



Spatial supports for comparison in educational science images

Benjamin D. Jee¹ · Bryan J. Matlen² · Monica Greenlaw¹ · Nina Simms³ · Dedre Gentner³

Received: 5 March 2021 / Accepted: 20 July 2022
© The Author(s) 2022

Abstract

Images, such as photographs and diagrams, play an important role in the teaching and learning of science. To optimize student learning, educational science images should be designed to facilitate the cognitive processes relevant to comprehension. One such process is comparison, which involves aligning multiple representations on the basis of their common relational structure. This structural alignment process can be facilitated by cognitive supports that are inherent to an image, including its spatial layout. Yet, little is known about the extent to which students must engage in comparison to learn from science images, and whether widely-used educational materials are conducive to structural alignment. To address these issues, we sampled multiple chapters from each of three popular U.S. middle school life science textbooks. We coded each image for the presence of prompts for comparison using cues within the images and surrounding text. For each image that prompted comparison, we coded whether its layout facilitated relevant structural alignment (*direct* placement of matched pairs) or obscured alignment (*impeded* placement). Overall, we found that comparisons were prompted for more than a third of the images. However, fewer than half of the images that required comparison had a spatial layout that provided strong support for comparison—that is, direct placement of matched objects/parts. We propose that, in concert with other cognitive supports for learning from multiple representations, spatial supports for comparison could be applied broadly to increase the effectiveness of educational science images.

Keywords Comparison · Structural alignment · Science learning · Education · Textbooks · Images · Multiple representations · Cognitive supports

✉ Benjamin D. Jee
bjee@worchester.edu

¹ Department of Psychology, Worcester State University, 486 Chandler St, Worcester, MA 01602, USA

² WestEd, San Francisco, CA, USA

³ Department of Psychology, Northwestern University, Evanston, IL, USA

Introduction

Images—photos, illustrations, diagrams, etc.—play a central role in the teaching and learning of science (e.g., Ainsworth, 2006; Ainsworth et al., 2011; Castro-Alonso & Uttal, 2019; Forbus et al., 2011; Jee et al., 2014; Newcombe, 2017). Compared to text alone, learning is consistently superior when text is accompanied by pictures, the *multimedia principle* (Mayer, 2021). Yet, interpreting and comprehending science images can be challenging for students (Rau, 2017). In many cases, explanatory images in textbooks are not accompanied by written explanations, or the relevant text is not in the same location as the image (Betrancourt et al., 2012; Nyachwaya et al., 2016; Slough et al., 2010). When studying unfamiliar diagrams, novice learners tend to overlook key information (Jee et al., 2014), and are more likely to experience cognitive overload (Chen et al., 2017; Sweller, 2010). Less-experienced students are also more likely to misinterpret visual conventions—such as before-and-after, and part-whole—that appear in instructional images (Boucheix et al., 2020). To optimize student learning, educational materials should be designed to facilitate the cognitive processes relevant to comprehension (Kintsch, 1994; Mayer, 2021). In this paper we consider the extent to which educational science images support one such process, *comparison*.

Comparison and structural alignment in science learning

To comprehend a scientific image, a student often needs to determine the relationships that link its various parts (Ainsworth, 2006; Boucheix et al., 2020; Jee et al., 2010). Comparing across examples can help learners zero in on conceptually-relevant information, such as geological structures in photos (Jee et al., 2013), diseases in x-rays (Kok et al., 2013), anomalies in skeletons (Kurtz & Gentner, 2013), and features of buildings that reflect elementary engineering principles (Gentner et al., 2016). Comparison of examples also facilitates the learning of concepts defined by common relations, such as *catalyst* or *adaptation*, which are pervasive in science (Goldwater & Schalk, 2016).

According to Structure-mapping theory, comparison involves aligning two examples in terms of their common relational structure, a process of *structural alignment* in which elements of the examples are placed into correspondence according to their role in a common system of relations (Falkenhainer et al., 1989; Gentner, 1983, 2010; Gentner & Markman, 1997). As an illustration, Fig. 1 is a textbook diagram that shows the bones in the forelimbs of several kinds of animals. To glean the important information, the student must align the spatial-relational patterns across the forelimbs of the different animals, placing the individual bones in correspondence according to their roles in the shared structure. Thus, the human humerus aligns with the cat humerus, the cat ulna with the bat ulna, and so on (these alignments are supported by the colors in the image, as well as the spatial layout). Even though the humerus, ulna, and radius vary in size and shape across the five animals, carrying out a structural alignment across the different forelimbs highlights their common structure. This alignment process also paves the way for further inferences. In this case, the similar structure of the limbs (termed *homology* in the biological sciences) suggests that these organisms share a common ancestor in evolutionary history.

Structural alignment also highlights differences that are connected to a shared system of relations (Markman & Gentner, 1993a, 1993b; Sagi et al., 2012); for example, the greater relative length of the ‘finger’ bones in the bat than in the human. More broadly, comparison

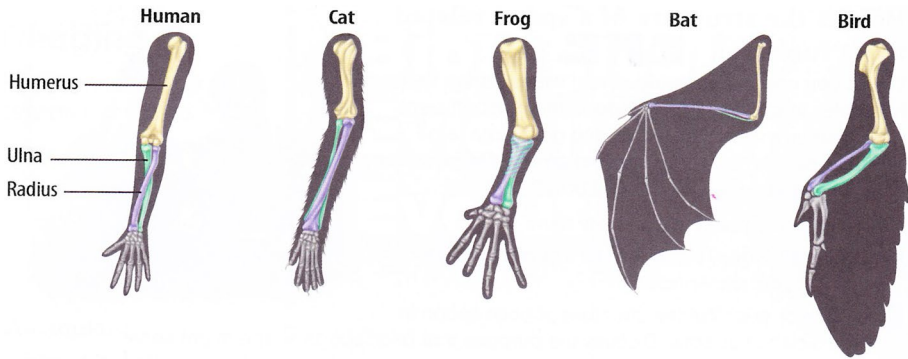


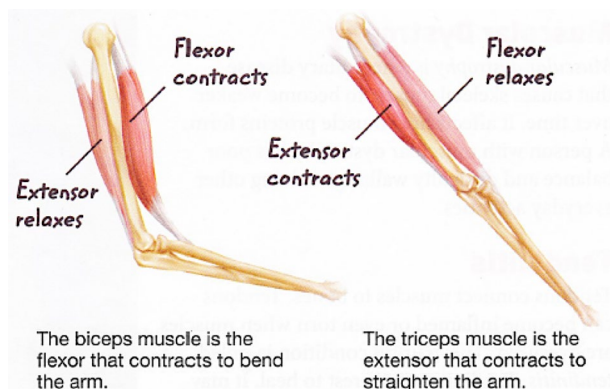
Fig. 1 Structure diagram showing bones of forelimbs of different animals. *Note* Illustration from McGraw Hill Integrated iScience Course 2, Student Edition. Copyright © 2016. All rights reserved. Reprinted by permission of McGraw Hill Education

promotes the selection, organization, and integration of relevant information—key aspects of meaningful learning—both from text materials (Goldwater & Gentner, 2015; Sweller, 2010) and from multimedia (Mayer, 2021; Schnotz, 2005; Schnotz & Bannert, 2003; Seufert & Brünken, 2006). Yet, comparison is effective only if students create accurate structural alignments (e.g., Kurtz et al., 2001). This requires the mapping of correspondences on the basis of shared relationships, while rejecting correspondences that might be suggested by more superficial details, such as local perceptual matches. For example, to understand how flexor and extensor muscles cooperate to bend and straighten an arm (see Fig. 2), a student must align the two flexor muscles across the two examples. This alignment must be based on shared structural roles, and avoid the incorrect matches suggested by perceptual similarity.

