Journal of Educational Psychology

ISSN: 0022-0663

2023, Vol. 115, No. 2, 267–289
https://doi.org/10.1037/edu0000774

Exploring the Effects of Gesture-Based Collaboration on Students' Benefit From a Perceptual Training

Martina A. Rau and Joel P. Beier
Department of Educational Psychology, University of Wisconsin–Madison

Collaboration enhances conceptual learning with multiple representations. However, learning with multiple representations also involves perceptual learning processes. These often-overlooked learning processes are the target of perceptual trainings, which expose students to short nonverbal tasks so that students can induce visual patterns across representations. Given the focus of perceptual trainings on nonverbal learning, we investigate the impact of collaboration via gestures without allowing students to talk. On the one hand, gesture-based collaboration may be effective because a partner's gestures may direct students to meaningful visual features. On the other hand, gesture-based collaboration might be ineffective because gesturing may trigger verbal thought, which has been shown to detract from perceptual processing in prior research on the verbal overshadowing effect. We investigated this question in a quasi-experiment with $N = 438$ chemistry undergraduate students. Students either worked on a perceptual training individually or collaborated using only gestures. Posttest data show an advantage of students working individually. Mediation analysis based on log data revealed a positive mechanism of collaboration enhancing learning gains by reducing students' errors during the training. Gesture analysis showed that students used gestures to nonverbally explain their thinking and that representational gestures reduced error rates whereas other types of gestures did not. This might have detracted students from perceptual processing of the stimuli, creating a "nonverbal overshadowing" effect analogous to the verbal overshadowing effect. Altogether, our findings identify boundary conditions of the benefits of collaboration while also revealing possible pathways for future research to explore perceptual learning in social situations.

Educational Impact and Implications Statement

The study revealed that gesture-based collaboration resulted in lower learning outcomes from a perceptual training compared with individual learning. This finding contributes to research to verbal overshadowing of perceptual processing, pointing to an analogous nonverbal overshadowing effect. This finding extends prior theory about collaboration by showing that the well-established benefits of collaboration do not extend to perceptual learning. It highlights the importance of better understanding perceptual learning in social contexts. For instructional practice, the findings suggest that students should work on perceptual trainings alone.

Keywords: collaboration, gestures, perceptual learning, visual representations

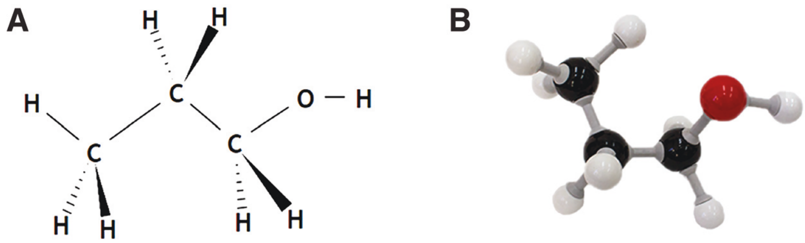
Many professional tasks involve the ability to fluently perceive visual information (Stokes, 2021b). Perceptual fluency describes the ability to quickly and effortlessly extract relevant information from visual representations (Gibson, 1969, 2000; Kellman & Massey, 2013). Although perceptual fluency plays a major role in many professions, little research has focused on the role of perceptual fluency in education.

Yet, perceptual fluency is an important driver of students' learning of content knowledge because most instruction includes visual representations (NRC, 2006). Visual representations can make content accessible to students (Ainsworth, 2008). For example, math students use pie charts and number lines to learn about fractions, engineering students use phasor graphs and spectrograms to learn about sinusoids, and chemistry students use wedge-dash structures and ball-and-stick models (see Figure 1) to learn about molecular geometry. Further, visual representations play a central role in disciplinary discourse (Gilbert, 2005; Kozma & Russell, 2005; Wertsch, 1997). For example, chemists regularly use the visual representations in Figure 1 to communicate about molecular geometry.

However, students often fail to learn content knowledge from visual representations (Ainsworth et al., 1998; NRC, 2006; Rau,

This article was published Online First October 10, 2022.
Martina A. Rau https://orcid.org/0000-0001-7204-3403
This work was funded by the National Academy of Education / Spencer foundation. Special thanks to Tiffany Herder, Edward Misback, Claudia Ramly, and Hanall Sung for their help in conducting the experiment, for the instructors and the Tas for their help, and to Rosanne Luu for his helpful comments on the manuscript.
Correspondence concerning this article should be addressed to Martina A. Rau, Department of Educational Psychology, University of Wisconsin–Madison, Madison, WI 53706, United States. Email: marau@wisc.edu

Figure 1
Visual Representations of Chemical Molecules



Note. (A) Two-dimensional wedge-dash structure. (B) Three-dimensional ball-and-stick model. See the online article for the color version of this figure.

2017). To address this issue, prior research has focused on supporting students' conceptual understanding of the visual representations; that is, students' ability to map visual features to the constructs they depict (Ainsworth, 2006; Rau, 2017). For example, a chemistry student may conceptually understand how the two-dimensional (2D) wedge-dash structure (Figure 1A) uses wedges and dashes to denote three-dimensional (3D) information about whether a bond is oriented toward or away from the viewer, as shown in the ball-and-stick model (Figure 1B).

However, conceptual understanding of visual representations is not sufficient; students also need *perceptual fluency* (Kellman & Massey, 2013; Rau, 2017). Without perceptual fluency, extracting information from representations takes considerable cognitive effort that can detract from students' learning of content knowledge (Anderson & Bodner, 2008). By contrast, perceptual fluency frees cognitive resources that students can invest in learning of content knowledge (Gilbert, 2005; Taber, 2014). For example, a perceptually fluent student who is presented with a 2D wedge-dash structure (Figure 1A) can immediately see the 3D structure it depicts (Figure 1B), without having to conceptually reason about this relationship. This allows the student to invest cognitive resources to learn about chemical properties that arise from a particular molecular geometry.

Students acquire perceptual fluency through *perceptual learning processes* that are nonverbal in nature (Kellman & Massey, 2013). That is, verbal reasoning is not necessary and can even interfere with perceptual learning processes (Kellman et al., 2008; Rau, 2017). This interference results from verbalization creating an additional verbally encoded memory trace that impedes people's ability to retrieve their memory of the original perceptual stimuli (Chin & Schooler, 2008; Schooler et al., 1997).

Building on research on perceptual processes, *perceptual trainings* aim to help students acquire perceptual fluency (Kellman et al., 2010). Perceptual trainings expose students to a series of simple tasks that present a large variety of visual representations and ask students to quickly categorize them based on visual features. Importantly, perceptual trainings emphasize nonverbal learning and discourage students from verbally explaining their thinking.

Given the emphasis on nonverbal learning, research on perceptual trainings has focused on individual learning rather than collaborative learning (e.g., Kellman et al., 2008; Rau & Wu, 2018). However, perceptual learning processes naturally occur in social interactions (e.g., Northedge, 2002), which involve not only verbal communication but also nonverbal communication via gesturing (Singer, 2017). This raises the question whether gesture-based collaboration enhances the effectiveness of perceptual trainings. Although we know of no research that has addressed this question, two lines of prior research yield conflicting predictions.

On the one hand, perceptual fluency can emerge from social interactions where students induce meaning of visual representations based on observing others communicate with visual representations (e.g., Wertsch & Kazak, 2011). Gestures may play a crucial role in the acquisition of perceptual fluency in social contexts: Gesturing can help students access perceptual-motor information (Hostetter & Alibali, 2008). Further, observing a partner's gestures can give students access to their partner's visuospatial knowledge (e.g., Kita, 2000). Thereby, gestures may help students induce meaning from representations.

On the other hand, collaboration has been thought to be most effective for complex tasks (Kirschner et al., 2009a). Given that perceptual trainings focus on simple tasks, they seem unsuitable for collaboration. Further, gesturing may be inseparable from verbal dialogue (e.g., Churchill et al., 2000). Therefore, nonverbal collaboration via gesturing may interfere with perceptual processes in a similar way as verbal explanations interfere with perceptual processes, as described above (see Chin & Schooler, 2008).

The goal of this article is to investigate whether gesture-based collaboration enhances students' benefit from a perceptual training. This goal aligns with Wise and Schwarz's (2017) call to investigate "if, when, and for what ends collaboration [is] beneficial" (p. 433). We conduct our research on a perceptual training for undergraduate chemistry.

Literature Review

In the following, we review theories related to perceptual fluency. Then, we detail the arguments about whether collaboration might enhance or impede perceptual fluency.

Perceptual Fluency

What Is Perceptual Fluency? Traditional modularity theories considered cognition and perception as separate, independent systems (e.g., Fodor, 1985; Pylyshyn, 1999). They proposed a one-directional flow of information: perception delivers information to the cognitive system, but cognition was thought *not* to affect perception. Although modularity theories conflict with more recent empirical findings (Stokes, 2021b), they are worth mentioning because they have a lasting effect on theories of learning with representations. For example, Mayer's (2005, 2009) Cognitive Theory of Multimedia Learning proposes a one-directional pathway where information is loaded from sensory memory into working memory. Similarly, Schnotz's (2005, 2014) Integrated Model of Text and Picture Comprehension focuses the perception-to-cognition pathway.

In contrast, theories of cognitive penetration of perception posit that there is a cognition-to-perception pathway, which accounts for phenomena of rapid perception of conceptually meaningful information. In other words, what we know affects what we see (Goldstone & Barsalou, 1998; Stokes, 2021a, 2021c). At the very least, to perceive facts, one needs relevant concepts (Stokes, 2021c). For example, to visually distinguish an apple from a pear, one needs the concepts APPLE and PEAR. Once one has learned that the perceived apple is an APPLE, the concept APPLE is part of the visual experience: one cannot perceive an apple without identifying it as an APPLE (Siegel, 2006). Further, seeing the apple as an APPLE happens quickly and automatically, without inference making or conceptual reasoning (Stokes, 2021c). Similarly, when presented with visual representations of their area of expertise, experts see meaning at a glance (Chi et al., 1981; Dreyfus & Dreyfus, 1986; Richman et al., 1996). For example, chemists immediately see that the visual representations in Figure 1 show the same molecule.

We refer to this high level of efficiency and automaticity with which experts see meaning in visuals as *perceptual fluency* (Kellman & Massey, 2013; Massey et al., 2011; Rau, 2017). Perceptual fluency is characterized by *involuntary, fast* use of conceptual knowledge, which speaks to cognitive effects on a phenomenon that is perceptual in nature (termed "cognitive penetration of perception"[1]). We define conceptual knowledge as the understanding of principles and relationships that govern a given domain, which may include knowledge about what type of fruit apples are and what varieties exist, or knowledge about how different visual representations depict information about molecular geometry and what inferences can be drawn based on molecular geometry about chemical behaviors. Conceptual knowledge about objects depicted in images affects performance on visual identification tasks (Meyer et al., 2007; Moores et al., 2003). This effect is modulated by saccadic eye movements at around 200-300ms poststimulus (Meyer et al., 2007; Moores et al., 2003) and EEG activity at around 200 ms poststimulus (Telling et al., 2010), which is too fast for voluntary attention direction and thus suggests that the effect is involuntary. Similarly, experts need fewer eye movements and fixate more quickly on relevant visual features than novices (Vogt & Magnussen, 2007), with short time lags indicative of involuntary effects (Kundel et al., 2007). This speed of processing is at least partly explained by *holistic processing* of visuals (Curby & Gauthier, 2010; Richler et al., 2011); that is, experts treat the entire perceived object as a perceptual chunk, as opposed to treating each visual feature as a separate chunk. Holistic processing is *automatic* because experts cannot turn it off: they perform better than novices on visuals that show whole objects, but no better than novices if the objects are partially occluded, when parts are presented sequentially, or spatially distributed (Busey & Vanderkolk, 2005; Curby et al., 2009; Diamond & Carey, 1986; Rossion et al., 2007). These effects correlate with EEG activity 150–200 ms poststimulus, again suggesting an involuntary effect (Rossion et al., 2002; Scott et al., 2006; Tanaka & Curran, 2001). Thus, perceptual fluency is characterized by involuntary, fast, holistic, automatic processing of visuals that is influenced by conceptual knowledge (Kellman & Garrigan, 2009; Rau, 2017; Stokes, 2021a).

