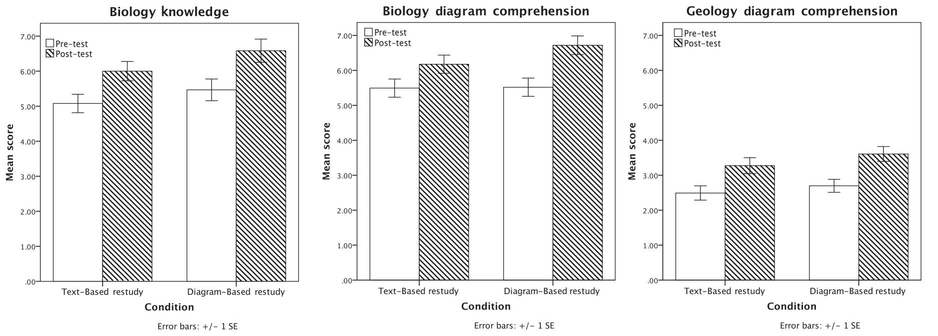

Figure 1. Mean Pre- and Post-test Scores by Treatment Group. Maximum score for biology diagram knowledge and biology diagram comprehension was 25. Maximum score for geology diagram comprehension was 10.

Analysis of covariance

Number of warm-ups completed as a covariate

A repeated-measures ANCOVA on biology knowledge adding the number of warm-ups completed as a covariate showed no significant effect of time, $F(1, 109) = 1.11, p = .29$, $MSE = 2.35$, no significant effect of treatment, $F(1, 109) = 1.68$, $MSE = 7.88$, $p = .20$, and a non-significant time × treatment interaction, $F(1, 109) = 0.22$, $MSE = 2.35$, $p = .64$. As expected, entering number of warm-ups attempted as a covariate removed the main effect of time. Growth in biology knowledge scores was associated with completion of warm-ups, not simply with the passage of time.

A repeated-measures ANCOVA on biology diagram comprehension adding the number of warm-ups completed as a covariate showed no significant effect of time, $F(1, 108) = 0.50$, $MSE = 2.83$, $p = .48$, no significant effect of treatment, $F(1, 108) = 2.47$, $MSE = 4.77$, $p = .12$, and a non-significant time × treatment interaction, $F(1, 108) = 1.99$, $MSE = 2.83$, $p = .16$. As expected, entering number of warm-ups attempted as a covariate removed the main effect of time. Growth in biology diagram comprehension scores was associated with completion of warm-ups, not simply with the passage of time.

A repeated-measures ANCOVA on geology diagram comprehension adding the number of warm-ups completed as a covariate showed no significant effect of time, $F(1, 106) = 1.75$, $MSE = 3.21$, $p = .19$, no significant effect of treatment, $F(1, 106) = 1.63$, $MSE = 4.89$, $p = .20$, and a non-significant time × treatment interaction, $F(1, 106) = .001$, $MSE = .002$, $p = .97$. As expected, entering number of warm-ups attempted as a covariate removed the main effect of time. Growth in geology diagram comprehension scores was associated with completion of warm-ups, not simply with the passage of time.

Quality of warm-up answers as a covariate

In a repeated-measures ANCOVA on biology knowledge adding the quality of warm-up answers out of attempted answers as a covariate, there were no significant main effect of time, $F(1, 103) = 0.32$, $MSE = 0.73$, $p = .58$, no significant effect of treatment, $F(1, 103) = 3.11$, $MSE = 23.67$, $p = .08$, and no significant time × treatment interaction, $F(1, 103) = 0.78$, $MSE = 1.83$, $p = .78$. As expected, entering quality of warm-up responses as a covariate removed the main effect of time. Growth in biology knowledge was associated with the quality of warm-up responses, not simply with the passage of time.