Spatial supports for comparison

Students often require assistance to select and cognitively integrate the relevant information in a lesson. One effective method is to explicitly teach students about the deep, relational connections between two representations (Seufert & Brünken, 2006). Similarly, prompting students to “self-explain” pairs of related images can encourage deeper, conceptual

Fig. 2 Structure diagram showing muscle activity and arm movement. *Note* Illustration from Houghton Mifflin Harcourt SCIENCE FUSION: The Human Body, Student Edition. Copyright © 2017. All rights reserved. Reprinted by permission of Houghton Mifflin Harcourt Publishing Company



learning (Rau et al., 2015). Students may also benefit from general training that teaches them how to cognitively process and integrate multiple representations (Bodemer et al., 2005), though lower-knowledge students may lack the conceptual knowledge to benefit from such training (Rau, 2018; Seufert, 2019). Indeed, inexperienced students may require relatively high levels of cognitive support, because they cannot rely on prior knowledge to make accurate inferences (Kintsch, 1994; Vosniadou & Skopeliti, 2017) or to reduce cognitive load during learning (Chen et al., 2017; Kalyuga, 2007).

With comparison in particular, novice students benefit when an instructor indicates the correspondences between examples through the use of language and/or gestures (e.g., Gentner & Rattermann, 1991; Gentner et al., 2011, 2016; Jee & Anggoro, 2019; Richland et al., 2007; Rittle-Johnson et al., 2019; Yuan et al., 2017). Yet, these helpful cues are seldom available when students engage in self-paced or out-of-classroom learning. In these contexts, cognitive supports embedded within the instructional materials are especially critical.

A number of *spatial* factors have been found to support student learning from images and multimedia materials. When the structure of a diagram aligns with the intended mental model of the depicted system—a “task-appropriate” image—comprehension is enhanced relative to a structurally-incompatible diagram (Schnotz et al., 2003). When pairs of science images are shown, comparison is faster and more accurate when the orientations of the images are the same as opposed to different (Kurtz & Gentner, 2013). In addition, people tend to learn more when corresponding words and pictures are displayed in close proximity to one another—the *spatial contiguity principle* of multimedia learning (Mayer, 2021; Moreno & Mayer, 1999). Spatial contiguity reduces the burden of dividing attention between separate sources of information, and facilitates cognitive integration (Ayres & Sweller, 2014; Mayer & Moreno, 2003; Schroeder & Cenkcı, 2018).

In the present work, we focus on a spatial factor that can affect the ease of structural alignment involving images: the *spatial placement* of the visual units—objects, parts, etc.—that are being compared (Kurtz & Gentner, 2013; Matlen et al., 2011, 2020). (We further define “visual unit” and other key terms in the section “Methods”). To illustrate spatial placement, Fig. 3 shows different layouts for the same pairs of units: shape triplets, box plots, and skeletal structures. In the top row, the visual units have horizontal axes, in the bottom row, vertical. In each case, the unit in the upper left can be paired with the unit directly below it, and with the one on its right. For the horizontal units, the correspondences between each pair (represented by connecting lines in Fig. 3) are clear and direct when the layout is vertical, with one unit above the other. We refer to this as a *direct placement* of the paired units.

In direct placement, the paths between matched pairs of visual units are clear and unobstructed. This arrangement minimizes competing matches, facilitating the intended alignment and reducing the chances of an incorrect alignment. In contrast, the alignment of the horizontal units is difficult when the layout is horizontal, with units side by side, because finding the optimal correspondences is impeded by nonmatching elements (thus the crossed connecting lines). We refer to this as an *impeded placement*. For the vertical units, the opposite is true: a horizontal layout constitutes a direct placement, and a vertical layout an impeded placement.

Prior research indicates that comparison is not only faster, but is consistently more accurate when placement is direct as opposed to impeded (Matlen et al., 2020). Substantial improvements in the accuracy of same/different judgments have been found with shape triplet pairs like those in Fig. 3 ($d = .49$; Matlen et al., 2020, Experiment 4), and with purely relational pairs, such as *square-square-circle* to *blue-blue-red* (Matlen et al., 2020). The

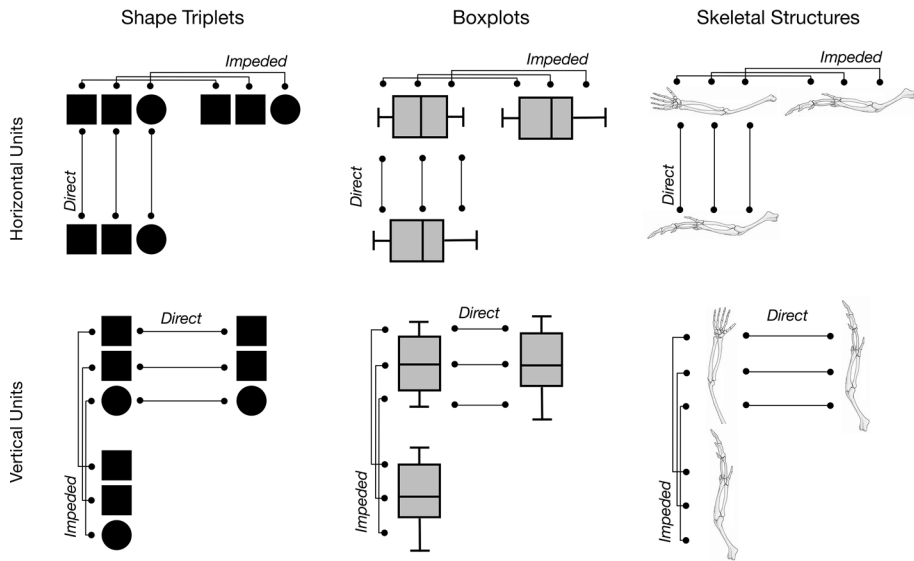


Fig. 3 Examples of different spatial placements for pairs of units with horizontal (top row) and vertical axes (bottom row). *Note* Lines indicate correspondences for direct and impeded placements

benefit of direct placement for the comparison of figures extends to young children, aged 6 and 8 (Zheng et al., 2020, in press), and to middle schoolers' and adults' performance on a comparison task involving complex images, similar to the skeletal structures in Figs. 1 and 3 (Simms et al., 2019).

Whereas direct placement facilitates structural alignment, impeded placement makes it more difficult. Compared to a baseline condition in which paired units were laid out diagonally, impeded placement led to slower and less accurate comparisons (Matlen et al., 2020). The root of the difficulty is that impeded placement introduces irrelevant competing element matches that interfere with structural alignment. Putting up a solid visual barrier between a matched pair does not counteract the benefits of direct placement¹; however, introducing a third structured example between the pair results in slower and less accurate responses (Matlen et al., 2020).

As with impeded placement, intervening visual units with detailed structure can introduce competing, nonmatching elements that derail the intended alignment. To illustrate, a comparison between the human and frog forelimbs in Fig. 1 could be hindered by the presence of the cat forelimb between them—an intervening visual unit with comparably detailed structure. Comparing the human and bird pair could be more problematic, because three forelimbs are intervening. The deleterious effect of intervening visual units is consistent with the finding that irrelevant visual objects in a display are more likely to distract attention when they match the properties of the target/relevant object (Folk et al., 1992; Gibson & Kelsey, 1998).

¹ Though a sufficiently large barrier could make comparison less efficient by increasing the spatial distance between a matched pair (e.g., Pouw et al., 2019).

Exploring comparisons in real-world educational science images

Prior research sheds light on the kinds of spatial layouts that facilitate comparison. Yet, we don't know how often students are expected to perform comparisons when learning from authentic educational science images, let alone whether the layouts of these images are conducive to structural alignment. This research aims to fill that gap by analyzing images that students are likely to encounter in formal education: the images that appear in science textbooks. Textbooks are one of the most popular mediums of instruction (McDonald, 2016; Valverde et al., 2002; Woodward, 1993), and are widely used around the world (Betrancourt et al., 2012; Boucheix et al., 2020; Slough et al., 2010; Wiley et al., 2017). Though their name suggests otherwise, textbooks contain an abundance of images. In fact, images take up about the same amount of page space as does text (Betrancourt et al., 2012; Mayer, 1993; Slough et al., 2010). Middle and high school science texts contain about 1.5 images per page (Liu & Treagust, 2013). This number has increased over the past several decades, while the number of words per page has gone down (Lee, 2010).