While cognition affects perceptual fluency, perceptual fluency also benefits cognitive processing of conceptual information. Efficiency and automaticity in visual processing makes cognitive resources available for conceptual reasoning about complex problems (Goldstone & Barsalou, 1998; Richman et al., 1996) and allows experts to think creatively and react adaptively to novel situations (Dreyfus & Dreyfus, 1986; Gibson, 1969, 2000; Richman et al., 1996). These impacts of perceptual fluency on cognitive processing have been demonstrated in many domains, such as math (Goldstone et al., 2008), chess (Chase & Simon, 1973), and reading (Baron, 1978).

In the context of education, perceptual fluency has been shown to enhance learning of content knowledge in STEM (Gilbert, 2005; Taber, 2014), likely by freeing cognitive resources for higher-order conceptual thinking (Rau, 2017). On the flipside, a lack of perceptual fluency can impede students' learning of content knowledge. For example, Anderson and Bodner (2008) present a case study of a chemistry student who demonstrated conceptual knowledge about how visual representations show domain-relevant concepts but who was not perceptually fluent with the visuals. He was at a considerable disadvantage because he had to invest considerable cognitive effort in making sense of the visuals, which slowed his thinking and impeded his ability to learn new content.

Through Which Processes Is Perceptual Fluency Acquired?

Even though cognition and perception are intertwined, their distinction is important because they involve qualitatively different learning processes (Koedinger et al., 2012). Specifically, the learning processes that lead to conceptual knowledge (i.e., a cognitive mechanism) are often willful and conscious because they involve explicit, verbal reasoning and explanations (Rau, 2017). In contrast, the learning processes that lead to perceptual fluency (i.e., a perceptual mechanism) tend to be inductive and nonverbal (Gibson, 1969, 2000; Kellman & Massey, 2013). Induction of perceptual patterns occurs as the emerging expert encounters a large variety of visual representations (Kellman & Garrigan, 2009). Based on this experience, she starts to recognize recurring perceptual patterns that carry conceptually meaningful information (Kellman & Garrigan, 2009). The perceptual encapsulation of meaningful visual patterns occurs unintentionally and often unconsciously (Rau, 2017).

Perceptual learning processes are characterized as nonverbal because research suggests that they do not require verbal explanations (Koedinger et al., 2012; Rau, 2017) and that verbal explanations interfere with the induction of perceptual patterns (Chin & Schooler, 2008; Schooler et al., 1997). Research on this "verbal overshadowing" effect has been conducted on many types of tasks, including facial recognition, wine tasting, and music classification (for an overview, see Chin & Schooler, 2008). Methodologically, typical studies on the verbal overshadowing effect prompt

[1] This idea and the following findings conflict with modularity theories, which interpreted perceptual fluency as a cognitive effect where experts' conceptual knowledge allows them to *voluntarily* direct their visual attention to conceptually relevant visual features (Fodor, 1985; Pylyshyn, 1999).

participants to verbally describe aspects of a perceptual stimulus at the time of its encoding (e.g., when a picture is presented, describe what a person in the picture looks like in as much detail as possible). Compared with participants who are not prompted to describe the stimulus, prompted participants tend to perform more poorly on later tasks related to the perceptual stimulus (see Chin & Schooler, 2008). The verbal overshadowing effect has been explained by people's tendency to rely on verbal information when it is available at the expense of the original perceptual information (Chin & Schooler, 2008; Schooler & Engstler-Schooler, 1990). Specifically, verbalization creates a verbally encoded memory trace of the stimulus that exists alongside the original perceptual memory. When people attempt to recall the stimulus, they default to retrieving the verbal memory and fail to retrieve the original perceptual memory. Because perceptual information is difficult to describe, the verbally encoded memory is often inaccurate.

How Is Perceptual Fluency Supported? Owing to the inductive and nonverbal nature of perceptual fluency, instructional interventions that aim to help students acquire perceptual fluency emphasize the importance of nonverbal learning (Kellman & Massey, 2013; Rau, 2017). Perceptual trainings are a type of computer-based instructional intervention that build on several principles based on extant research on perceptual learning processes (Kellman & Massey, 2013; Rau, 2017). First, perceptual trainings expose students to a large variety of visual representations that are commonly used in the given domain (Kellman & Garrigan, 2009; Kellman et al., 2008). Because they focus on quick perceptual processing, perceptual training tasks are simple and usually ask students to categorize or classify visual representations. For example, chemistry students might be exposed to several tasks that ask them to quickly identify one of several ball-and-stick models (Figure 1B) that shows the same molecule as the one shown by a wedge-dash structure (Figure 1A). These tasks are typically very short (e.g., students may complete about 20 short tasks in about 5 minutes). Second, the visual representations present contrasting cases (Chase et al., 2010) so that they repeat visual features that carry meaningful information and vary conceptually irrelevant visual features. For example, if the number of black spheres is a meaningful visual feature in a ball-and-stick model (Figure 1B), then students should encounter ball-and-stick models with an incorrect number of black spheres so that they learn to extract information from the visual representation based on this feature. Third, students are discouraged from verbalizing or explaining their thinking. For example, students may be prompted to solve the tasks quickly and to rely on their perceptual intuitions. Finally, students receive immediate feedback on their use of visual information, but this feedback does not provide verbal explanations so as not to disrupt perceptual processing. For example, students may see color highlights that signal whether their response was correct.

In line with the notion that perception and cognition are intertwined, perceptual trainings are typically offered after students have acquired conceptual knowledge about how the visual representations show domain-relevant information (Kellman & Garrigan, 2009; Kellman et al., 2008). Indeed, prior experimental research shows that perceptual trainings are most effective when students have acquired conceptual knowledge about the representations (Rau, 2018; Rau et al., 2017a) and may even be ineffective for students who lack conceptual knowledge (Rau et al., 2017b; Rau & Wu, 2018). Research suggests that conceptual knowledge enables students to attend to meaningful visual features, which enhances their benefit from perceptual trainings (Rau, 2018).

Further, the goal of perceptual trainings is usually not the acquisition of perceptual fluency alone but rather to enable students to efficiently use representations to solve domain-relevant tasks (Kellman & Garrigan, 2009; Kellman et al., 2008). Therefore, studies on perceptual trainings usually assesses students' performance on domain-relevant tasks that provide disciplinary information through visual representations, for example in math (Kellman & Garrigan, 2009; Rau et al., 2017b) or chemistry (Rau & Wu, 2018; Wise et al., 2000). Indeed, these studies show that perceptual trainings enhance students' performance on such tasks.

Collaborative Learning With Visual Representations

Although extant research shows that collaboration can enhance students' conceptual knowledge about visual representations (e.g., Rau et al., 2017; Schwartz, 1995; Strickland et al., 2010), we know of no studies that have tested how collaboration affects the acquisition of perceptual fluency. Yet, two lines of research suggest competing predictions about this question.

Collaboration May Enhance Perceptual Fluency via Nonverbal Gesturing. Sociocultural perspectives on learning view visual representations as a tool that supports social interactions (Hakkarainen et al., 2013; Rau, 2017). Hence, visual representations are viewed as socially constructed cultural tools that enable communication and problem-solving practices within the disciplinary community (Wertsch, 1997). Specifically, members of the disciplinary community express themselves through visual representations and use them as thinking tools (Airey & Linder, 2009; Braden & Hortin, 1982; Schönborn & Anderson, 2006). Different disciplines have different practices in using visual representations, which involve the ability to "quickly register perceptual features that are relevant to their particular practice, features invisible at a glance to nonexperts" (Stevens & Hall, 1998, p. 109).

From this perspective, the purpose of acquiring perceptual fluency is fundamentally social as well (Rau, 2017). Participating in disciplinary practices necessarily involves becoming fluent in using visual representations (Northedge, 2002; Postman & Wiengartner, 1971). Students' visual communication needs to be so fluent that their use of visual representations is "unproblematic, almost second nature" (Airey & Linder, 2009). Otherwise the task of interpreting representations during disciplinary discourse may overwhelm students (Taber, 2002). Indeed, the sociocultural literature describes the ability to fluently use visual representations as an important instructional goal (Goodwin, 1994; Kozma & Russell, 2005; Wertsch & Kazak, 2011).

Moreover, the process through which students acquire perceptual fluency is inherently social (Rau, 2017). Students acquire disciplinary representational practices by participating in community practices (Greeno & Hall, 1997; Northedge, 2002, 2003; Roth, 2014; Roth & McGinn, 1997). Participation in community practices allows students to inductively learn disciplinary ways of knowing and communicating as well as ways of seeing and perceiving (Braden & Hortin, 1982; Goodwin, 1994). Such inductive learning processes unfold as students come to appropriate communication and perception practices that happen at the "intermental plane" (i.e., between persons) until they appear at the "intramental plane" (i.e., within a person; Vygotsky, 1962; Wertsch & Kazak, 2011). Hence, it is possible and even desirable for students to participate in representational social interactions before they fully understand the meaning of the representations (diSessa, 1993; Wertsch & Kazak, 2011). During participatory interactions, a more knowledgeable member of the community "disciplines" the student's perception by nudging them toward normative use of representations (Stevens & Hall, 1998). Notably, in describing the emergence of perceptual fluency from social interactions, this literature documents how students incrementally induce meaning as they observe and imitate others while they use representations in ways that do not necessarily involve verbal communication (Airey & Linder, 2009; Wertsch & Kazak, 2011). Thus, the observation that perceptual learning processes can occur in social contexts does not necessarily conflict with the notion that they are nonverbal.

Although sociocultural research typically does not seek to isolate which specific components of discourse affect specific types of learning outcomes, it documents the critical role of gestures in the acquisition of perceptual fluency in social contexts (Stevens & Hall, 1998). Indeed, gesturing has been shown to facilitate inductive processes while students participate in social practices (e.g., Carraher & Schliemann, 2002; Roschelle, 1992). The role of gestures in social interactions has been widely studied in the cognitive and sociocultural literatures (e.g., Cook & Fenn, 2017; Nathan et al., 2007; Singer, 2017). Although much of this research has focused on how gesturing enhances conceptual learning processes (e.g., Alibali & Nathan, 2007; Alibali et al., 2014; Singer, 2017), research suggests two mechanisms through which gestures may enhance perceptual learning processes.

First, gesturing may help the student who is performing the gesture to engage in perceptual learning processes. Evidence for this claim comes from studies showing that prohibiting students from gesturing results in disfluencies in speech (Graham & Heywood, 1975), particularly in conveying spatial information (Rauscher et al., 1996) and in including imagery references in speech (Rimé et al., 1984). Indeed, gestures occur frequently when speakers access perceptual-motor information (Hostetter & Alibali, 2008). Further, gestures seem to play a particularly important role in contexts where the goal is for students to engage in perceptual learning processes (Cope et al., 2015). Also, students who spontaneously gesture (Chu & Kita, 2008) or who are instructed to gesture (Broaders et al., 2007; Cook et al., 2008) demonstrate higher learning outcomes from visuospatial instructional activities. These effects may be partially moderated by spatial skills: Students with high spatial skills tend to gesture more often than students with low spatial skills (Hostetter & Alibali, 2007).