In a repeated-measures ANCOVA on biology diagram comprehension adding the quality of warm-up answers out of attempted answers as a covariate, the significant effect of time was maintained, $F(1, 108) = 5.29$, $MSE = 2.82$, $p = .02$, and there were no significant effect of treatment, $F(1, 108) = 3.66$, $MSE = 4.46$, $p = .06$, and no significant time $\times$ treatment interaction, $F(1, 108) = 5.03$, $MSE = 2.82$, $p = .18$. Contrary to expectations, entering the quality of warm-up answers did not remove the main effect of time.

In a repeated-measures ANCOVA on geology diagram comprehension adding the quality of warm-up answers out of attempted answers as a covariate, the significant effect of time was maintained, $F(1, 106) = 8.84$, $MSE = 15.77$, $p < .01$, and there were no significant effect of treatment, $F(1, 106) = 1.94$, $MSE = 5.77$, $p = .17$, and no significant time $\times$ treatment interaction, $F(1, 106) = .07$, $MSE = .12$, $p = .80$. Contrary to expectations, entering the quality of warm-up answers did not remove the main effect of time.

Knowledge × treatment and spatial × treatment interactions

We regressed post-test scores on pre-test scores to create a residualized gain score for biology knowledge, biology diagram comprehension, and geology diagram comprehension. Residualized gain scores quantify the extent to which individual scores change from pre- to post-test after accounting for mean change for the group. Residualized gain scores are useful because they preserve individual variability in growth scores. Results from separate regressions of residualized gain scores on knowledge and on spatial scores split by conditions are shown in Table 2. Results indicate that DBR generally produced better results for higher-knowledge and higher-spatial students. For example, students with higher biology knowledge scores in the DBR condition grew more in biology diagram comprehension than did students with lower biology knowledge. Similarly, students with lower spatial scores in the DBR condition were disadvantaged on biology knowledge and geology diagrams. The TBR condition, by contrast, was equitable for students across a range of knowledge and spatial scores.

Table 2. *R*-Squared values for regressions of residualized gain scores on knowledge and spatial predictors

	On biology knowledge score gains		On biology diagram score gains		On geology diagram score gains	
Predictor	DBR	TBR	DBR	TBR	DBR	TBR
Background knowledge	n/a	n/a	.29*	.21	.34*	.08
From PFT	.29*	<.01	.08	.16	.25†	.08

Note. Values in the table signify R^2 values for each regression; gain scores are unstandardized *b*'s created by regressing post-scores on pre-scores; background knowledge = post-intervention knowledge. DBR, Diagram-Based Restudy; TBR, Text-Based Restudy; PFT, Paper Folding Task *significant at $p < .05$; †significant at $p < .10$.

Discussion

This study combined principles of learning from diagrams with principles of restudy of previously learned material using delay schedules to compare the effectiveness of restudy with diagram-based or text-based review materials. Using two 4-week long restudy treatments with feedback implemented *in situ* in 10th grade biology classes, we showed equal, significant, and moderate- to large-sized effects on biology knowledge (d = .46–.51), biology diagram comprehension (near transfer; d = .31–.63), and geology diagram comprehension (far transfer; d = .59–.67). Covarying out the number of warm-ups attempted showed that warm-up completion, not just the passage of time, was associated with increases in biology knowledge and biology diagrams. However, covarying out the score on quality of answers did not show any effects. Whether delivered in a text-based format or a diagram-based format, restudy warm-ups were associated with significant growth in biology knowledge.

Our results are consistent with the many laboratory studies (e.g., Arnold & McDermott, 2013; Carpenter *et al.*, 2012) and a small number of classroom studies (Carpenter *et al.*, 2009) that have examined the restudy effect. Our results extend prior literature on restudy by demonstrating effects on transfer measures (i.e., uninstructed diagrams) and by replicating findings in the context of high school science classrooms with actual course concepts and materials. Our findings for DBR are also consistent with findings from laboratory-based delayed retrieval studies that used visual representations (Carpenter & Kelly, 2012; Carpenter & Pashler, 2007; Rohrer *et al.*, 2010), and expand this literature to restudy from images. Consistent with this research, our warm-ups required students to use information they had learned previously, and the act of restudy led to the formation of stronger memory traces for key concepts. Furthermore, by modifying existing instructional practices to include restudy with visual representations, our findings demonstrate the feasibility of applying these principles with minimal teacher training and minimal disruption to classroom instruction. It is possible that restudy via visual representation allowed for variety in types of memory traces that could be retrieved.