Science textbooks also contain different kinds of images, such as photos and diagrams. These different types tend to play distinct instructional roles (Anagnostopoulou et al., 2012; Mayer, 1993; Wiley et al., 2017). Photos and illustrations are often used to represent objects or categories (Lee, 2010); however, some convey relational structure, such as how the parts of an object or system are connected (Pozzer & Roth, 2003). Diagrams are typically used to convey relational information, such as spatial configurations or causal processes. In this research, we separated *structure* diagrams—which convey static spatial structure—from *process* diagrams, which convey causal processes and changes over time (Heiser & Tversky, 2006). In addition to exploring the overall frequency of spatial supports for comparison, we consider how these supports are distributed across different image types.

Research overview

Our main research questions concern how often students are expected to engage in comparison when learning from educational science images, and whether image layouts tend to support relevant structural alignments. To address these questions, we sampled two chapters each from the three most popular middle school science textbooks in the U.S., identifying 313 educational images in total. We focused on middle school science textbooks, because middle school (typically grades 6–8 in the U.S.) is a time when students are taught a number of complex scientific topics, including evolution and physiology. We coded each image for whether comparison was elicited, using cues in the images and surrounding text.

For each image that involved comparison, we determined whether matched visual units had *direct* or *impeded* placement (Matlen et al., 2020). Structural alignment should be enhanced when matched visual units have direct placement, and should be hampered when placement is impeded. We also identified when the space between a matched pair was occupied by *intervening* units (visual units from *another* comparison), as these may interfere with the structural alignment of a matched pair. By examining the presence of spatial supports for comparison in widely-used educational science images, our findings address whether such images are likely to facilitate students' comprehension of scientific ideas, and could be used to improve how educational science images are designed.

Methods

Materials

Our goal was to examine authentic educational images that students are likely to encounter in school. The 2012 National Survey of Science and Mathematics Education: Status of Middle School Science found that about 80% of US middle school science classes used commercially-published materials (Weis, 2013). The most commonly used science textbooks came from three publishers: McGraw-Hill, Houghton Mifflin Harcourt, and Pearson (Weis, 2013). We selected a current edition (circa. 2017) of the middle school science book from each publisher. We coded physical, bound copies of each textbook. We did not include any digital or other supplemental materials in our coding, as many included dynamic images (videos) and interactive activities that fell outside the scope of the current project.

We focused on two topics in each book, *human anatomy* and *evolution*. We chose these topics because together they cover a range of concrete and abstract concepts, and because prior research suggests that visual comparison may be consequential to learning them (Kok et al., 2013; Kurtz & Gentner, 2013). In total, we coded 176 textbook pages across the 3 books. Of these, 95 pages were about human anatomy and 81 about evolution.

Coding process

We defined an image as any photo, diagram, chart, or graph appearing on the page, excluding background pictures and patterns. The coding of each image was separated into two phases. We describe each coding phase below, and clarify key terms related to analyzing each comparison. Table 1 contains a summary of these terms along with examples based on Fig. 1.

Coding phase 1

The first phase involved classifying the type of image, whether it involved comparison, and (if so) identifying all visual units being compared within the image. A *unit* was defined as a whole object or part within an image, as in the forelimbs in Fig. 1.

The categories for *type of image* included: (1) Photo: a picture taken by a camera; (2) Structure diagram: a stylized drawing that shows the organization (and sometimes behavior) of the parts of an object or system; (3) Process diagram: a stylized drawing that shows a causal process and/or sequence of events; (4) Chart: a table containing information; and (5) Graph: a representation of quantitative data in the form of lines, bars, points, etc. This coding system mainly distinguishes images by their form; however, given our focus on relational structure, we also wanted to distinguish between diagrams that primarily expressed spatial relations (structure diagrams) vs. those that expressed causal relations (process diagrams). This distinction aligns well with Mayer's (1993) science image coding, which distinguished "organizational" and "explanative" illustrations, and the approach of Wiley et al. (2017), which distinguished between "depictive" and "explanatory" graphics.

For each image, we determined whether it *involved comparison* by using cues in the surrounding text (including figure captions and labels) and cues within the image itself. These cues included terms that explicitly related objects or parts within the image (e.g.,

Table 1 Summary and examples of key terms related to the analysis of visual comparisons

Term	Definition	Example from Fig. 1
Units	The objects or parts compared within an image	Each animal forelimb (human, cat, frog, bat, and bird.)
Paired units	Each pair of units being compared within a figure	Each unique pair of forelimbs (e.g., human and cat, human and frog, etc.)
Elements	The components of a unit	The bones (or bone groups) of each forelimb
Axis [of unit]	The intrinsic axis along which the unit changes substantially; typically, its longest axis	The axis of the human forelimb runs from its shoulder to its fingertips
Orientation [of unit]	The direction along which the axis is displayed: Vertical, horizontal, or diagonal with respect to the page	The human forelimb is oriented vertically
Layout [of paired units]	How units are arranged: Vertical (above/below), horizontal (next to each other), diagonal positive (leftmost unit below rightmost), or diagonal negative (leftmost above rightmost)	The human and cat forelimb pair has a horizontal layout
Spatial placement [of paired units]	Image layout with respect to resulting correspondences between paired units: Direct or impeded	The human and cat forelimb pair exemplifies direct placement
Intervening unit	A unit that occupies the space between a matched pair	The cat forelimb is an intervening unit between the human and frog pair

“How do the two types of bone *differ?*”) or generalized across several visual units (e.g., “Fig. 9 illustrates examples of all *three types* of adaptations in the desert jackrabbit”). We also used nonverbal symbols, such as arrows or color highlighting, that conveyed a connection between visual units within an image. Some images involved comparisons at multiple levels of granularity, with comparisons of both larger units and their component parts. Figure 1 has this multi-level structure—the limbs of the different animals are to be compared, as are the individual bones within the limbs. When multiple levels of comparison occurred within the same image, we focused our coding on the largest-scale visual units; in most cases, the spatial layout of smaller units was consistent with that of the larger ones.

For each image that involved comparison, we identified every *pair of units* that was being compared. The coders relied on both text-based cues (as in the examples just above) and non-text cues to identify the pairs. For example, when a diagram contained a sequence of three units, as in a causal chain of events (e.g., Unit 1 → Unit 2 → Unit 3), all neighboring steps (1:2, 2:3) were identified as paired units. When the image contained 3 units to be compared, each combination (1:2, 2:3, 1:3) was considered a matched pair.

Coding phase 2

The second phase of coding focused on dimensions relevant to spatial placement. For each pair of units in an image, we coded the orientation of each unit (vertical, horizontal, or n/a) and the layout of the pair (vertical, horizontal, diagonal positive, or diagonal negative). Together, the orientation and layout of the units determined whether the spatial placement of the pair was direct or impeded (see Fig. 3). We also coded the number of intervening

units between each matched pair. As with impeded placement, intervening units invite incorrect correspondences that could interfere with the intended structural alignment.

We coded the *orientation* of each visual unit by first identifying its *axis*, the line along which the greatest change occurs. In most cases this was operationalized as the longest straight line that intersected at least two points on the unit. For example, in Fig. 1, the axis of the human forelimb runs from shoulder to fingertip. Some shapes, like circles and squares, have no single longest intersecting line, and so the axis was coded as “not applicable (n/a)” in such cases. When the unit had two competing longest lines, we selected the one that fell within its largest area/part. The orientation of each unit’s axis was coded as vertical or horizontal with respect to the layout of the textbook page on which it appeared. The axis parallel to the height of the page was considered vertical, and the axis parallel with the width was considered horizontal. Units with a primary axis within 44° of the page’s vertical axis (in the clockwise or counterclockwise direction) were coded as vertical. Those with a primary axis within 44° of the page’s horizontal axis were coded as horizontal. In rare cases in which the primary axis was a 45° diagonal, it was coded as n/a. We aimed to avoid assigning different orientation codes to units that were close to, but on opposite sides of the 45° line. In the few cases in which this issue arose, we assigned both units to the same category (V or H) using the unit axis farthest from the 45° line as the determinant.