Second, watching a partner's gestures may help students engage in perceptual learning processes. Gesture research generally distinguishes representational gestures (i.e., gestures that resemble the referent, such as hand shapes that resemble object shapes) and deictic gestures (i.e., movements that indicate a concrete referent in the environment, such as pointing at an object; McNeill, 1992; Stephens, 1983). Representational gestures convey a speaker's internal representations (Cienki, 2005; Hostetter & Boncoddo, 2017; Nathan, 2008). They can encode implicit knowledge and spatial information that the speaker may not be able to express verbally (Broaders et al., 2007; Goldin-Meadow & Alibali, 2013; Kita, 2000; Schwartz & Black, 1996; Zhen et al., 2019). Getting insights into a partner's internal representation and implicit visuospatial knowledge may thus help students induce meaning from external representations. Further, deictic gestures (i.e., pointing at something) play an important role in directing a partner's attention to features of the environment (Fussell et al., 2000, 2004). Deictic gestures have been documented in sociocultural research on the acquisition of perceptual fluency in social contexts where a more knowledgeable person points at parts of a visual representation to direct the student's attention to relevant features (Stevens & Hall, 1998). Collaborative interactions in which students help each other attend to specific visual features have been shown to enhance students' ability to extract accurate information from visual representations (Wu et al., 2019). Indeed, deictic gestures play an important role in helping students induce meaning from visual representations (Rau & Patel, 2018; Stevens & Hall, 1998).

Collaboration May Impede Perceptual Fluency via Explanation-Based Processes. Sociocognitive research, in contrast, suggests that collaboration might interfere with the acquisition of perceptual fluency. One argument draws on the Knowledge, Learning, and Instruction framework (KLI; Koedinger et al., 2012), which distinguishes knowledge types based on the complexity of the involved knowledge structures. KLI characterizes perceptual fluency as a "simple" type of knowledge because it relies on pattern recognition that involves mapping a single visual pattern to a single construct (e.g., an image of a chemical molecule can be mapped to only one molecule). In contrast, "complex" types of knowledge involve learning of rules and schemas that have many application contexts and many possible solutions (e.g., how chemical bonding works may be applied to numerous chemical structures and can be explained in multiple ways). According to KLI, instruction is effective if it engages students in interactions that match the complexity of the to-be-learned knowledge (Koedinger et al., 2012).

The learning processes fostered by collaboration are arguably complex. When students collaborate, they may build on each other's contributions, explaining their ideas, defending and challenging each other's positions, thereby prompting each other to explain their thinking (Chi, 2009). Students engaged in such collaborative exchanges are more actively engaged in explanation-based learning processes than students who work alone (Chi, 2009). The benefits of these collaborative exchanges outweigh any so-called "process losses" that result from group members having to coordinate and communicate with each other (Kirschner et al., 2009a). Therefore, the overall effect of collaboration is positive for learning from complex tasks (Kirschner et al., 2010; Nokes-Malach et al., 2012).

By contrast, when students learn simple types of knowledge, there is nothing to explain (Koedinger et al., 2012). Consequently, collaboration may not offer a benefit (Koedinger et al., 2012; Nokes-Malach et al., 2012; Wylie et al., 2009). If a task can be completed by an individual alone, the costs associated with process loss outweigh any benefit of collaboration, resulting in an overall negative effect: Collaborating increases cognitive effort and time that students must invest in coordinating and communicating with each other compared with students working individually (Ciborra & Olson, 1988; Clark & Brennan, 1991; Kirschner et al., 2009b; Yamane, 1996). Indeed, research on simple verbal recall tasks shows that collaborative learning is less effective than individual learning because the cognitive effort invested in coordination and communication does not pay off (Clark & Brennan, 1991; Kirschner et al., 2009b). By the same argument, collaboration may reduce the efficiency of perceptual trainings by increasing time and effort without offering a benefit. Yet, this research has not empirically examined perceptual trainings.

A second argument emerges from the finding that verbalization interferes with perceptual processing (Chin & Schooler, 2008; Schooler et al., 1997). Benefits of collaboration have largely been attributed to verbally mediated coconstruction of knowledge as students explain ideas to each other (Chi, 2009; Dillenbourg et al., 1996). As detailed above, the verbal overshadowing effect suggests that verbalization interferes with perceptual processing because the availability of verbal information detracts from perceptual information (Chin & Schooler, 2008; Schooler & Engstler-Schooler, 1990). Thus, ordinary collaboration that involves verbal communication seems inappropriate for perceptual trainings.

What about collaboration that involves only nonverbal communication via gesturing? To our knowledge, prior sociocognitive research has focused only on verbally mediated collaboration and has not tested whether nonverbal collaboration via gestures could enhance learning of simple types of knowledge. However, prior research on gesturing provides two related arguments suggesting that gesturing may interfere with perceptual fluency.

First, gesturing might trigger internal verbalization. Indeed, several gesture theories propose that gesture and speech are part of the same system that cannot be separated (Goldin-Meadow, 2003; Hostetter & Alibali, 2008; McNeill, 1992) and that gesturing facilitates speech production (Krauss et al., 2000). Similarly, research on artifact-centered discourse and anchored discourse considers gesturing an integral part of dialogue (Churchill et al., 2000; Guzdial, 1997; Suthers & Xu, 2002). Dialogue not only involves speech but also nonverbal communication devices such as gestures (Cassell et al., 2000). If gestures are linked to speech production and inseparable from verbal dialogue, then gesturing may trigger verbal thought even when speech is prohibited. For example, gesturing about the orientation of a molecule representation may trigger internal verbalization such as “the molecule is rotated.” As argued above, such verbalization may create a verbally encoded memory trace that interferes with retrieval of the original perceptual information.

Second, a slightly more elaborate argument builds on the first by suggesting gesturing may constitute a complex explanation process. For example, Fussell and colleagues (2004) consider gestures a shortcut to verbal explaining because they provide a quick and efficient way to communicate about a collaborative task. Specifically, representational gestures serve to encapsulate complex ideas (Bekker et al., 1995; Emmorey & Casey, 2001; Tang & Leifer, 1988). Deictic gestures serve to indicate task-relevant referents (Bauer et al., 1999; Fussell et al., 2000). When not allowed to gesture, these ideas and referents are communicated via verbal explanations (Bauer et al., 1999; Fussell et al., 2000). Thus, gesturing may be a nonverbal form of explanation. For example, suppose a student is presented with a wedge-dash structure and a ball-and-stick model showing a planar molecule (i.e., all atoms reside on the same two-dimensional plane). To make sense of these representations, the student may use a representational gesture that positions the thumb and index finger of both hands at 120-degree angles to each other while silently explaining to herself how this corresponds to a planar molecule. In this way, the gesture supports a complex explanation process, which—as argued above—might interfere with perceptual learning processes. This may hold even if the student does not verbalize the explanation.

Research Questions and Hypotheses

Our review of prior research reveals a gap in our understanding of how nonverbal collaboration via gesturing may affect students’ benefit from a perceptual training. Two different lines of research suggest conflicting arguments about whether gesture-based collaboration may enhance or impede perceptual fluency. We note that both arguments regarding this question are speculative because they have no direct empirical basis yet. This article addresses this question in the context of a perceptual training for chemistry students. While we examine whether students demonstrate perceptual fluency in the sense that they become better at solving the perceptual tasks in the training, our main outcome of interest is performance on chemistry tasks that involve visual representations. This focus is in line with prior research on perceptual trainings, which have the goal to enhance students’ efficient use of visual representations when solving domain-relevant tasks (Kellman & Massey, 2013). Hence, we investigate:

> *Research Question 1:* How does nonverbal collaboration via gesturing affect the effectiveness of a perceptual training with respect to (a) students’ performance on perceptual tasks during the perceptual training and (b) their performance on chemistry tasks after the perceptual training?

Hypothesis 1.1 emerges from sociocultural research and predicts that gesture-based collaboration enhances the effectiveness of the perceptual training, whereas Hypothesis 1.2 emerges from sociocognitive research and predicts that gesture-based collaboration decreases its effectiveness.

In addition, we seek to understand the mechanisms underlying potential effects. To this end, we investigate how gesture-based collaboration changes students’ interactions with the perceptual training compared with students who work individually on the perceptual training, and whether this accounts for potential effects on the effectiveness of the perceptual training. We focus on interactions that can be gleaned from the log data from the perceptual training (i.e., training time, error rates). Specifically, we ask:

> *Research Question 2:* Are students’ interactions with the perceptual training (a) affected by gesture-based collaboration, and (b) do they mediate potential effects of gesture-based collaboration on students’ benefit from the perceptual training?

Finally, to get a better understanding of how students use gestures in their nonverbal communication when collaborating on the perceptual training, we qualitatively explore:

> *Research Question 3:* How do students use gestures during the perceptual training?

Method

Transparency and Openness

In the following sections, we describe in detail the methods we used to obtain the reported data, how we determined our sample size, and any data exclusion criteria. Data were analyzed using IBM SPSS Statistics Version 27 and Tetrad Version 6.8.1. This study’s design and its analysis were not preregistered.

Participants

Participants were $N = 527$ first-year undergraduate students in a semester-long introductory general chemistry course at a large Midwestern university in Fall 2019. The course was designed for students who intend to major in a STEM subject. Prerequisites for the course included at least one year of high-school chemistry and demonstration of high mathematics achievement, for instance, by qualifying for an advanced mathematics course. Students enrolled in this course had an average GPA of 3.20 (on a scale from 0–4).

All students who were enrolled in the course participated in the experiment activities; however, their data were used for analysis only if they provided informed consent to their data being used for research purposes. This procedure was approved by our institution's ethical internal review board.

Although we were unable to assess demographics of study participants owing to restrictions placed by our internal review board, we report demographics of undergraduate students majoring in chemistry at the university: 43.79% of these students are female; 68.38% are Caucasian, 1.80% are African American, 3.60% are Hispanic, 5.70% are Asian, 18.71% are International, and 1.80% are unknown; 59.96% are state residents, and 43.04% nonresidents.

Our experiment was conducted at the end of a 3-hour lab session in Week 5 of the chemistry course. The course involved two weekly 50-minute lectures, a weekly 3-hour lab session, and two weekly 50-minute discussion sessions. The lecture was taught by three professors, and each student was assigned to the lecture of one of the professors. Lab and discussion sessions were held in smaller groups; namely 32 sections of about 18 students. Lab and discussion sessions were led by 18 teaching assistants (TAs). The professor was present in discussion sessions, but not in lab sessions. Throughout the semester, students worked collaboratively in small groups during both discussion and lab sessions with a partner (or, in rare cases, two partners, if an uneven number of students was enrolled in the section) who was assigned for the duration of the semester.

Experimental Design

Using a quasi-experimental design, we assigned 15 sections of the course to the individual condition ($n = 253$ students) and 17 sections to the collaborative condition ($n = 270$ students). Students selected a course section at the beginning of the semester. They tend to select sections so that they fit well into their class schedule. We have no reason to believe that systematic differences exist between sections. In addition, we took the following steps to ensure equivalency of the conditions. To counterbalance potential effects of time of day, we assigned the same number of early and late morning, early and late afternoon, and evening sections to each condition. Further, to counterbalance potential effects of TA, each TA led one session for the individual and one for the collaborative condition, with one exception owing to scheduling constraints, which led to the slightly unequal assignment of 15 versus 17 sections to the conditions. In addition, to counterbalance potential effects of the lecturing professors, we ensured that students who attended each professor's lecture were equally distributed across the two conditions.

Students in the *individual* condition worked individually on the perceptual training (detailed below). Students accessed the perceptual training on their personal laptop. Students in the *collaborative* condition worked in small groups (i.e., dyads, or in very few cases in groups of three, if there was an uneven number of students in the given section) on the perceptual training (detailed below). They used the laptop of one of the students in their small group. Students worked in the small groups they were assigned to for the duration of the semester. If the partner was absent from the class, the TA assigned them to an alternative small group, never exceeding three students per group.

Materials

Prior Class Activities

In keeping with the common use of perceptual trainings after conceptual instruction, students received the perceptual training after working on conceptually focused class activities that regularly occur in the given week of the chemistry lab session. These class activities focused on conceptual understanding of how to translate between 3D ball-and-stick models and 2D wedge-dash drawings (see Figure 1). The regular class activities took up most of the class period. Our experimental manipulation took place after these activities.

Perceptual Training

Our experimental manipulation regarded the perceptual training, to which the last 15 minutes of the chemistry lab session was dedicated (the perceptual training took about 5 minutes).