Contrasting results from our two analyses that covaried learning process variables (number of questions attempted and correctness of attempted questions) provide important insights into boundary features that make restudy effective. Specifically, more *attempts* at restudy were associated with more gain in biology knowledge, biology diagram comprehension, and geology diagram comprehension from both DBR and TBR treatments. Similarly, the correctness of warm-up responses was associated with gains in biology knowledge. By contrast, students did not need to have highly correct answers on warm-up attempts in order to gain in biology diagram comprehension and geology diagram comprehension from the treatments. This is consistent with findings from the retrieval literature that even retrieving incorrect information can lead to better memory for correct information, presumably because the correct information is semantically linked to the (now highly activated) incorrect information (Grimaldi & Karpicke, 2012).

Why would restudy practice with diagrams lead to increases in both knowledge and diagram comprehension? First, prior research suggests that students often skip diagrams entirely (Cromley *et al.*, 2010a). The DBR warm-up questions were only answerable by looking at the diagrams, thereby possibly giving students additional opportunities to learn the content. Second, the warm-ups in the DBR and TBR conditions included questions that required students to combine different pieces of information and make inferences. The DBR treatment therefore fostered inferential processing via the diagrams. Third, the relation between biology knowledge and increases in biology diagram scores suggests that the DBR treatment especially demands that students draw on prior knowledge, to the extent they have this knowledge. While prior research has established that topic knowledge is vital for both comprehension of written text (Cromley & Azevedo, 2007; Cromley, Snyder-Hogan, & Luciw-Dubas, 2010b; Tarchi, 2010) and diagram comprehension (Cromley, Bergey, *et al.*, 2013; Cromley, Perez, *et al.*, 2013; Münzer, Seufert, & Brünken, 2009), knowledge appears to have particularly strong effects when learners attempt to make sense of diagrams. Fourth, practice with diagrams in the DBR condition seems to be associated with a larger effect size for growth in biology diagram comprehension, although this difference is non-significant with our relatively small sample.

Interactions with spatial ability have not been investigated to date in classroom restudy research. Similar to basic laboratory research with diagrams, the spatial skills measured by the Paper Folding Task (PFT) were related to diagram comprehension for these high school students. Specifically, our measure of spatial visualization was significantly related to knowledge gains and marginally related to geology diagram comprehension gains, but only for the DBR condition where students used visual representations to retrieve the previously covered biology content. Paper Folding Test scores were not related to the pre-test geology diagram scores by themselves, but were related to the shift from pre-test to post-test – the higher the PFT score, the greater the shift in geology diagram comprehension within the DBR condition. The spatial visualization skill tapped by the PFT appears to enable students to gain more diagram comprehension skills from the restudy-with-diagram tasks. One possibility is that gaining scientific knowledge from the diagrams for the covered topics – such as structure of DNA and protein synthesis – is best understood if one can mentally rotate the structures, such as imagining the DNA strand unfurling during replication, and this mental rotation of a 3D object depicted in 2D is common to the PFT and the diagrams in the DBR condition.

Limitations

We note several limitations to our study. First, students restudied concepts only twice – although they did answer a total of four questions for each key concept – whereas for retrieval, multiple opportunities to retrieve are associated with better learning (Carpenter *et al.*, 2009). Second, our intervention was brief, constituting 5–7 min per day repeated for approximately 10 warm-ups per student, which represents 50–70 min of intervention spread over 4 weeks of class time. Third, we were not able to provide immediate feedback to students; feedback was provided once per week. The literature suggests that more immediate feedback on the correctness of retrieved information might produce stronger learning effects. However, Carpenter and Kelly (2012) found that in delayed retrieval with visual representations, retest-only and retest-with-feedback conditions led to equally superior performance compared to retrieval only. Fourth, we interpreted the ANCOVA results as an indication that involvement in the intervention was responsible for growth, not simply the passage of time. However, the effects of involvement in the intervention may be confounded with other factors that may have influenced the number of warm-ups completed, such as student motivation.