The *layout of unit pairs* refers to the relative spatial location of a pair of units within an image. We coded pair layout as either vertical (one above the other), horizontal (one next to the other), diagonal positive (the leftmost unit below the rightmost), or diagonal negative (the leftmost unit above the rightmost). When there were competing coding alternatives, we considered the amount of vertical, horizontal, and diagonal overlap between the units, and selected the layout category for which overlap was greatest.

Intervening units are those units occupying the space between a matched pair. We operationalized “intervening” as falling on the path of the shortest possible straight line between a pair of units. We counted only those units that were also involved in comparison within the image, excluding background objects and parts of the image that were irrelevant to a comparison. The total number of intervening units was counted for each pair.

Results

Coding reliability and finalization

The research team collaborated to create a comprehensive coding guide, refining the criteria for each category through iterative testing and adjustment with sample pages of the textbooks. After studying the coding guide and performing several rounds of practice, members of the research team coded each textbook page independently. Given the novelty of our coding scheme, we opted to have two coders for every page. There were three coders altogether—one who coded all of the images on each coding variable for both Phase 1 and 2, a second who coded all of Phase 1, and a third who coded all of Phase 2. Table 2 shows the intercoder reliability for the full set of Phase 1 and 2 variables. The kappa values are all at or above .70, suggesting adequate reliability (Landis & Koch, 1977). Furthermore, the research team discussed all coding inconsistencies until they reached 100% agreement. The data we report are based on the finalized, agreed-upon coding. Phase 1 of the coding was completed before the coders began Phase 2.

Table 2 Coding variables, category values, and intercoder reliabilities

Variable	Categories	Reliability
Phase 1: for each image		
Type of image	– Photo – Structure diagram – Process diagram – Chart (tables and lists) – Graph	K = .70 95% CI [.62, .78]
Comparison involved	– Yes – No	K = .75 95% CI [.67, .83]
Paired units	– Open-ended (based on image content)	83% full agreement
Phase 2: for each pair of units		
Unit orientation	– Vertical – Horizontal – n/a	K = .88 95% CI [.85, .91]
Layout of unit pairs	– Vertical – Horizontal – Diagonal positive – Diagonal negative	K = .89 95% CI [.85, .93]
# Intervening units	– Any number ≥ 0 (actual range: 0–7)	K = .84 95% CI [.79, .89]

We report intercoder reliability in terms of Cohen's Kappa (K) except for the open-ended coding of *paired units*, for which we report % agreement

Types of images

We identified 313 total images across 176 textbook pages, a mean of 1.78 images per page (1.75 in the anatomy chapters, and 1.81 for evolution). Figure 4 shows the number of each type of image for each topic and overall.

The five types of images differed widely in frequency, $\chi^2(4, N=313)=206.98, p<.001$. Overall, photos were the most common type, followed closely by structure diagrams, and then process diagrams. Charts and graphs seldom appeared. Altogether, structure and process diagrams accounted for 163 (52%) of the 313 images. The different types of images were distributed similarly in the anatomy and evolution chapters, $\chi^2(4, N=313)=3.56, p=.47$; though, as shown in Fig. 4, structure diagrams were especially prevalent in the anatomy chapters. The three textbooks varied somewhat in terms of the types of images that were included, $\chi^2(8, N=313)=15.34, p=.05$, the largest disparity being that one of the textbooks (HMH Science Fusion) had relatively few process diagrams ($n=4$) compared to the other two texts [$n=16$ (Pearson), and $n=24$ (McGraw-Hill)].

Images involving comparison

Figure 4 also shows the number of each type of image that involved comparison, by topic and overall. Of the 313 total images, 116 (37%) were identified as involving comparison. Images in the evolution chapters were slightly more likely to involve comparison

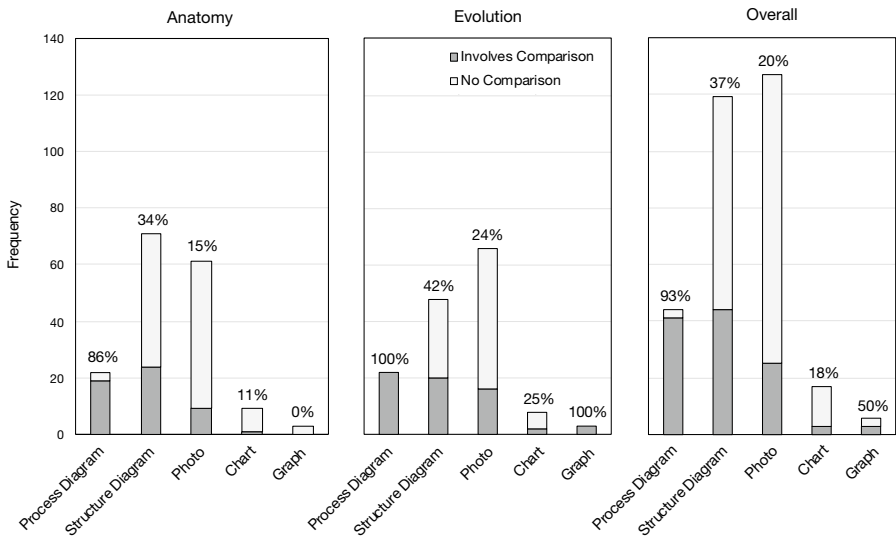


Fig. 4 Frequencies of types of images and number (with percentage) that involved comparison

than images in the anatomy chapters, $\chi^2(1, N=313)=3.99, p=.05$. There was a considerably stronger relationship between frequency of comparison and the image type, $\chi^2(4, N=313)=79.03, p<.001$. Comparison was invoked for 93% of the process diagrams. In contrast, only 20% of photos—the most common image type—involved comparison. The proportion of images involving comparison did not significantly vary between the three textbooks, $\chi^2(2, N=313)=1.58, p=.46$.

The coders noted text-based cues for about 85% of the comparisons, such as, “Do you see any similarities between the bones of the bat and cat limbs and the bones of the human arm?”; “Examine the pictures and observe how the item has changed over time”; “Compare and Contrast: The photos show two bones. Label the healthy bone and the bone with osteoporosis.” Image-based symbols, such as arrows and matching colors, accounted for the remaining 15% of the comparison cues.

Altogether, we find that science textbooks included a substantial proportion of images—nearly 40% overall—for which students must engage in comparison to extract key information. Though comparison was most often required for diagrams (especially process diagrams), comparison was also prompted for several photos. Thus, we find comparison to be involved across the full range of visual representations in the texts.

Spatial supports for comparison

The following analyses explore spatial supports for comparison in the images—in particular, the presence of direct vs. impeded spatial placements, and intervening visual units. Spatial placement and intervening information data are available via the Open Science Framework: https://osf.io/pg3vz/?view_only=ac6c426b961c4b7e96a181c170f38c35.

Given the small number of charts and graphs that involved comparison (only 3 of each type; see Fig. 4), we focused on the other types—process diagrams, structure diagrams,

and photos. We removed from our analysis four images that had extraordinarily high numbers of unit pairs (more than 22.1 pairs—2 *SD* above the mean). We also excluded three images that consisted of large arrays of objects, for example, a display of dozens of shells with the prompt, “Even though the snail shells in Fig. 7 are not all exactly the same, they are all from snails of the same species.” We considered these broad comparisons inappropriate for our analysis. This left us with 103 images in our working data set.