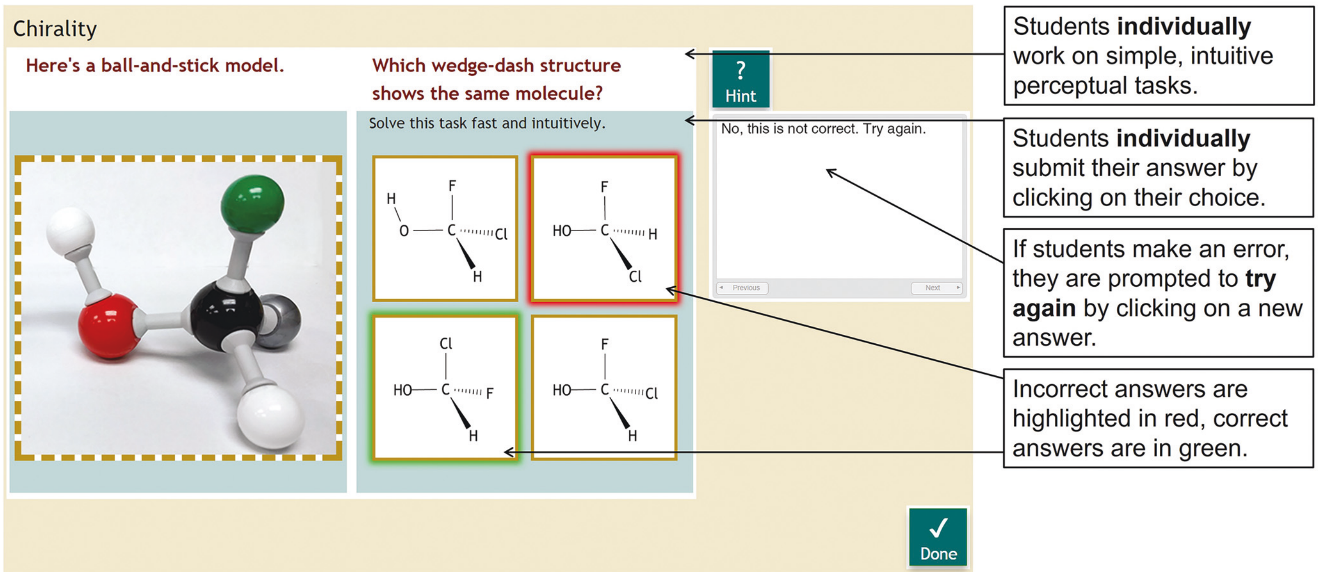
In consultation with the chemistry course instructors, we specifically created the perceptual training to align with the content covered in the regular class activities relating to translating between ball-and-stick models and wedge-dash structures. Hence, the training aimed at helping students become perceptually fluent at translating between these visual representations. The visual representations depicted molecules that are common examples of stereoisomerism and included molecules that students had and had not encountered during the class activities.

The perceptual training was designed based on the principles described above. To encourage students to rely on nonverbal, inductive processes, students were asked to solve the perceptual tasks "fast and intuitively without overthinking it." Students received 20 short tasks that provided one visual representation (e.g., a ball-and-stick model) and asked students to select one of four visual representations (e.g., a wedge-dash drawing) that showed the same molecule. The choice options varied a range of visual features that are relevant to isomerism (e.g., different bonding arrangements) and irrelevant features (e.g., rotations around single bonds) so that students would learn to attend to the relevant features.

Individual Version of the Perceptual Training. Before students in the individual condition started working on the individual version of the perceptual training, they watched a short video. The video explained the purpose of the perceptual training and the importance of following one's own perceptual intuitions and of solving the tasks quickly without overthinking them. The video ended with an example of an individual student solving perceptual tasks. Further, each perceptual task included a text prompt to solve the task fast and intuitively (see Figure 2).

Collaborative Version of the Perceptual Training. Before students in the collaborative condition started working on the collaborative version of the perceptual training (see Figure 3),

Figure 2
Individual Version of the Perceptual Training

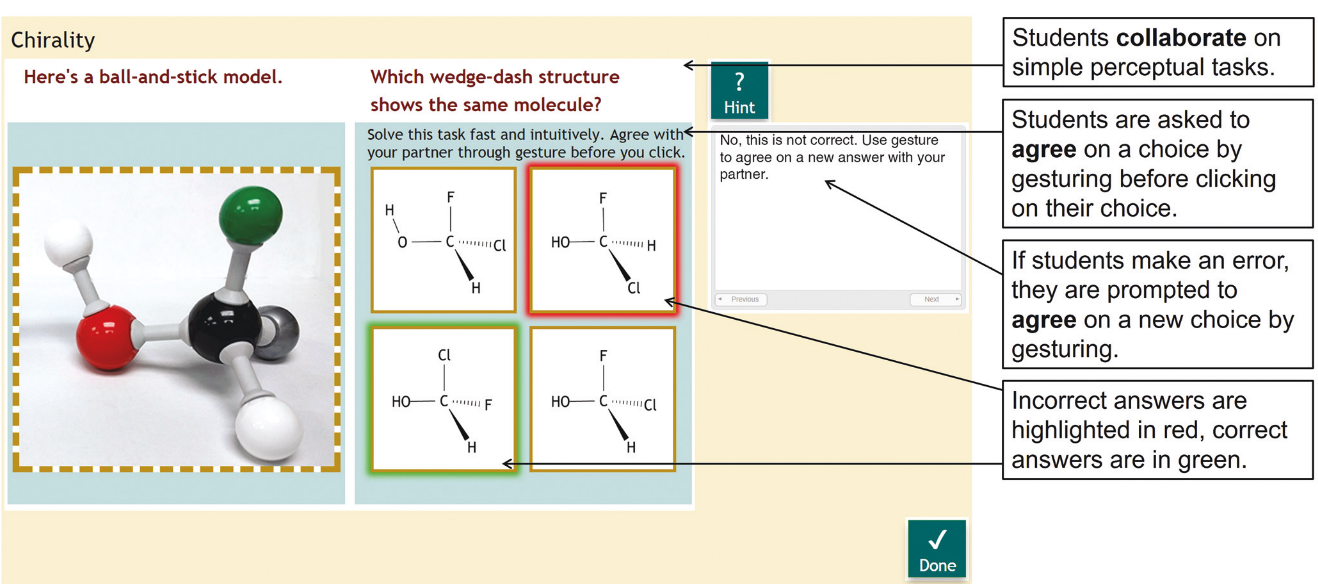


Note. See the online article for the color version of this figure.

they watched a brief video that was equal in length to the video provided along with the individual version. The video explained the purpose of the perceptual training as engaging students in inductive, nonverbal processes involved in perceptual pattern recognition. It also explained that verbalization can impede the acquisition of perceptual fluency as well as why perceptual fluency is important for chemistry. Students were told that they should collaborate using only gestures to communicate and that they should not talk to each other. Students were instructed that one of them would operate the mouse but that they would have to reach consensus on the answer choice before clicking. As illustrated in Figure 4, the video ended with an example of a dyad who used gestures to collaborate, hence serving as a model of nonverbal collaboration.

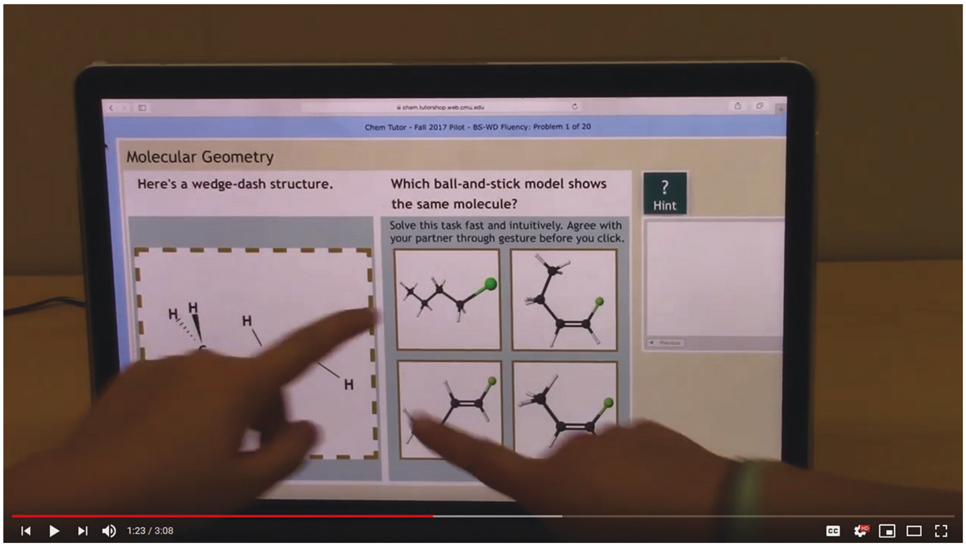
Further, each of the perceptual tasks included a text prompt for students to solve the tasks fast and intuitively and to agree with

Figure 3
Collaborative Version of the Perceptual Training



Note. See the online article for the color version of this figure.

Figure 4
Collaborating Students Were Shown an Example of a Dyad Using Gestures to Collaborate on the Perceptual Training



Note. See the online article for the color version of this figure.

their partner to gesture before clicking on an answer (see Figure 3). Once students had reached an agreement, they could submit the answer. If their answer was wrong, the script prompted students to agree on a new answer using gestures only (see Figure 3).

We conducted two pilot tests of the collaborative version of the perceptual training. In the first pilot study (Rau & Patel, 2018), 20 undergraduate students (10 dyads) were recruited from an introductory chemistry course to take the perceptual training in a research laboratory. We evaluated whether students used gestures indicating features of visual representations that corresponded to the learning goals of the perceptual training. Results showed that students' gestures exclusively corresponded to features that carried meaningful information (e.g., the number of atoms or bonds), thus matching the learning goals of the perceptual training. Further, we interviewed students about their experiences with nonverbal, gesture-based collaboration. Students reported that it felt strange at first to not be allowed to talk but that they quickly became comfortable with the procedure. They also reported that they found their partner's gestures useful because they helped them attend to visual features they would have otherwise missed, because they increased their confidence in their own perceptions, and because they aligned with the use of nonverbal communication in their chemistry courses. Finally, we compared how students' error rates on the perceptual tasks decreased over time and compared the error rates to data from 28 undergraduate chemistry students who had worked individually on the same perceptual training. Results showed that students' error rates decreased across the 20 perceptual tasks, indicating that students became perceptually fluent. Students who worked collaboratively on the perceptual training also showed significantly lower error rates than students who worked individually. This pilot study showed that nonverbal, gesture-based collaboration in the context of a perceptual training is feasible, that students engage in the desired form of collaboration, and that doing so results in perceptual fluency. A limitation of this pilot study is that it was conducted in the controlled context of a research laboratory.

Therefore, the second pilot study tested whether nonverbal, gesture-based collaboration on the perceptual training works in the context of a chemistry course, especially when students are used to collaborating verbally. Participants were 22 undergraduate students in a small elective chemistry course. The pilot-study took place in Week 14 of the semester. Students worked on the perceptual training at the end of the course session and had collaborated verbally on a related chemistry task before. We observed that all students complied with the instructions not to talk and relied on gestures to communicate with their partner. They demonstrated the desired form of collaboration by agreeing with their partner before making a choice while equally contributing gestures to the communicative exchange. Further, we observed that error rates decreased over the course of the 20 perceptual tasks, indicating that they became perceptually fluent. This pilot-test shows the feasibility of nonverbal, gesture-based collaboration in the context of a chemistry class.

Assessments

Performance on Perceptual Tasks. To assess whether collaboration affects students' performance on perceptual tasks (Research Question 1a), we examined their learning curves (Koedinger & Mathan, 2004; Stamper et al., 2013). The learning curves depict how students' performance on the perceptual tasks changes over time. Specifically, we considered how the proportion of errors made on the first attempt at solving each perceptual task declines across the 20 tasks of the perceptual training. We consider a decline in error rates an indicator of perceptual fluency gains. This assessment was obtained at the level of the small group for the collaborative condition and at the level of the individual student for the individual condition.