Fifth, embedding the intervention in the context of typical high school science classrooms required that the central experimental manipulation – DBR versus TBR – involved additional differences in warm-up instructions and questions. While warm-up characteristics were matched across conditions, such as balancing questions by type (open vs. forced choice) and difficulty (factual vs. inferential) and focusing questions on the same concept in the same sequence across the two conditions, diagram-based and text-based questions could not be identical. Minor differences in the material presented in the diagrams and text required that questions differ slightly in their content. In addition, diagram warm-ups included diagram decoding tips to scaffold students' comprehension of diagrams, while text-based warm-ups did not include comparable comprehension tips. Diagram decoding tips were included for the DBR condition based on prior literature that indicated that diagrams are often misunderstood (Canham & Hegarty, 2010; Hegarty *et al.*, 2003) and comprehension of science diagrams can be improved by providing high school students with tips about the conventions of diagrams (Cromley, Perez, *et al.*, 2013). Because the goal of the study was to compare DBR to business-as-usual TBR, we opted not to include text-based comprehension tips, as such tips would not be typically included in restudy activities. As a result, differences in questions or the presence or absence of comprehension tips may have contributed to differential gains for the two conditions.

Finally, while the warm-up instructions directed students to locate either text or diagrams, we could not prevent students in the DBR treatment from looking at text – although the questions were only answerable from the diagram – nor could we prevent students in the TBR treatment from looking at diagrams. The similar results could be due to bleed-over from one condition to the other. Collecting student discourse during learning could provide insights into the extent to which students are actually using diagrams whether they are prompted (DBR) or not prompted (TBR) to do so.

Conclusion

Brief TBR or DBR warm-ups can be a low-cost but powerful aid for learning high school biology. Unlike end-of-chapter exercises, restudy warm-ups required students to revisit and use previously learned material on an uneven schedule, leading to better memory for the studied content. The DBR condition posed questions that could only be answered with the diagrams, and for some students, this may have been the first time they inspected the diagram. While both restudy conditions were effective, the differential effects of knowledge and spatial ability in the DBR condition might complicate the picture for implementing these interventions in classrooms, raising questions for educators and researchers: Should the diagram-based condition only be used with students with higher knowledge or spatial skills? How can students with low knowledge or spatial skills be effectively supported to learn from the diagrams in their textbooks? Future research should try to replicate these findings and might also consider whether extended practice with diagram instruction could itself increase spatial ability. Further developing the diagram-based condition could capitalize on its strengths – especially for increasing both biology and geology diagram comprehension – while helping us better understand which students can benefit most from it.

Acknowledgements

The research reported here was supported by funds from NSF award #0814901; the opinions stated here are the authors' own and do not represent the opinions of the agency. We thank the students who participated in the study, their parents, teachers, and school administrators. We thank John Romano, Steven Peterson, Jonathan Howe, Dar Sheth, and Reham Raja for assistance with scoring and data entry.

References

Arnold, K. M., & McDermott, K. B. (2013). Test-potentiated learning: Distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 940–945. doi:10.1037/a0029199

Bertin, J. (1983). *Semiology of graphics (W Berg Trans.)*. Madison, WI: University of Wisconsin Press.