Spatial placement of matched pairs

Because direct placement has been found to enhance structural alignment—whereas impeded placement undermines this process—our chief concern was the number of images that contained each placement type. First, we identified pairs in which the axes of both units had the same orientation (both units either horizontal, vertical, or n/a). Out of 395 total pairs, 350 (about 89%) were coded as having the same orientation. Of these 350 pairs, 168 (48%) were coded as both horizontal, 140 (40%) as both vertical, and 42 (12%) as both n/a. Second, the *layout* of the paired units—that is, their relative spatial position—was coded as either horizontal, vertical, diagonal positive, or diagonal negative, focusing on the 308 pairs for which the units’ axis orientations were both vertical or both horizontal. Prior research and theory on comparison makes a clear prescription for the optimal layout in such cases: horizontal pairs should have a vertical layout (one above the other), and vertical pairs should have a horizontal layout (one next to the other). Pairs with an optimal layout were coded as *direct placements*. Horizontal pairs with horizontal layout and vertical pairs with vertical layout were coded as *impeded placements*. When the layout was diagonal positive or negative, we coded the placement as “other.”

Of the 103 images in our working data set, 72 (70%) contained at least one direct or one impeded placement. Because direct and impeded placements were not normally distributed, we used nonparametric analyses to compare the frequencies of images containing each placement type. Figure 5 shows the frequencies of images with direct spatial placements only, impeded only, and those with a mixture of direct and impeded. Overall, images with direct placements only were the most common, followed by those with impeded placements only, and those with a mixture of direct and impeded, $\chi^2(2, N=72)=35.58, p<.001$. The frequencies of the three spatial placement types varied with the chapter topic, $\chi^2(2, N=72)=6.86, p=.03$, mainly because anatomy images were more likely than evolution images to have direct placement only (see Fig. 5). The frequencies of the spatial placement types did not vary significantly with the type of image (photo, structure diagram, process diagram), $\chi^2(4, N=72)=0.26, p=.99$, or the textbook source, $\chi^2(4, N=72)=0.47, p=.98$.

Based on our results, one might conclude that images were more likely to have an optimal layout for comparison—i.e., paired units in direct placement—than a suboptimal layout. Yet, if we consider the number of images with direct placements only ($n=47$) relative to the total number of images involving comparison ($n=103$), the results seem less favorable: fewer than half had an optimal layout for comparison. Moreover, several images with impeded placements appeared to compromise conceptually-relevant information. For example, in the left side of Fig. 6a, the leg and foot bones of horses across evolution are displayed in a vertical orientation and a vertical layout, resulting in impeded placements. In contrast, in Fig. 6b, from a different textbook, the bones are displayed vertically but in a horizontal layout, resulting in direct placements. Though the two diagrams convey similar information, the layout of Fig. 6b should make it easier to compare the shape of the bones and number of toes across species.

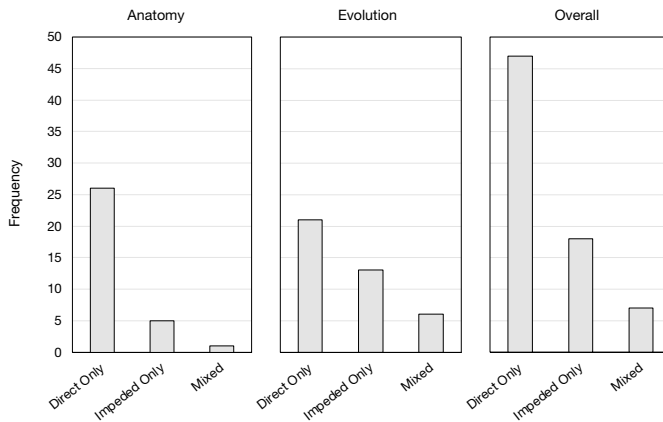


Fig. 5 Frequencies of images with different types of spatial placement

Intervening units between matched pairs

Besides spatial placement, competing element matches between a matched pair—*intervening visual units*—may also affect structural alignment. Our main interest was whether the presence of intervening units varies with spatial placement type. If matched pairs with direct placement are more likely to include intervening visual units, this could indicate a tradeoff between two factors that facilitate structural alignment. Of the 47 images with direct placement only, 14 contained one or more intervening units (30%) compared with 4 out of the 18 images with impeded placement only (22%) and 0 out of the 7 images with mixed placement (0%). Although this pattern is consistent with the possibility of a tradeoff, the presence of intervening units did not vary significantly with spatial placement type, $\chi^2(2, N=72)=2.98, p=.23$.

We also explored whether the presence of intervening units varied with the type of image, the topic, and the textbook source. There was a slight but significant relationship between the frequency of images with intervening units and the figure type, $\chi^2(2, N=72)=7.12, p=0.03$. Structure diagrams were more likely to contain intervening units (12 of 31; 39%) than process diagrams (3 of 31; 10%) and photos (3 out of 10; 30%). The frequency of images with intervening units did not vary significantly with the topic (anatomy and evolution), $\chi^2(1, N=72)=0.00, p=1.00$, or the textbook source of the image, $\chi^2(2, N=72)=0.04, p=0.98$. In sum, intervening units were moderately common in the textbook images, and appeared more often in structure diagrams.

Discussion

Summary of results

In this study we examined instructional images through a cognitive lens, asking how often a student must engage in comparison to understand an image, and how often comparisons are supported by the spatial layout of matched pairs within an image. We found that

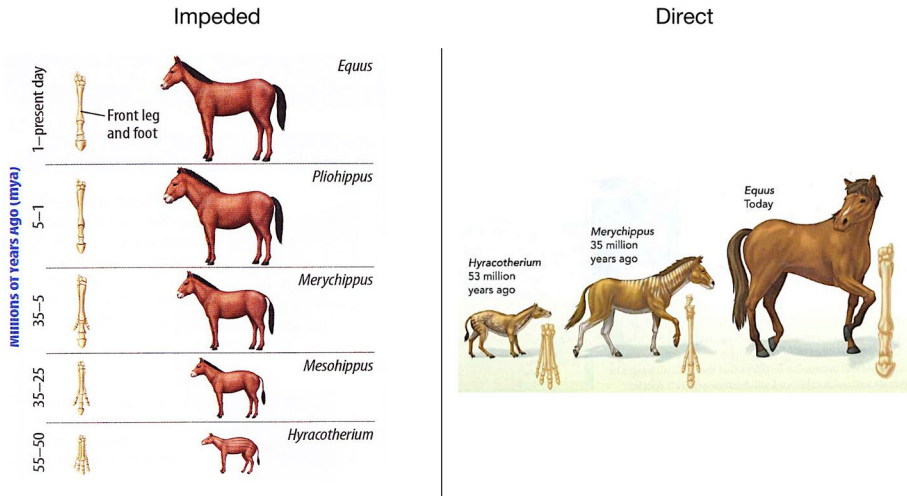


Fig. 6 Diagrams of horse evolution with impeded and direct placement of leg bones. *Note* (a) Illustration from McGraw Hill Integrated iScience Course 2, Student Edition. Copyright © 2016. All rights reserved. Reprinted by permission of McGraw Hill Education; (b) Illustration from Pearson Interactive Science, Life Science Student Edition. Copyright © 2016 by Savvas Learning Company LLC. All rights reserved. Reprinted by permission of Savvas Learning Company LLC

students are frequently required to perform comparisons to comprehend images in middle school science textbooks. Textbooks contained almost two images per page, and over a third of these involved comparisons. Although comparisons were common, images often were not optimally designed to support structural alignment. Intervening visual units were detected in a quarter of the images that had direct and/or impeded placement (18 out of 72). More significantly, fewer than half of the images (47 of 103) were arranged such that matched pairs of visual units were optimally placed—that is, in direct placement. In about a quarter of the images (25 out of 103), one or more matched pairs were in impeded placement.

Implications for learning from science images

Our findings highlight the prevalence of visual comparison in student learning and suggest ample room for improving the spatial design of science images. Given the important role of images in teaching and learning, there may be a considerable payoff for doing so. When a student's goal is to understand a specific topic in a chapter, they spend relatively more time looking at images than at text (Schnotz & Wagner, 2018; Zhao et al., 2020). Successful students are aware of the types of images—such as explanatory process diagrams—that support comprehension, and can budget their attention accordingly (Wiley et al., 2017). Students also treat instructional images differently than text, engaging in more high-level thinking, such as inference-making, when studying diagrams (Cromley et al., 2010). If images appear without written explanations—as is frequently the case (Betancourt et al., 2012)—supports for structural alignment could play a pivotal role in students' efforts to comprehend visual material.