Performance on Chemistry Tasks. To assess whether collaboration affects students' performance on chemistry tasks (Research Question 1b), we created a chemistry test that involved the visual representations targeted in the chemistry lab session. The test included visual representations that were not part of the perceptual training. Because the goal of perceptual trainings is to enhance students' efficient use of visual representations to solve domain-relevant tasks (Kellman & Massey, 2013), this was our main outcome of interest. The test contained 19 items, 17 of which were multiple-choice questions and two of which were short answer questions. The items covered a range of difficulty levels, including the ability to recall and understand information about chemical isomers and the ability to apply and analyze knowledge about isomers. We created two versions of the test and counterbalanced which version students received as the pretest and the posttest. The test had high reliability with Cronbach's $\alpha = .81$.

To get additional insights into students' ease of solving the chemistry tasks, we asked them to rate their cognitive load on a 9-point scale, using an instrument from prior research (Kirschner et al., 2010; Paas & Van Merrienboër, 1994). Because the measure contained only one item, its reliability cannot be quantified. We also assessed the time students took to solve the tasks based on log files. These assessments were obtained at the level of the individual student.

Interactions With the Perceptual Training. To assess how students interacted with the perceptual training (Research Question 2), we used the log data generated by the perceptual training. Specifically, we computed students' error rates as the proportion of perceptual tasks on which students made a mistake on the first attempt at solving the task. Further, we computed training time as the amount of time between students' start and finish of the perceptual training. These metrics were obtained at the level of the small group for the collaborative condition, and at the level of the individual student for the individual condition.

Gesturing. To assess how students used gestures during the perceptual training (Research Question 3), we videotaped 12 small groups, randomly selected from different sections. We then combined a top-down approach with a bottom-up, grounded theory approach (Muller, 2014) to develop a coding scheme. Specifically, the only theory-driven, top-down aspect of our approach was to distinguish deictic and representational gestures. Following McNeill (1992), we defined deictic gestures as those gestures that indicate a concrete referent in the environment. In our videos, these were exclusively gestures where students pointed at an object on the screen. We defined representational gestures as gestures that resembled the referent or indicated a movement. In our videos, representational gestures were gestures where students used their hand(s) to form a shape or indicated an angle between fingers, or where they used one hand to indicate a rotational movement or a flip.

The bottom-up, grounded theory aspect of our approach involved viewing the videos and identifying patterns of students' use of deictic and representational gestures that recurred across multiple student groups. We then formalized these patterns into a coding scheme. Next, we segmented the video data into attempts at solving the perceptual training tasks (e.g., if students made one mistake while solving a perceptual task and tried again, this yielded two segments; if they made two mistakes and tried again after each mistake, this yielded three segments). Finally, the coding scheme was applied to the resulting 303 segments.

The resulting coding scheme differentiated whether students' gestures occurred in response to disagreement or hesitation, after students made an error, after students submitted a correct answer, or spontaneously. We operationalized disagreement as students shaking their head or the hand; hesitation as pauses of 1 second or longer, as tilting the top of the head from side to side, or as slowly raising the hand and holding it up; errors and correct answers as steps that were highlighted as incorrect or correct by the perceptual training, respectively; and spontaneous gestures as gestures that occurred in the absence of the instances just described.

Further, the coding scheme distinguished deictic-holistic gestures (i.e., a student points at a whole answer choice, such as the top-right choice in Figure 3; usually this involved pointing at only one answer choice), deictic-feature gestures (i.e., a student points at visual features, such as the green sphere in Figure 3; usually this involved pointing at a few features sequentially), and representational gestures (i.e., a student made a hand movement or a hand shape that resembled the referent, such as flipping two fingers to indicate that the molecule is flipped or holding the fingers at an angle to indicate an angle between two or more bonds). These codes were mutually exclusive. In addition, for cases where students only made deictic-holistic gestures (i.e., pointing at an answer choice), the coding scheme differentiated common interaction patterns such as whether all group members pointed or only one of them pointed. Finally, the coding scheme noted cases where no gesturing occurred. The codes were mostly treated as mutually exclusive (i.e., each segment was assigned only one code), except if there was a visible shift in interaction patterns (e.g., one student pointed at an answer choice that the other student disagreed with and ruled out as the potential answer, and then showed hesitation about the remaining choices), or if gesturing occurred after students had submitted the correct answer. Table 1 shows the resulting coding scheme.

Interrater reliability was established on 16% of the data by two independent coders. Interrater reliability on distinguishing deictic-holistic, deictic-feature, and representational gestures was high with $\kappa = .92$. Interrater reliability for applying the coding scheme as a whole was high with $\kappa = .90$.

Spatial Skills. Spatial skills are a predictor of students' gesturing (Hostetter & Alibali, 2007), as well as of students' learning with visual representations (NRC, 2006), particularly in chemistry (Stieff, 2013; Wu & Shah, 2004). Therefore, we assessed students' spatial skills with the Vandenberg & Kuse test (Peters et al., 1995), which measures students' ability to mentally rotate objects; a type of spatial skills necessary for translating between 2D and 3D representations. This test has been used in prior research on the role of spatial skills for chemistry learning (e.g., Stieff, 2007). The test had acceptable reliability with Cronbach's $\alpha = .66$. This assessment was obtained at the level of the individual student.

Procedure

Students completed the pretest and the spatial skills test prior to the chemistry lab session. They were given access to the tests 24 h before the class via their online course management system. During the lab session, they first worked on the regular class activities. The last 15 minutes of the lab session were dedicated to the perceptual training, which students completed according to their

Table 1
Gesture Codes With Definitions

Gesture code	Definition
Deictic-holistic gestures only	
D_Both-point	All students point at an answer, one of them clicks.
D_One-points-and-clicks	One student points, the other nods, the first clicks.
D_One-points-other-clicks	One student points, the other one clicks.
D_One-points-alone:	One student points, the other does not contribute.
Spontaneous gesturing	
D_Spontaneous-deicticFeature	A student spontaneously uses deictic-feature gestures to explain choice.
R_Spontaneous-representational	A student spontaneously uses representational gestures to explain choice.
DR_Spontaneous-deicticFeature-representational	A student spontaneously uses deictic-feature and representational gestures to explain choice.
Disagreement	
D_Disagreement-deicticFeature	Students use deictic-feature gestures to resolve a disagreement.
D_Disagreement-deicticFeature-representational	Students use deictic-feature and representational gestures to resolve a disagreement.
N_Disagreement-noGestures	Students have a disagreement but do not use any gestures to resolve it.
Hesitation	
D_Hesitation-deicticHolistic	One student hesitates, then at least one of them then makes deictic-holistic gestures.
D_Hesitation-deicticFeature	One student hesitates, then at least one of them uses deictic-feature gestures.
R_Hesitation-representational	One student hesitates, then at least one of them uses representational gestures.
DR_Hesitation-deicticFeature-representational	One student hesitates, then they use deictic-feature and representational gestures.
After an error	
D_Error-deicticFeature	Students make an error and then at least one of them uses deictic-feature gestures.
DR_Error-deicticFeature-representational	Students make an error and then use deictic-feature and representational gestures to figure it out.
Error-noGestures:	Students make an error and do not gesture to resolve it but simply try again.
After submitting the correct answer	
D_After-deicticFeature	After submitting the right answer, a student uses deictic-feature gestures.
R_After-representational	After submitting the right answer, a student uses representational gestures.
No gesturing	
N_nodding	No gesturing, but one student nods.
N_none	No gesturing occurred.

Note. Gesture codes preceded by "D" indicate deictic gestures; gesture codes preceded by "R" indicate representational gestures; gesture codes preceded by "N" indicate no gesturing.

experimental condition. If students had finished the regular class activities earlier, they were permitted to complete the perceptual training right away and were free to leave the class when they were finished. Finally, students were given access to the content posttest immediately after the lab session and had to complete it within 24 h.

Analyses

Effects on Performance on Perceptual Tasks

To address Research Question 1a (how collaboration affects students' performance on perceptual tasks), we fitted an Additive Factor model logistic regression (Cen et al., 2007) to the error rates by task for each condition. This yields a slope for each condition that describes how quickly students' error rates decline over the course of the 20 tasks. We then compared these slopes.

Effects on Posttest Measures

To examine effects of condition on the posttest measures, we modeled the effect of condition on students' posttest scores. In creating a statistical model, we first examined whether a hierarchical linear model (HLM; Raudenbush & Bryk, 2002) would be appropriate for our analyses. An HLM could account for potential nested sources of variance owing to students being nested in small groups, section, TA, or lecturing professor had an effect on students' test scores. To this end, we used an HLM with scores on the posttests as dependent measure. Level 1 modeled student characteristics (spatial skills, pretest scores both as continuous variables). Level 2 modeled student-group variables (condition as categorical variable). To test whether an adjustment was needed to account for students being nested in small groups, section, TA, or professor, we calculated intraclass correlations (ICCs) to estimate the degree of clustering due to these factors. The ICCs for our primary outcome variable of interest (i.e., the content knowledge posttest) were so small (ICCs $< .10$) that adjustments for nonindependence were not needed (Cress, 2008). Consequently, we conducted analyses of covariance (ANCOVAs).

Second, we tested whether spatial skills should be included as a moderator in the ANCOVA model. To this end, we used an ANCOVA with students' posttest scores as the dependent measure, their pretest scores and spatial skills as the covariate, and condition as the independent variable. We then added an interaction between the continuous spatial skills variable and the categorical condition variable to test for aptitude-treatment interactions (Park & Lee, 2003) that would indicate a moderation effect. We found no aptitude-treatment interaction between spatial skills and condition ($F < 1$). However, when removing the aptitude-treatment interaction, spatial skills remained a significant predictor (p = .002). Thus, we addressed Research Question 1b (how collaboration affects the effectiveness of the perceptual training with respect to performance on chemistry tasks) using the ANCOVA model with spatial skills as a covariate but without the aptitude-treatment interaction. Further, we used the same ANCOVA model to test effects on cognitive load and on time spent on the posttest, except that cognitive load / time spent on the posttest was the respective dependent measure and cognitive load/time spent on the pretest was added as a covariate. We checked for interactions between cognitive load or time spent with condition, but found no such effects (Fs < 1) and hence did not include such interactions in our ANCOVA model.

To address Research Question 2a (whether students' interactions with the perceptual training are affected by collaboration), we used a multivariate ANCOVA with error rates and training time as the dependent measures, students' pretest scores and spatial skills as covariates, and condition as the independent variable.

Mediation Effects

We investigated Research Question 2b (whether students' interactions with the perceptual tasks mediate potential effects of collaboration on students' benefit from the perceptual training) for those interaction metrics where the ANCOVA model revealed a significant condition effect. In that case, we then conducted the mediation analysis using causal path analysis. In contrast to conventional mediation analysis (e.g., Baron & Kenny, 1986), causal path analysis provides a unified framework for simultaneously estimating multiple effects and separating direct from indirect effects in one coherent statistical model (Spirtes et al., 2000). Causal path analysis models contain all known effects in the given data set. In contrast, conventional mediation analyses use a series of independent regression analyses that do not take all known effects into account. This can result in unreliable estimates because it does not accurately describe the data.

To construct the causal path analysis model, we used the Tetrad 6.8.1 program, which allows users to search for models that are theoretically plausible and consistent with the data. The *independent variable* was condition, *dependent variables* were pretest, posttest, and spatial skills. As *mediators*, we considered error rates and training time, after verifying that these variables were significantly affected by condition because this is a prerequisite for a mediation. Specifically, in Tetrad, the user first specifies assumptions that constrain the space of theoretically plausible models. The assumptions we specified included that pretest and spatial skills cannot cause condition; that condition cannot cause pretest or spatial skills; that pretest, spatial skills, and condition can cause error rates and training time, but not vice versa; and that the pretest, spatial skills test, condition, error rates, and training time can cause the posttest, but not vice versa. Second, the user can select an algorithm that find the model with the best model fit among models in the search space (Spirtes et al., 2000). We used the Fast Greedy Equivalence Search algorithm (FGES).