Bodemer, D., Ploetzner, R., Bruchmüller, K., & Häcker, S. (2005). Supporting learning with interactive multimedia through active integration of representations. *Instructional Science*, *33*(1), 73–95. doi:10.1007/s11251-004-7685-z

Butcher, K. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology*, *98*(1), 182–197. doi:10.1037/0022-0663.98.1.182

Canham, M., & Hegarty, M. (2010). Effects of knowledge and display design on comprehension of complex graphics. *Learning and Instruction*, *20*(2), 155–166. doi:10.1016/j.learninstruc.2009.02.014

Carpenter, S. K., Cepeda, M. J., Rohrer, D., Kang, S. H., & Pashler, H. (2012). Using spacing to enhance diverse forms of learning: Review of recent research and implications for instruction. *Educational Researcher*, *24*, 369–378. doi:10.1007/s10648-012-9205-z

Carpenter, S. K., & DeLosh, E. L. (2005). Application of the testing and spacing effects to name-learning. *Applied Cognitive Psychology*, *19*, 619–636. doi:10.1002/acp.1101

Carpenter, S. K., & Kelly, J. W. (2012). Tests enhance retention and transfer of spatial learning. *Psychonomic Bulletin and Review*, *19*, 443–448. doi:10.3758/s13423-012-0221-2

Carpenter, S. K., & Pashler, H. (2007). Testing beyond words: Using tests to enhance visuospatial map learning. *Psychonomic Bulletin and Review*, *14*, 474–478. doi:10.3758/BF03194092

Carpenter, S. K., Pashler, H., & Cepeda, N. J. (2009). Using tests to enhance 8th grade students' retention of US history facts. *Applied Cognitive Psychology*, *23*, 760–771. doi:10.1002/acp.1507

Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. *Memory & Cognition*, *20*, 633–642. doi:10.3758/bf03202713

Chi, M., Feltovich, P. J., & Glaser, R. (1981). Categorization and representations of physics problems by experts and novices. *Cognitive Science*, *5*(2), 121–152. doi:10.3758/BF03202713

Congleton, A., & Rajaram, S. (2012). The origin of the interaction between learning method and delay in the testing effect: The roles of processing and conceptual retrieval organization. *Memory & Cognition*, *40*, 528–539. doi:10.3758/s13421-011-0168-y

Cranney, J., Ahn, M., McKinnon, R., Morris, S., & Watts, K. (2009). The testing effect, collaborative learning, and retrieval-induced facilitation in a classroom setting. *European Journal of Cognitive Psychology*, *21*, 919–940. doi:10.1080/09541440802413505

Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, *99*, 311–325. doi:10.1037/0022-0663.99.2.311

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

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

Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010b). Reading comprehension of scientific text: A domain-specific test of the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, *102*, 687–700. doi:10.1037/a0019452

Dimopoulos, K., Koulaidis, V., & Sklaveniti, S. (2005). Towards a framework of socio-linguistic analysis of science textbooks: The Greek case. *Research in Science Education*, *35*, 173–195. doi:10.1007/s11165-004-8162-z

Eitel, A., Scheiter, K., Schüler, A., Nyström, M., & Holmqvist, K. (2013). How a picture facilitates the process of learning from text: Evidence for scaffolding. *Learning and Instruction*, *28*, 48–63. doi:10.1016/j.learninstruc.2013.05.002

Ekstrom, R. B., French, J. W., & Harman, H. H. (1979). Cognitive factors: Their identification and replication. *Multivariate Behavioral Research Monographs*, *79*(2), 3–84.

Ekstrom, R. B., French, J. W., Harman, H., & Derman, D. (1976). *Kit of factor-referenced cognitive tests* (Rev ed.). Princeton, NJ: Educational Testing Service.