Our approach to image-based comparison is complementary to research on the cognitive integration of information across representations and modalities (e.g., Mayer, 2021; Rau, 2018; Schnotz, 2005). Several studies have found that students benefit from prompts and training to make conceptual connections between visual representations (Ardac & Akaygun, 2004; Bodemer et al., 2005; Linenberger & Bretz, 2012; Seufert, 2019; Stieff & Wilensky, 2003; Van der Meij & de Jong, 2011). When students are explicitly taught to compare and contrast the information within diagrams, their understanding tends to improve (Cromley & Mara, 2018). Spatial supports for comparison could be incorporated into such training. For example, students, especially early in learning, may benefit from easier-to-align materials earlier in instruction, followed by more challenging alignments later—a method known as progressive alignment (Gentner et al., 2011; Kotovsky & Gentner, 1996; Thompson & Opfer, 2010). The initial presence of spatial supports for alignment could compensate for a novices' lack of domain knowledge, which can otherwise hinder their attempts to integrate multiple representations (e.g., Rau, 2018; Seufert, 2019). Similarly, spatial supports for alignment could extend to instructional methods in which structurally-related representations are presented systematically to convey the underlying deep structure; for example, concreteness fading—a progression from concrete/grounded representations to abstract/formal representations (e.g., Day & Goldstone, 2012; Fyfe et al., 2014), and bridging analogy—building from an intuitive to a nonintuitive case through successive comparison (Clement, 1993; Jee et al., 2010). These applications are an interesting possibility for future research.

The benefits of direct spatial placement appear similar to those of spatial contiguity between text and images, which reduces the cognitive burden of sorting through alternative matches (Mayer, 2019). Students spend less time looking at irrelevant images and report lower extraneous cognitive load when spatial contiguity is incorporated into a lesson (Makransky et al., 2019). Similarly, direct spatial *placement* could make it easier for a student to match related objects/parts within an image, and avoid irrelevant, competing correspondences. In future research it will be interesting to explore spatial placement effects in more depth; for example, by tracking participants' eye movements as they use matched pairs of images for simple judgments (as in Matlen et al., 2020) or higher-order sense making. In addition to revealing the time course of the comparison process (e.g., Thibaut & French, 2016), eye tracking could shed light on the effects of direct vs. impeded placements in different tasks and for images with different levels of familiarity and conceptual complexity.

Limitations and future directions

Of course, spatial support for comparison is not the only consideration for producing effective science images. Indeed, other properties, like color and labeling, can facilitate the comparison of visual representations in science (Jee & Anggoro, 2021). Our finding that fewer than half of the textbook images had an optimal spatial layout for comparison does not imply that these images lacked other cognitive supports. Besides supporting comparisons, image designers may be interested in preserving realism, capturing attention, or following established visual conventions in a domain. Designers may also have to fit their images within pre-specified dimensions on a page, and may therefore sacrifice some desirable qualities out of necessity. Nevertheless, there is still likely to be a substantial number of images that could be enhanced by adjusting spatial placement. If student learning is the main priority, then designers of educational images should aim to support relevant

cognition first and foremost (Ainsworth, 2006; Postigo & López-Manjón, 2019). From a broader educational policy perspective, even interventions with small to moderate effects could be cost-effective if they are also cost-efficient (e.g., Kraft, 2020). Incorporating cognitive supports in textbooks places little burden on developers, and can produce positive effects on student learning over the course of a year-long curriculum (Davenport et al., 2020).

The present research involved the development of and implementation of coding methods to identify and classify visual comparisons in real-world instructional images. These methods could be applied and extended in future research in STEM and related disciplines. Yet, we focused on describing spatial supports for comparison, not testing their effects. Whether and how spatial placement affects students' learning of science is an important question that we are actively pursuing. Encouragingly, recent work indicates that spatial placement improves both the efficiency and efficacy of the comparison process (e.g., Matlen et al., 2020; Simms et al., 2019; Zheng et al., 2020, 2022). Also, though we sampled several types of textbook images, we concentrated on middle school material, biology in particular. Texts for younger children tend to contain mostly iconic/realistic pictures, while those for older children have more abstract diagrams (Wiley et al., 2017). Although we did not find substantial differences in the frequency of direct spatial placement across types of images (process diagrams, structure diagrams, and photos), spatial supports for alignment in textbooks may depend on the topic and the audience for which the images are intended.

Another open question is the degree to which a student's preexisting domain knowledge affects the importance of optimal placement. Prior evidence suggests that cognitive supports, such as spatial contiguity, are especially beneficial when students are unfamiliar with the topic of instruction (Ayres & Sweller, 2014; Mayer & Fiorella, 2014). Similarly, students with little domain knowledge are highly likely to benefit from spatial supports for comparison (Gentner et al., 2016; Jee & Anggoro, 2019). Direct spatial placement could help to illuminate the deeper relational structure in an image, drawing the student's attention to the elements involved in a shared system of relations. Novices are also more likely to experience cognitive overload when learning new, complex topics (Chen et al., 2017; Seufert, 2019). Comparison in particular can require novices to maintain a substantial amount of information in working memory while also inhibiting irrelevant material (Begolli et al., 2018). Spatial support for alignment could reduce the burden of these cognitive activities. It is possible that supports for alignment will be less helpful for students with high levels of domain knowledge. Indeed, instructional methods that enhance novice learning can produce negligible or even detrimental results for more knowledgeable learners (Kalyuga, 2007). This is a possibility to explore in future research.

The impact of spatial supports for comparison may also depend on a student's general cognitive skills, such as spatial thinking. Those who are adept at spatial thinking processes, such as mental rotation and spatial perspective taking, tend to perform better on science tasks (Downs & DeSouza, 2006; Newcombe, 2017; Uttal & Cohen, 2012; Wai et al., 2009). However, instructional supports can help students overcome the limits of their spatial skills (Hegarty et al., 2007; Jaeger et al., 2016; Sanchez & Wiley, 2014; Taylor & Hutton, 2013). For example, the use of concrete models was found to help chemistry students—especially those with low spatial skill—translate between different 2D representations of molecules (Stull et al., 2012). Optimal spatial placement might confer a similar benefit by eliminating the need to mentally rotate or rearrange visual objects in order to align them.

Conclusions

Images are central to science instruction. Our research found that students often must compare parts or objects within an image to extract relevant scientific ideas. Yet spatial supports for comparison—direct spatial placements, in particular—were often lacking in science textbook images. Greater attention to image-based supports for comparison could facilitate student learning without adding further burden to instructors.

Acknowledgements This research was supported by the Institute of Education Sciences at the U.S. Department of Education through Grant R305A170165 to WestEd. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. We thank Kelsey Bedard, Ngoc Nguyen, Madison Marceau, and Michaela White for their contributions to the textbook coding. We thank Laura Novick for feedback on an early version of the coding rubric.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16(3), 183–198.
- Ainsworth, S., Prain, V., & Tytler, R. (2011). Drawing to learn in science. *Science*, 333(6046), 1096–1097.
- Anagnostopoulou, K., Hatzinikita, V., & Christidou, V. (2012). PISA and biology school textbooks: The role of visual material. *Procedia-Social and Behavioral Sciences*, 46, 1839–1845.
- Ardac, D., & Akaygun, S. (2004). Effectiveness of multimedia-based instruction that emphasizes molecular representations on students' understanding of chemical change. *Journal of Research in Science Teaching*, 41(4), 317–337.
- Ayres, P., & Sweller, J. (2014). The split-attention principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 206–226). New York: Cambridge University Press.
- Begolli, K. N., Richland, L. E., Jaeggi, S. M., Lyons, E. M., Klostermann, E. C., & Matlen, B. J. (2018). Executive function in learning mathematics by comparison: Incorporating everyday classrooms into the science of learning. *Thinking & Reasoning*, 24(2), 280–313.
- Betrancourt, M., Ainsworth, S., De Vries, E., Boucheix, J., & Lowe, R. (2012). Graphicacy: Do readers of science textbooks need it? In *Staging knowledge and experience: How to take advantage of representational technologies in education and training* (pp. 37–39).
- 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.
- Boucheix, J. M., Lowe, R. K., & Thibaut, J. P. (2020). A developmental perspective on young children's understandings of paired graphics conventions from an analogy task. *Frontiers in Psychology*, 11, 2032.
- Castro-Alonso, J. C., & Uttal, D. H. (2019). Science education and visuospatial processing. In J. Castro-Alonso (Ed.), *Visuospatial processing for education in health and natural sciences*. Springer. https://doi.org/10.1007/978-3-030-20969-8_3
- Chen, O., Kalyuga, S., & Sweller, J. (2017). The expertise reversal effect is a variant of the more general element interactivity effect. *Educational Psychology Review*, 29(2), 393–405.
- Cromley, J. G., & Mara, K. (2018). Comparing and contrasting within diagrams: An effective study strategy. In *10th international conference on the theory and application of diagrams, diagrams 2018* (pp. 492–499). Springer-Verlag.

- Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Cognitive activities in complex science text and diagrams. *Contemporary Educational Psychology*, 35(1), 59–74.
- Clement, J. (1993). Using bridging analogies and anchoring intuitions to deal with students' preconceptions in physics. *Journal of Research in Science Teaching*, 30(10), 1241–1257.
- Davenport, J. L., Kao, Y. S., Matlen, B. J., & Schneider, S. A. (2020). Cognition research in practice: Engineering and evaluating a middle school math curriculum. *The Journal of Experimental Education*, 88(4), 516–535.
- Day, S. B., & Goldstone, R. L. (2012). The import of knowledge export: Connecting findings and theories of transfer of learning. *Educational Psychologist*, 47(3), 153–176.
- Downs, R., & DeSouza, A. (Eds.). (2006). *Learning to think spatially: GIS as a support system in the K-12 curriculum*. National Academies Press.
- Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence*, 41(1), 1–63.
- Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, 18(4), 1030–1044. <https://doi.org/10.1037/0096-1523.18.4.1030>
- Forbus, K., Usher, J., Lovett, A., Lockwood, K., & Wetzel, J. (2011). CogSketch: Sketch understanding for cognitive science research and for education. *Topics in Cognitive Science*, 3(4), 648–666.
- Fyfe, E. R., McNeil, N. M., Son, J. Y., & Goldstone, R. L. (2014). Concreteness fading in mathematics and science instruction: A systematic review. *Educational Psychology Review*, 26(1), 9–25.
- Gentner, D. (1983). Structure mapping: A theoretical framework for analogy. *Cognitive Science*, 7(2), 155–170.
- Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science*, 34(5), 752–775.
- Gentner, D., Anggoro, F. K., & Klibanoff, R. S. (2011). Structure-mapping and relational language support children's learning of relational categories. *Child Development*, 82(4), 1173–1188.
- Gentner, D., Levine, S. C., Ping, R., Isaia, A., Dhillon, S., Bradley, C., & Honke, G. (2016). Rapid learning in a children's museum via analogical comparison. *Cognitive Science*, 40(1), 224–240.
- Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, 52(1), 45.
- Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on thought and language: Interrelations in development* (pp. 225–277). Cambridge University Press.
- Gibson, B. S., & Kelsey, E. M. (1998). Stimulus-driven attentional capture is contingent on attentional set for displaywide visual features. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3), 699–706. <https://doi.org/10.1037/0096-1523.24.3.699>
- Goldwater, M. B., & Gentner, D. (2015). On the acquisition of abstract knowledge: Structural alignment and explication in learning causal system categories. *Cognition*, 137, 137–153.
- Goldwater, M. B., & Schalk, L. (2016). Relational categories as a bridge between cognitive and educational research. *Psychological Bulletin*, 142(7), 729–757.
- Hegarty, M., Keehner, M., Cohen, C., Montello, D. R., & Lippa, Y. (2007). The role of spatial cognition in medicine: Applications for selecting and training professionals. In *Applied spatial cognition* (pp. 285–315). Psychology Press.
- Heiser, J., & Tversky, B. (2006). Arrows in comprehending and producing mechanical diagrams. *Cognitive Science*, 30(3), 581–592.
- Jaeger, A. J., Wiley, J., & Moher, T. (2016). Leveling the playing field: Grounding learning with embedded simulations in geoscience. *Cognitive Research: Principles and Implications*, 1(1), 23.
- Jee, B. D., & Anggoro, F. K. (2019). Relational scaffolding enhances children's understanding of scientific models. *Psychological Science*, 30(9), 1287–1302. <https://doi.org/10.1177/0956797619864601>
- Jee, B. D., & Anggoro, F. K. (2021). Designing exhibits to support relational learning in a science museum. *Frontiers in Psychology*, 12, 636030.
- Jee, B. D., Gentner, D., Uttal, D. H., Sageman, B., Forbus, K., Manduca, C. A., Ormand, C. J., Shipley, T. F., & Tikoff, B. (2014). Drawing on experience: How domain knowledge is reflected in sketches of scientific structures and processes. *Research in Science Education*, 44(6), 859–883.
- Jee, B. D., Uttal, D. H., Gentner, D., Manduca, C., Shipley, T. F., & Sageman, B. (2013). Finding faults: Analogical comparison supports spatial concept learning in geoscience. *Cognitive Processing*, 14(2), 175–187.
- Jee, B. D., Uttal, D. H., Gentner, D., Manduca, C., Shipley, T. F., Tikoff, B., Ormand, C. J., & Sageman, B. (2010). Commentary: Analogical thinking in geoscience education. *Journal of Geoscience Education*, 58(1), 2–13. <https://doi.org/10.5408/1.3544291>

- Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review*, 19(4), 509–539.
- Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, 49(4), 294.
- Kok, E. M., de Bruin, A. B., Robben, S. G., & van Merriënboer, J. J. (2013). Learning radiological appearances of diseases: Does comparison help? *Learning and Instruction*, 23, 90–97.
- Kotovskiy, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development*, 67, 2797–2822.
- Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, 49(4), 241–253.
- Kurtz, K. J., & Gentner, D. (2013). Detecting anomalous features in complex stimuli: The role of structured comparison. *Journal of Experimental Psychology: Applied*, 19(3), 219–232. <https://doi.org/10.1037/a0034395>
- Kurtz, K. J., Miao, C. H., & Gentner, D. (2001). Learning by analogical bootstrapping. *The Journal of the Learning Sciences*, 10(4), 417–446.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- Lee, V. R. (2010). Adaptations and continuities in the use and design of visual representations in US middle school science textbooks. *International Journal of Science Education*, 32(8), 1099–1126.
- Linenberger, K. J., & Bretz, S. L. (2012). Generating cognitive dissonance in student interviews through multiple representations. *Chemistry Education Research and Practice*, 13(3), 172–178.
- Liu, Y., & Treagust, D. F. (2013). Content analysis of diagrams in secondary school science textbooks. In M. Khine (Ed.), *Critical analysis of science textbooks* (pp. 287–300). Dordrecht: Springer. https://doi.org/10.1007/978-94-007-4168-3_14
- Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Role of subjective and objective measures of cognitive processing during learning in explaining the spatial contiguity effect. *Learning and Instruction*, 61, 23–34.
- Markman, A. B., & Gentner, D. (1993a). Structural alignment during similarity comparisons. *Cognitive Psychology*, 25(4), 431–467. <https://doi.org/10.1006/cogp.1993.1011>
- Markman, A. B., & Gentner, D. (1993b). Splitting the differences: A structural alignment view of similarity. *Journal of Memory and Language*, 32, 517–535.
- Matlen, B. J., Gentner, D., & Franconeri, S. L. (2020). Spatial alignment facilitates visual comparison. *Journal of Experimental Psychology: Human Perception and Performance*, 46(5), 443–457.
- Matlen, B. J., Vosniadou, S., Jee, B. D., & Ptouckina, M. (2011). Enhancing the comprehension of science text through visual analogies. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 33, No. 33).
- Mayer, R. E. (1993). Illustrations that instruct. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 5, pp. 253–284). Lawrence Erlbaum.
- Mayer, R. E. (2019). How multimedia can improve learning and instruction. In J. Dunlosky & K. A. Rawson (Eds.), *The Cambridge handbook of cognition and education* (pp. 460–479). Cambridge University Press. <https://doi.org/10.1017/9781108235631.019>
- Mayer, R. E. (2021). *Multimedia learning* (3rd ed.). Cambridge University Press.
- Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 43–52.
- Mayer, R. E., & Fiorella, L. (2014). 12 principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In *The Cambridge handbook of multimedia learning* ((2nd ed., pp. 279–315). New York, NY: Cambridge University Press.
- McDonald, C. V. (2016). Evaluating junior secondary science textbook usage in Australian schools. *Research in Science Education*, 46(4), 481–509.
- Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, 91(2), 358–368.
- Newcombe, N. (2017). Harnessing spatial thinking to support STEM learning. *OECD education working paper no. 161*. OECD Publishing.
- Nyachwaya, J. M., & Gillaspie, M. (2016). Features of representations in general chemistry textbooks: A peek through the lens of the cognitive load theory. *Chemistry Education Research and Practice*, 17(1), 58–71.
- Postigo, Y., & López-Manjón, A. (2019). Images in biology: Are instructional criteria used in textbook image design? *International Journal of Science Education*, 41(2), 210–229.
- Pouw, W., Rop, G., De Koning, B., & Paas, F. (2019). The cognitive basis for the split-attention effect. *Journal of Experimental Psychology: General*, 148(11), 2058