Exploratory Analyses of Interaction Patterns

We conducted exploratory analyses to shed additional light into how students' interactions may affect their learning gains. First, we were interested in how gesturing affects students' interactions with the perceptual-fluency training. To this end, we correlated the different types of students' gestures (i.e., deictic-holistic, deictic-feature, and representational gestures) with error rates and training time.

Second, we were interested in whether one student in a small group might have been more dominant than another, and whether this might have led to one student having higher learning gains. To address this question, we created three measures that we consider proxies for dominance because they describe the impact one student had on the collaborative interactions. First, we assessed the proportion of problems on which each student was the first to gesture. Second, for problems on which students had a disagreement, we computed the proportion of problems for which each student convinced the other student that their answer was correct. Third, we created a categorical variable that indicated whether or not the student controlled the mouse and was responsible for clicking on the answer while working on the perceptual training. We computed these measures for the 12 dyads for whom we had video data. We then correlated these measures with posttest scores.

Owing to the small sample size, we used Kendall's tau correlations (r_τ) for the exploratory analyses.

Results

In the following analyses, we report η^2 (η_p^2) and d for effect sizes. According to Cohen (1988), an effect size η_p^2 of .01 or d of .2 corresponds to a small effect, η_p^2 of .06 or d of .5 to a medium effect, and η_p^2 of .14 or d of .8 to a large effect. Table 2 shows the means and standard deviations for all conditions and measures.

Prior Checks

Students were excluded from the analysis if they had failed to complete the pretest or the posttest, or if they were absent during the

Table 2
Means and Standard Deviations (in Parentheses) for Each Dependent Measure by Condition

Measure	Individual (n = 213)	Collaborative (n = 243)	Average (Total N = 456)
Spatial skills	.93 (.09)	.93 (.11)	.93 (.10)
Content knowledge			
Pretest	.64 (.14)	.65 (.14)	.64 (.14)
Posttest	.79 (.13)	.77 (.12)	.78 (.12)
Cognitive load			
Pretest	4.39 (1.38)	4.44 (1.41)	4.42 (1.39)
Posttest	4.15 (1.44)	4.22 (1.47)	4.19 (1.45)
Duration (in minutes)			
Pretest	12.03 (6.19)	11.42 (5.88)	11.7 (6.03)
Posttest	9.58 (6.07)	9.06 (5.00)	9.31 (5.52)
Error rates	.35 (.16)	.22 (.14)	.30 (.16)
Training time (in minutes)	4.51 (2.53)	5.29 (1.89)	4.92 (2.25)

chemistry lab session. Of the 527 students, 45 failed to complete the pretest. An additional 21 students failed to complete the posttest. Four students did not attend class, and one student had to leave class before the perceptual training. Hence, the final sample was $N = 456$.

We checked for differences prior to the intervention. A multivariate ANOVA with pretest scores and spatial skills as dependent measures showed no differences between conditions on pretest, $F(1, 454) = 1.77$, $p = .18$ or spatial skills ($F < 1$). Further, we verified that there were pretest-to-posttest learning gains. A repeated measures ANOVA with pretest and posttest as repeated dependent measures showed significant learning gains, $F(1, 455) = 375.16$, $p < .001$, $\eta_p^2 = .45$.

Finally, we checked whether students complied with the collaboration instructions. Besides our informal observations in the classroom that confirmed students' compliance with the collaboration instructions, we examined the video data of the 12 small groups who had been selected for videotaping. We found that students did not talk to each other. Further, in 88.10% of the observed cases, students agreed with their partner before clicking, either by both contributing gestures or by nodding agreement. In the 12 small groups we observed, all students within the small groups contributed to the collaborative task.

Performance on Perceptual Tasks

To address Research Question 1a (whether collaboration affects students' performance on perceptual tasks), we inspected the learning curves for each condition. Figure 5 shows the learning curve that was estimated by the Additive Factor Model. The slope estimates were .03 for the individual condition and .03 for the collaborative condition. Hence, students in both conditions showed improvements in perceptual fluency. Their improvement happened at virtually the same rate. This finding does not support either Hypothesis 1.1a or 1.2a. However, Figure 5 shows that students in the individual condition had overall higher error rates (see analyses for Research Question 2).

Performance on Chemistry Tasks

To address Research Question 1a (whether collaboration affects students' performance on chemistry tasks), we examined the condition effects of the ANCOVA model using the content knowledge posttest as dependent measure. Results showed a significant effect of condition with a small effect size, $F(1, 452) = 4.93$, $p = .03$, $\eta_p^2 = .01$, such that students who worked individually on the perceptual training outperformed students who worked collaboratively. This finding is in line with Hypothesis 1.2b that was derived from sociocognitive research. There were no significant effects of condition on cognitive load or test duration (Fs < 1).

Interactions With Perceptual Tasks

To investigate Research Question 2a (whether students' interactions with the perceptual training are affected by collaboration), we inspected the condition effects on students' error rates and training time. Results showed significant effect of condition with a large effect on error rates, $F(1, 342) = 73.68$, $p < .001$, $\eta_p^2 = .18$, and with a small effect on training time, $F(1, 342) = 10.00$, $p < .01$, $\eta_p^2 = .03$. Students in the collaborative condition made significantly fewer errors and spent significantly more time on the perceptual training than students in the individual condition.

To address Research Question 2b (whether students' interactions with the perceptual tasks mediate effects of collaboration on students' benefit from the perceptual training), we examined the path analysis model outputted by the model search described above. Figure 6 summarizes the best-fitting model the FGES algorithm found. The model fits the data well $\chi^2(6, N = 346) = 4.30$, $p = .64$).[2] The model reveals that error rates partially mediated the effect of condition on the posttest. Specifically, the average student in the collaborative condition made .13 fewer errors per perceptual task than the average student in the individual condition. For each error on the perceptual tasks, the average student's performance on the posttest dropped by .11. Thus, taken together, the impact of the mediated effect of collaboration on the posttest was an increase of 1.43% ($-.13 \times -.11 = .0143$). With respect to training time, the causal path analysis revealed no significant mediation. Additionally, there was a direct negative effect of collaboration on the posttest with a strength of $-.04$, which outweighed the mediated positive effect. Hence, the model is in line with the overall finding that collaboration decreased learning outcomes. The net effect is thus a decrease of 2.57% on the posttest due to collaboration ($-.04 + .0143 = -.0257$), which corresponds to about a third of a letter grade. Note that these effects are achieved in the context of a five-minute intervention and control for the effects of pretest and spatial skills.

Use of Gesture During the Perceptual Training

To investigate Research Question 3 (how students use gestures during the perceptual training), we examined the prevalence of the codes that emerged from the grounded theory approach described above. Table 3 shows the frequency of each code. In the following, we report patterns that stand out and provide a qualitative description of representative examples from the video data.

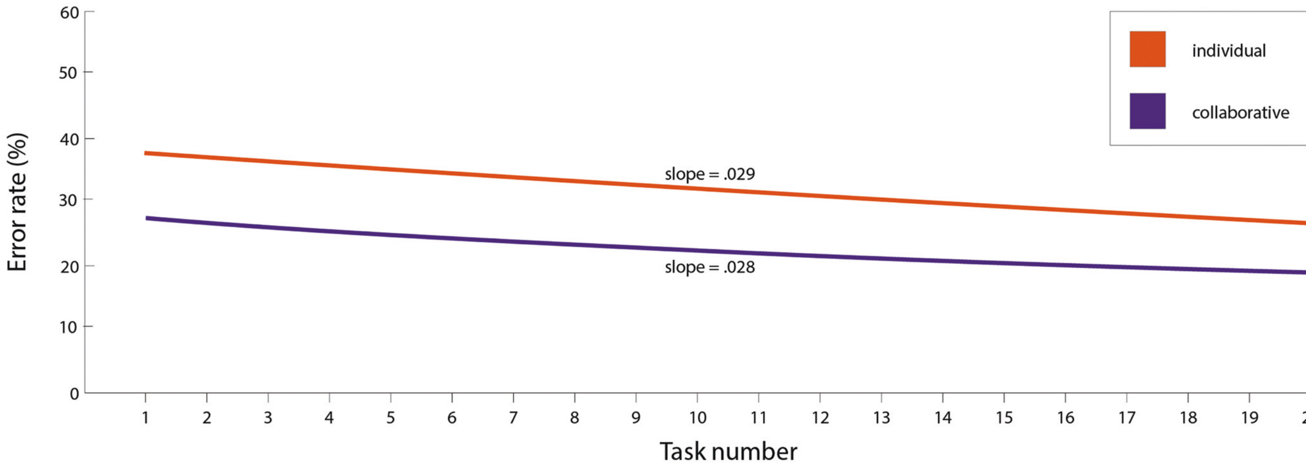
In most cases ($n = 129$), students pointed holistically at an answer. For example, Keith[3] and Hakeem viewed the perceptual task for 3 seconds, and then both pointed at an answer choice at the same time. Three seconds later, Hakeem clicked on that answer choice. Their answer was correct, and 1 second later, Hakeem clicked on the "Done" button to move to the next task.

All partners in a small group were generally involved in the decision to click ($n = 127$). Involvement in the decision was explicit if both students pointed (as in Keith and Hakeem's example) or if one student nodded agreement. For example, Emma and Seiji viewed the perceptual task for 6 seconds, then Seiji pointed at an answer choice. 5 seconds later, Emma nodded in agreement. One second later, Seiji clicked on the answer choice. Alternatively,

[2] To clarify the logic of hypothesis testing in causal path analysis: A nonsignificant p value is desirable. That is, the usual logic of hypothesis testing is inverted in causal path analysis. The p value reflects the probability of seeing as much or more deviation between the covariance matrix implied by the estimated model and the observed covariance matrix, conditional on the null hypothesis that the model that we estimated was the true model. Thus, a low p value means the *model* can be rejected, and a high p value means it cannot. Conventional thresholds are .05 or .01, but like other alpha values, this is somewhat arbitrary. The p value should be higher at low sample sizes and lowered as the sample size increases, but the rate is a function of several factors, and generally unknown.

[3] All names are pseudonyms to protect students' anonymity.

Figure 5
Learning Curves Estimated by the Additive Factor Model for Each Condition



Note. See the online article for the color version of this figure.

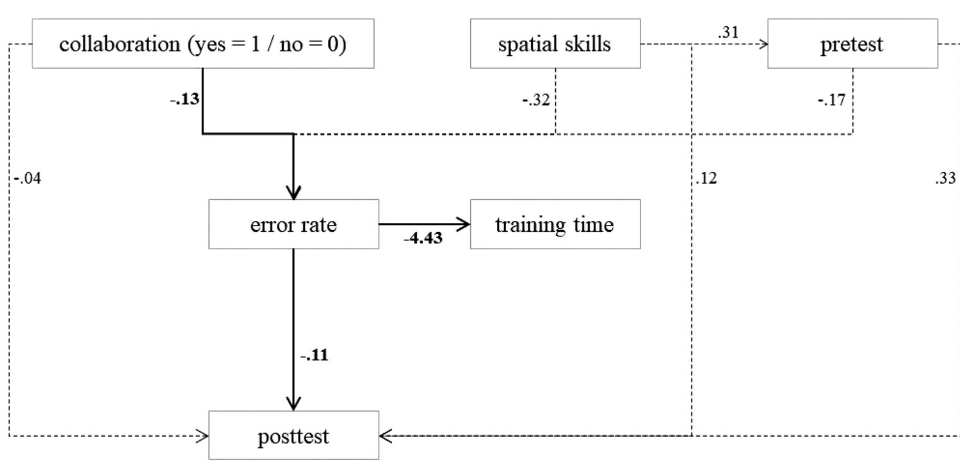
involvement in the decision was implicit if one student pointed and the other student clicked. For example, Charleen and Isaak viewed the perceptual task for 5 seconds. Then, Charleen pointed at the top-right choice. 2 seconds later, Isaak clicked on that choice, hence implicitly agreeing with Charleen's contribution.