Grimaldi, P. J., & Karpicke, J. D. (2012). When and why do retrieval attempts enhance subsequent encoding? *Memory & Cognition*, *40*, 505–513. doi:10.3758/s13421-011-0174-0

Hegarty, M., & Kozhevnikov, M. (1999). Types of visual-spatial representations and mathematical problem solving. *Journal of Educational Psychology*, *91*, 684–689. doi:10.1037/0022-0663.91.4.684

Hegarty, M., Kriz, S., & Cate, C. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition and Instruction*, *21*, 325–360. doi:10.1207/s1532690xci2104_1

Hegarty, M., & Sims, V. (1994). Individual differences in mental animation during mechanical reasoning. *Memory & Cognition*, *22*, 411–430. doi:10.3758/BF03200867

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

Huk, T., & Steinke, M. (2007). Learning cell biology with close-up views or connecting lines: Evidence for the structure mapping effect. *Computers in Human Behavior*, *23*, 1089–1104. doi:10.1016/j.chb.2006.10.004

Hulleman, C. S., & Cordray, D. S. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness*, *2*, 88–110. doi:10.1080/19345740802539325

Kozhevnikov, M., Motes, M. A., & Hegarty, M. (2007). Spatial visualization in physics problem solving. *Cognitive Science: A Multidisciplinary Journal*, *31*, 549–579. doi:10.1080/15326900701399897

Kriz, S., & Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *International Journal of Human-Computer Studies*, *65*, 911–930. doi:10.1016/j.ijhcs.2007.06.005

Madden, S. P., Jones, L. L., & Rahm, J. (2011). The role of multiple representations in the understanding of ideal gas problems. *Chemistry Education Research and Practice*, *12*, 283–293. doi:10.1039/C1RP90035H

McDaniel, M. A., & Masson, M. E. (1985). Altering memory representations through retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 371–385. doi:10.1037/0278-7393.11.2.371

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent variable analysis. *Journal of Experimental Psychology: General*, *130*, 621–640.

Münzer, S., Seufert, T., & Brünken, R. (2009). Learning from multimedia presentations: Facilitation function of animations and spatial abilities. *Learning and Individual Differences*, *19*, 481–485. doi:10.1016/j.lindif.2009.05.001

Pinker, S. (1990). A theory of graph comprehension. In R. Freedle & R. Freedle (Eds.), *Artificial intelligence and the future of testing* (pp. 73–126). Hillsdale, NJ: Lawrence Erlbaum.

Roediger, H., Agarwal, P. K., McDaniel, M. A., & McDermott, K. B. (2011). Test-enhanced learning in the classroom: Long-term improvements from quizzing. *Journal of Experimental Psychology: Applied*, *17*, 382–395. doi:10.1037/a0026252

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*, 249–255. doi:10.1111/j.1467-9280.2006.01693.x

Rohrer, D., Taylor, K., & Sholar, B. (2010). Tests enhance the transfer of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(1), 233–239. doi:10.1037/a0017678

Schwonke, R., Berthold, K., & Renkl, A. (2009). How multiple external representations are used and how they can be made more useful. *Applied Cognitive Psychology*, *23*, 1227–1243. doi:10.1002/acp.1526

Seufert, T., & Brünken, R. (2006). Cognitive load and the format of instructional aids for coherence formation. *Applied Cognitive Psychology*, *20*, 321–331. doi:10.1002/acp.1248

Stieff, M. (2007). Mental rotation and diagrammatic reasoning in science. *Learning and Instruction*, *17*, 219–234. doi:10.1016/j.learninstruc.2007.01.012

Tarchi, C. (2010). Reading comprehension of informative texts in secondary school: A focus on direct and indirect effects of reader's prior knowledge. *Learning and Individual Differences*, *20*, 415–420. doi:10.1016/j.lindif.2010.04.002

Toppino, T. C., Kasserman, J. E., & Mracek, W. A. (1991). The effect of spacing repetitions on the recognition memory of young children and adults. *Journal of Experimental Child Psychology*, *51*(1), 123–138. doi:10.1016/0022-0965(91)90079-8

Wainer, H. (1992). Understanding graphs and tables. *Educational Researcher*, *21*(1), 14–23. doi:10.3102/0013189X021001014

Received 1 May 2014; revised version received 24 October 2014

Supporting Information

The following supporting information may be found in the online edition of the article:

Figure S1. Examples of Text-Based (Left) and Diagram-Based (Right) Warm-ups.