- Pozzer, L. L., & Roth, W. M. (2003). Prevalence, function, and structure of photographs in high school biology textbooks. *Journal of Research in Science Teaching*, 40(10), 1089–1114.
- Rau, M. A. (2017). Conditions for the effectiveness of multiple visual representations in enhancing STEM learning. *Educational Psychology Review*, 29(4), 717–761.
- Rau, M. A. (2018). Making connections among multiple visual representations: How do sense-making skills and perceptual fluency relate to learning of chemistry knowledge? *Instructional Science*, 46(2), 209–243.
- Rau, M. A., Aleven, V., & Rummel, N. (2015). Successful learning with multiple graphical representations and self-explanation prompts. *Journal of Educational Psychology*, 107(1), 30–46.
- Richland, L. E., Zur, O., & Holyoak, K. J. (2007). Cognitive supports for analogies in the mathematics classroom. *Science*, 316(5828), 1128–1129.
- Rittle-Johnson, B., Star, J. R., Durkin, K., & Loehr, A. (2019). Compare and discuss to promote deeper learning. In E. Manalo (Ed.), *Deeper learning, dialogic learning, and critical thinking: Research-based strategies for the classroom* (pp. 48–64). Routledge.
- Sagi, E., Gentner, D., & Lovett, A. (2012). What difference reveals about similarity. *Cognitive Science*, 36(6), 1019–1050.
- Sanchez, C. A., & Wiley, J. (2014). The role of dynamic spatial ability in geoscience text comprehension. *Learning and Instruction*, 31, 33–45.
- Schnotz, W. (2005). An integrated model of text and picture comprehension. *The Cambridge Handbook of Multimedia Learning*, 49, 69.
- Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, 13(2), 141–156.
- Schnotz, W., & Wagner, I. (2018). Construction and elaboration of mental models through strategic conjoint processing of text and pictures. *Journal of Educational Psychology*, 110(6), 850–863.
- Schroeder, N. L., & Cencki, A. T. (2018). Spatial contiguity and spatial split-attention effects in multimedia learning environments: A meta-analysis. *Educational Psychology Review*, 30(3), 679–701.
- Seufert, T. (2019). Training for coherence formation when learning from text and picture and the interplay with learners' prior knowledge. *Frontiers in Psychology*, 10, 193.
- Seufert, T., & Brünken, R. (2006). Cognitive load and the format of instructional aids for coherence formation. *Applied Cognitive Psychology*, 20, 321–331. <https://doi.org/10.1002/acp.1248>
- Simms, N., Matlen, B., Jee, B., Franconeri, S., & Gentner, D. (2019, April 5). Scaffolding visual comparison through spatial alignment. *Paper presented at the 2019 annual meeting of the American Educational Research Association*. Retrieved February 10, 2022, from the AERA Online Paper Repository.
- Slough, S. W., McTigue, E. M., Kim, S., & Jennings, S. K. (2010). Science textbooks' use of graphical representation: A descriptive analysis of four sixth grade science texts. *Reading Psychology*, 31(3), 301–325.
- Stieff, M., & Wilensky, U. (2003). Connected chemistry—Incorporating interactive simulations into the chemistry classroom. *Journal of Science Education and Technology*, 12(3), 285–302.
- Stull, A. T., Hegarty, M., Dixon, B., & Stieff, M. (2012). Representational translation with concrete models in organic chemistry. *Cognition and Instruction*, 30(4), 404–434.
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, 22(2), 123–138.
- Taylor, H. A., & Hutton, A. (2013). Think3d!: Training spatial thinking fundamental to STEM education. *Cognition and Instruction*, 31(4), 434–455.
- Thibaut, J. P., & French, R. M. (2016). Analogical reasoning, control and executive functions: A developmental investigation with eye-tracking. *Cognitive Development*, 38, 10–26.
- Thompson, C. A., & Opfer, J. E. (2010). How 15 hundred is like 15 cherries: Effect of progressive alignment on representational changes in numerical cognition. *Child Development*, 81(6), 1768–1786.
- Uttal, D. H., & Cohen, C. A. (2012). Spatial thinking and STEM education: When, why and how. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 57, pp. 147–181). Elsevier.
- Valverde, G. A., Bianchi, L. J., Wolfe, R. G., Schmidt, W. H., & Houang, R. T. (2002). *According to the book: Using TIMSS to investigate the translation of policy into practice through the world of textbooks*. Springer Science & Business Media.
- van der Meij, J., & de Jong, T. (2011). The effects of directive self-explanation prompts to support active processing of multiple representations in a simulation-based learning environment. *Journal of Computer Assisted Learning*, 27(5), 411–423.
- Vosniadou, S., & Skopeliti, I. (2017). Is it the Earth that turns or the Sun that goes behind the mountains? Students' misconceptions about the day/night cycle after reading a science text. *International Journal of Science Education*, 39(15), 2027–2051.

- 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.
- Wiley, J., Sarmento, D., Griffin, T. D., & Hinze, S. R. (2017). Biology textbook graphics and their impact on expectations of understanding. *Discourse Processes*, 54(5–6), 463–478.
- Woodward, A. (1993). Do illustrations serve an instructional purpose in US textbooks. In B. K. Britton, A. Woodward, & M. Binkley (Eds.), *Learning from textbooks: Theory and practice* (pp. 115–134). Routledge.
- Weis, A. M. (2013). 2012 National Survey of Science and Mathematics Education: Status of Middle School Science. Horizon Research, Inc
- Yuan, L., Uttal, D., & Gentner, D. (2017). Analogical processes in children's understanding of spatial representations. *Developmental Psychology*, 53(6), 1098–1114.
- Zhao, F., Schnotz, W., Wagner, I., & Gaschler, R. (2020). Texts and pictures serve different functions in conjoint mental model construction and adaptation. *Memory & Cognition*, 48(1), 69–82.
- Zheng, Y., Matlen, B. & Gentner, D. (2020). Spatial alignment facilitates visual comparison in children. In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings of the 42nd annual conference of the cognitive science society* (pp. 1546–1551) Cognitive Science Society.
- Zheng, Y., Matlen, B. & Gentner, D. (2022). Spatial alignment facilitates visual comparison in children. *Cognitive Science*, 46(8), e13182. <https://doi.org/10.1111/cogs.13182>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.