If gesturing went beyond deictic-holistic pointing at answers, deictic-feature gestures and representational gestures occurred equally often per segment (*n* = 62 for both). For example, Daphne and Judith used deictic-holistic gestures: They viewed the perceptual task for 12 seconds, then Daphne pointed at a specific atom in the given representation. Two seconds later, both pointed at one of the answer choices, and Daphne clicked on it 1 second later. As another example, Mateo and Ben used a representational gesture: They viewed the perceptual task for 4 seconds, then Mateo pointed at the top-right choice and made a representational gesture by orienting his finger to indicate the direction of a bond relative to other bonds. Seven seconds later, Ben clicked on that choice. Informally, we noticed that students seemed to make more representational gestures when the perceptual tasks were particularly difficult (e.g., involving molecules with stereocenters).

Gestures beyond deictic-holistic pointing usually occurred in response to an event (*n* = 99), as opposed to spontaneously (*n* = 38), as illustrated by Daphne and Judith's and Mateo and Ben's examples. A common event that triggered gesturing was disagreement among partners (*n* = 21). For example, Mehren and Daisy viewed the perceptual task for 4 seconds. Mehren pointed at an answer choice, but Daisy shook her head. For the next 11 seconds, both provided a series of deictic-feature gestures (e.g., pointing at bonds and atoms) and representational gestures (e.g., orienting the hand or a finger to indicate the angle between two bonds). Then, Mehren pointed at a different answer choice, and Daisy agreed by nodding. One second later, Mehren clicked the answer choice.

Another event that triggered gesturing, although less common, was hesitation by a partner (*n* = 13). For example, Keith and

Figure 6
Causal Path Analysis Found by the FGES Algorithm



Note. Values are unstandardized coefficients. Bold paths indicate condition effects.

Table 3
Frequency of Gesture Codes

Gesture code	Frequency
Deictic-holistic gestures only	
D_Both-point	38
D_One-points-and-clicks	15
D_One-points-other-clicks	75
D_One-points-alone	2
Total	129
Spontaneous gesturing	
D_Spontaneous-deicticFeature	4
R_Spontaneous-representational	25
DR_Spontaneous-deicticFeature-representational	9
Total	38
Disagreement	
D_Disagreement-deicticFeature	6
D_Disagreement-deicticFeature-representational	15
N_Disagreement-noGestures	4
Total	25
Hesitation	
D_Hesitation-deicticHolistic	13
D_Hesitation-deicticFeature	4
R_Hesitation-representational	3
DR_Hesitation-deicticFeature-representational	6
Total	26
After an error	
D_Error-deicticFeature	7
DR_Error-deicticFeature-representational	3
Error-noGestures	29
Total	39
After submitting the correct answer	
D_After-deicticFeature	8
R_After-representational	1
Total	9
No gesturing	
N_nodding	15
N_None	38
Total	53

Hakeem viewed the perceptual task for 4 seconds, then Hakeem pointed at two answer choices simultaneously, moving his fingers quickly up and down, as if to indicate that he was undecided between the two. One second later, Keith pointed at one of those answer choices and moved his finger to show a curve that matched the shape of the molecule. Three seconds later, Hakeem nodded and clicked on that answer choice.

Gesturing in response to errors occurred least often (n = 10). For example, Daphne and Judith viewed the perceptual task for 6 seconds, then they pointed at different answer choices. Four seconds later, Judith pointed at Daphne's choice, which she immediately clicked, but it was incorrect. One second later, Judith pointed at corresponding atoms in her original choice, the incorrect choice, and the given representation. One second later, Daphne clicked on Judith's original choice, which was correct.

Most often following an error, students clicked again on a different answer without gesturing (n = 29). For example, on a different perceptual task, Daphne and Judith's first attempt was incorrect. One second later, without further gesturing by either partner, Daphne clicked on a different choice, which was correct.

Gestures sometimes occurred after a correct response (n = 9), in which case they were mostly deictic-feature gestures (n = 8). For example, 2 seconds after clicking on the correct choice in the perceptual task just mentioned, Daphne pointed at distinct atoms in the given representation and clicked "Done" 3 seconds later.

Exploratory Analyses of Interaction Patterns

First, we computed correlations of the different types of gestures with error rates and training time for the 12 dyads for whom we had video data. All types of gestures were associated with longer training time (r_τ = .36, p = .01 for deictic-holistic gestures; r_τ = .50, $p < .001$ for deictic-feature gestures; r_τ = .62, $p < .001$ for representational gestures), but only representational gestures were associated with a reduction of error rates ($r_\tau = -.29$, p = .05).

Second, we analyzed whether our proxy measures of student dominance affected posttest scores. We found no significant correlation between the proportion of problems on which students were the first to gesture with their posttest scores (r_τ = .19, p = .29) or between the proportion of problems with disagreements on which they convinced their partner that their choice was the right one ($r_\tau = -.04$, p = .82). We found no significant effects of controlling the mouse on posttest scores ($F < 1$).

Discussion

Summary of Results

The goal of this article was to contrast conflicting predictions about whether nonverbal collaboration via gesturing enhances students' benefit from a perceptual training, compared with the common practice of working individually on perceptual trainings. Thereby, this article follows Wise and Schwarz's (2017) call to understand under what conditions collaboration is beneficial. First, we examined how gesture-based collaboration affects students' performance on perceptual tasks during the perceptual training (Research Question 1a). Our results revealed that students learning individually and collaboratively showed a parallel decrease of error rates on the perceptual tasks. This suggests that both conditions equally improved their ability to make perceptual mappings between the visual representations presented during the training. This supports neither Hypothesis 1.1a or Hypothesis 1.2a.

Second, in line with the goal of perceptual trainings to enhance students' efficient use of visual representations when solving domain-relevant tasks (Kellman & Massey, 2013), our main focus was on how gesture-based collaboration affects students' performance on chemistry tasks after the perceptual training (Research Question 1b). Our results revealed an advantage of individual learning over collaboration. Students who worked on the individual version of the perceptual training showed higher posttest performance than students who worked collaboratively on the perceptual training. Yet, they experienced the same levels of cognitive load and spent the same amount of time on the posttest as collaborating students. Taken together, this finding supports Hypothesis 1.2b, which was derived from prior sociocognitive research. Although the effect on posttest scores was small, the effect is sizable if one considers that it represents about a third of a letter grade resulting from a short intervention of about 5 minutes on average.

Third, we investigated whether gesture-based collaboration affects students' interactions with the perceptual tasks in terms of error rates and training time (Research Question 2a). Our results showed that collaborating students made fewer errors and spent more time on the perceptual training than students learning individually. We then investigated whether these changes in students' interactions with the perceptual tasks mediated condition effects on students' benefit from the perceptual training (Research Question 2b). Results from a causal path analysis showed that the decrease in collaborating students' error rates was associated with higher performance on the content knowledge posttest. Hence, decreased error rates mediated a *positive* effect of nonverbal collaboration on students' benefit from the perceptual training. However, there was an additional direct *negative* effect of gesture-based collaboration that outweighed the mediated positive effect. Thus, the model revealed a mechanism through which collaboration enhanced students' benefit from the perceptual training, but also indicates that there was an additional mechanism through which gesture-based collaboration impeded students' benefit from the perceptual training. The model found no mediating effect of increased training time.

Finally, we explored how students used gestures during the perceptual training (Research Question 3). Our analysis of video data from a subset of students showed that most gestures were holistic-deictic, and usually all partners were involved in the decision to click on an answer choice. This indicates that students' behavior was largely in line with the instructions students received to solve the perceptual tasks quickly without overthinking them while coming to an agreement with their partner. However, our analyses also showed that students engaged in a fair amount of deictic-feature and representational gesturing, in response to a variety of situations (e.g., disagreement with their partner, hesitation by their partner). Our inspection of these gestures revealed that deictic-feature and representational gestures were equally frequent and externalized students' thinking, either to convince the partner of a particular answer choice or to jointly make sense of the perceptual task. Thus, these kinds of gestures can be interpreted as a nonverbal form of explaining. Although all types of gestures were associated with increased training time, only representational gestures were associated with a reduction in error rates.

Interpretation of Findings

The main finding of the present study was that gesture-based collaboration during the perceptual training yielded lower outcomes on the posttest. Our other measures provide nuanced insights into this effect. It is noteworthy that although we found an advantage of individual learning on students' performance on domain-relevant tasks after the training, we found a parallel decline of error rates on the perceptual tasks during the training. The parallel decline suggests that gesture-based collaboration did not impede students' learning of the targeted perceptual mappings themselves. Nevertheless, the results from the posttest suggest that there were differences in the quality of perceptual-fluency acquisition: students in the individual condition were better at extracting meaningful information from visual representations when solving domain-relevant tasks, which is the ultimate goal of perceptual trainings.

One explanation of our findings is derived from sociocognitive research, which suggests that collaboration increases the effort and time needed for coordinating and communicating (Nokes-Malach et al., 2015; Kirschner et al., 2009a), and that this extra effort and time does not pay off for simple tasks (Kirschner et al., 2010). As mentioned, the perceptual tasks in our perceptual training are a simple type of task. Our findings indeed revealed that training time was increased in the collaborative condition. The positive correlation between gestures and training time indicates that communication with a partner increased training time. The causal path analysis showed that increased training time did not affect the effectiveness of the perceptual training. Thus, communicating with a partner increased training time without paying off, in essence reducing the time-efficiency of the perceptual training. At the same time, however, this finding rules out the possibility that negative effects of gesture-based collaboration might be explained by the coordination among partners detracting from learning. In other words, the extra training time itself did not impede students' benefit from the perceptual training, but neither did it enhance it.

A further explanation builds on theories suggesting that gesture and speech are inseparable processes (Goldin-Meadow, 2003; Hostetter & Alibali, 2008; McNeill, 1992). Therefore, gesturing may trigger verbal thought that could interfere with perceptual processing. Our gesture analysis revealed insights into this mechanism. We found that collaborating students used a fair amount of deictic-feature and representational gestures. These types of gestures seemed to allow students to externalize their thinking and to build on each other's externalizations, which suggests that they allowed students to engage in a nonverbal form of explaining. Thereby, these gestures may have offered a pathway for students to engage in complex explanations of the perceptual mappings.

Such explanations may have interfered with perceptual fluency in ways that are analogous to the verbal overshadowing effect. Recall that the verbal overshadowing effect is explained by verbalization creating a verbally encoded memory trace that is preferably recalled at the expense of the original perceptual memory, thereby interfering with perceptual performance after a training situation (Chin & Schooler, 2008; Schooler & Engstler-Schooler, 1990). Our finding that both conditions showed parallel declines in students' errors on the perceptual tasks during the training is consistent with this interpretation. The parallel decline suggests that gesturing may not have been directly detrimental to students' ability to create perceptual mappings between the visual representations they encountered during the training. Rather, gesturing might have triggered silent explaining that yielded a verbal memory trace, which students may have preferably retrieved at the expense of the perceptual mapping when working on the posttest. This may have caused difficulties in extracting information from visual representations when solving domain-relevant tasks on the posttest, thereby explaining lower posttest scores in the collaborative condition. The finding that collaborating students did not report higher cognitive load or spent more time on the posttest suggests that they may not have noticed these difficulties. If they had noticed that perceiving relevant visual information is difficult, they likely would have taken more time to process the representations or reported higher cognitive load while doing so. Possibly, they had an illusion of knowing how to perceive information as they had been exposed to the same number of visual representations as students in the individual condition, but they were less accurate at extracting relevant information from the representations.

Although these aspects of our results support the interpretation that gesture-based collaboration *impeded* students' benefit from the perceptual training, our results also point to a positive—albeit weaker—mechanism through which collaboration *enhanced* students' benefit from the perceptual training. This aspect of our results is in line with the view that gesturing is an important communicative mechanism that facilitates perceptual learning processes while students participate in social practices (e.g., Roschelle, 1992; Singer, 2017; Stevens & Hall, 1998). Specifically, we found a large effect of gesture-based collaboration reducing students' errors on the perceptual tasks during the perceptual training. The finding that collaborating students had altogether lower error rates than students in the individual condition is not necessarily surprising: It is likely that collaborating students (at least sometimes) caught and corrected each other's mistakes before an incorrect answer was submitted, which may explain the lower error rates in the collaborative condition. The noteworthy finding is that a reduction of errors was associated with higher outcomes on the posttest. The overall effect of collaboration on posttest scores is thus a composite of the positive mediated effect and the negative direct effect. The finding that two opposite effects account for the overall effect may explain the overall small effect size, which is, however, as argued above, of practical relevance given that it was achieved within 5 minutes. Taken together, the finding that a positive mediated effect exists alongside a direct negative effect indicates that some component of working collaboratively enhanced students' benefit from the perceptual training compared with students working individually.

What is this component? It is possible that partner feedback was more helpful than feedback from the perceptual training because it might have been more nuanced and dynamic. The feedback provided by the perceptual training only indicated that students had made a mistake, but not where the mistake was located. Hence, individual students had to discover by themselves how to correct their original response; and this may have been too difficult for them. In that case, they could have induced an incorrect perceptual mapping, which may have impeded their benefit from the perceptual training, as is a known risk of providing perceptual trainings to novices (Rau & Wu, 2018). In contrast, for collaborating students, the partner's gestures may have operated like a worked example of the perceptual task, which may have off-loaded the cognitive burden of searching for relevant visual information.

The fact that representational gestures correlated with lower error rates—whereas deictic-feature gestures did not—suggests that representational gestures offered particularly helpful feedback. Whereas deictic-feature gestures draw students' attention to static visual features that are visible on the screen (e.g., the location of a double bond or a carbon atom), representational gestures are more dynamic and use movement to indicate spatial operations such as rotation, flipping, or the shape of a molecule. The finding that deictic-feature gestures did not correlate with lower error rates might suggest that individual students were able to induce by themselves which static features of the visual representations carried important information. This finding stands in contrast to other research that has found associations between deictic gestures and learning outcomes (e.g., Cook et al., 2008). Those findings were, however, obtained in the context of conceptual learning tasks where explanations were encouraged. Our research therefore does not contradict those findings but merely suggests that deictic gestures were not particularly helpful in the context of our perceptual tasks.

In contrast, the correlation between representational gestures and error rates suggests that the promise of gesture-based collaboration lay in helping students perform spatial operations on the visual representations. This finding aligns with other studies that found an association between representational gestures and learning from spatial tasks (Chu & Kita, 2008). Further, we observed that representational gestures occurred more often during difficult parts of the perceptual training. This observation speaks to the interpretation that students may have used representational gestures to address difficulties when the perceptual tasks required them to mentally perform spatial operations on the molecules.

Additionally, it is possible that performing the representational gestures themselves allowed the gesturing student to induce the correct perceptual mappings. We note that we did not prohibit students in the individual condition from gesturing. However, we did not observe gestures among students who worked individually on the perceptual training. Thus, the social situation in the collaborative condition may have given rise to representational gestures, which might have enhanced perceptual-fluency acquisition for gesturing students. This interpretation is consistent with prior research showing that students who gesture tend to show higher learning gains from visuospatial tasks (Chu & Kita, 2008). By performing representational gestures, the student can externalize the spatial operations needed to perform to perceive mappings among the visual representations. This may have made the perceptual task accessible to the gesturing student.

Our interpretation that representational gestures may have helped prevent errors is consistent with our finding that gestures rarely occurred after errors: Because gesturing most often occurred before students submitted an erroneous response, they had already elaborated on the various options, which often prevented making an error to begin with. If an error occurred nevertheless, their gestures may have already identified a plausible alternative choice, so that additional gesturing may not have seemed necessary.

Implications for Research

Our findings expand prior research suggesting that verbal collaboration decreases students' benefit from simple verbal recall tasks (Clark & Brennan, 1991; Kirschner et al., 2009b) by showing that the same may be true for gesture-based collaboration on simple perceptual tasks. In contrast to prior research on simple recall tasks, which had suggested that the increased effort and time accounts for ineffectiveness of collaborative interventions, our results point to a different mechanism. Specifically, gesture-based collaboration may trigger verbal thought and explanation (albeit silent), which seems to interfere with perceptual fluency. Although this finding is in line with prior research on the verbal overshadowing effect (Chin & Schooler, 2008; Schooler et al., 1997), our findings show that even nonverbal gesturing can have a negative effect on students' benefit from perceptual training, pointing to a nonverbal overshadowing effect, so to speak.

In addition, our findings provide insights into the role of gesture-based communication in collaborative learning. Whereas studies that have compared collaborative to individual learning have mostly attributed benefits of collaboration to verbal communication (e.g., Chi, 2009; Dillenbourg et al., 1996), nonverbal communication via gesturing has a significant role in social interactions (e.g., Singer, 2017). Our results show that even in an instructional situation where collaboration resulted in lower posttest scores, representational gestures reduced students' errors on perceptual tasks. This finding suggests that gestures that illustrate spatial operations may be a particularly important driver of collaborative learning.

Our gesture analysis revealed that representational gestures often occurred in response to errors, disagreement with a partner, or a partner's hesitation. This finding suggests that difficulties in students' collaborative interactions may trigger gestures that may allow students to navigate a difficult learning situation with the help of their partner. The fact that difficulties offer opportunities for learning has been investigated by research on desirable difficulties (e.g., Bjork & Bjork, 2020; de Croock et al., 1998) as well as by research on productive failure (e.g., Sinha & Kapur, 2021). Future research may examine the role of representational gestures in the context of difficulties that occur in collaborative settings.

Furthermore, by revealing the impact of gesture-based collaboration on students' learning, our research connects to ongoing research on the role of gestures in learning. For example, recent findings suggest that students build gestures incrementally and that the spatial congruency of incrementally built gestures has an impact on their learning (DeLiema et al., 2021). Further, recent research shows that dialogic gestures play a role in students developing a scientific understanding of everyday experiences (Flood, 2018). To date, this research has focused exclusively on conceptual learning. Yet, it seems plausible that the acquisition of perceptual fluency may also be affected by students' incremental building of gestures and by the spatial congruency of such gestures, which could be explored in future research. More broadly, our research adds to the growing embodied cognition literature, which documents the educational impact of students' body movements such as gestures (Nathan et al., 2021).

In addition, the finding that gesture-based collaboration did not enhance students' benefit from the perceptual training opens new questions for future research on social learning situations. Many learning situations involve conceptual and perceptual learning processes in conjunction. For example, perceptual learning processes are centrally involved when surgery students assist an attending surgeon in the operating room, learning to see meaningful anatomical structures in the body, where they also listen to the surgeon's conceptual explanation of the operating procedures. Our findings suggest that more research is needed on perceptual learning in social situations because we should not take for granted that students can benefit from the perceptual aspect of such a training situation. Further, future research should investigate whether maximizing opportunities for representational gestures might increase students' benefit from perceptual learning in a collaborative context. We emphasize that we do not mean to extrapolate from our findings to social situations that involve speech but rather to highlight interesting venues for future research on perceptual learning in social contexts.

Further, our findings have implications for orchestrating individual and collaborative activities. Our findings suggest that gesture-based collaboration facilitates a positive mechanism of reducing students' errors on perceptual tasks. Future research should explore whether this effect could be leveraged by allowing students to collaborate initially until they have induced correct perceptual mappings and then transition to individual work on the perceptual training. Future research could also investigate whether such a combination might enhance students' acquisition of perceptual fluency when perceptual learning processes occur in a social context. For example, a surgery student might review a video recording of a surgery procedure after having participated in the social training situation in the operating room.

Taken together, our findings highlight the importance of understanding boundary conditions of when and why collaboration is effective and offers new venues for better understanding how collaboration affects students' benefit from perceptual training situations.

Implications for Instruction

Altogether, based on our results, we cannot recommend gesture-based collaboration for perceptual trainings. Nevertheless, our findings suggest that representational gestures could be important for perceptual learning processes, although more research is needed to explore this possibility. Future research pending, students themselves may opt to engage in additional individual practice with perceptual tasks if the only opportunity to practice efficient processing of visual information is in a social context.

For perceptual trainings, our research suggests that although they may be more effective when done individually, there is room for improvement. Specifically, when spatial processing is required to establish perceptual mappings, students may need more guidance than is commonly offered by perceptual trainings. Perceptual trainings may need to provide more nuanced guidance and feedback in a way that does not trigger elaborative processes. Future research should establish whether providing a perceptual worked example enhances the effectiveness of perceptual trainings.

Limitations and Future Directions

Our findings should be interpreted in the context of the following limitations. First, collaborating via gestures only without being allowed to talk was likely unfamiliar to students. Having to learn a new mode of communication might have affected our results. Although our first and second pilot studies indicated that students got used to this mode of communication very quickly and even in the span of the short duration of the perceptual training, future research should address this issue. For example, future studies could repeatedly expose students to collaborative versions of perceptual trainings and trace the outcomes relative to an individual version of the perceptual training over time.

Second, in order not to disturb perceptual processing, we were unable to ask students to rate their cognitive load during the perceptual training. However, cognitive load measurements provided at different times during a study can reveal different effects (Ayres, 2018). Hence, it is possible that cognitive load during the perceptual training was higher in the collaborative condition, and our study was unable to reveal such an effect. Future research could address this limitation by using alternative cognitive load measures, such as physiological measures that are unobtrusive and provide online assessments of cognitive load while students engage in other tasks (see Ayres, 2018). Doing so could yield insights into whether the task of coordinating with a partner leads to increased cognitive load during the training, as suggested by sociocognitive research (e.g., Kirschner et al., 2010).

Relatedly, the cognitive load measure we used to assess cognitive load during the posttest was a single-item rating that was repeatedly presented to students after each test item, following prior research (Kirschner et al., 2010; Paas & Van Merrienboër, 1994). Although this method is widely used, it may be less reliable than multiitem measures of cognitive load (Ayres, 2018). Future research could address this limitation by using more comprehensive cognitive load measures.

Third, it is possible that students in the collaborative condition were somewhat disadvantaged because they had to work individually on the posttest. In that respect, the difference between training and test situation was larger for students in the collaborative condition than for students in the individual condition. We think providing an individual test was appropriate because most test situations involve individual work; and this has not changed since the push toward more collaborative work in classrooms (e.g., Freeman et al., 2014). Nevertheless, future research should examine whether there are benefits of collaborative perceptual trainings that become visible only in collaborative situations. Given that ultimately, the goal of STEM instruction is not to enhance test performance but to enable students to become valuable members of scientific and professional communities, this is a worthwhile goal for educational research more broadly.

Further, our study was not designed to separate effects of collaborating from effects of gesturing. Future research should compare effects of gesturing without collaborating or collaborating without gesturing to individual learning to further disambiguate our findings. Similarly, our methods did not allow us to separate effects of observing versus performing gestures. As discussed above, we believe the role of representational gestures in perceptual learning processes is worth exploring further. Future research could investigate whether representational gestures are effective for the performing student, the receiving student, or both, by comparing students who are allowed to perform but not watch, watch but not perform, or to perform and watch representational gestures.

Likewise, it would be interesting to further investigate why collaborating via gesturing was ineffective. For example, it is possible that gestures triggered verbal thought in a way that interfered with students' perceptual learning during the training. Another possibility is that gesturing itself distracts students from the perceptual experience, without necessarily triggering verbal thought.

In addition, we note that, because the gesture analysis was performed on a subset of student groups, we cannot use it to establish causal claims. Rather, we view our findings from the gesture analysis as providing complementary insights to the quantitative analysis of learning outcome and log data, serving to identify possible mechanisms that need to be verified by future research. To do so, future research would need to collect video data on a larger sample of students and verify, for example, whether representational gestures indeed mediate a positive effect of collaboration on learning gains, relative to students working individually.

Finally, and related to the previous point, we caution against overgeneralizing our findings. As any study, ours was conducted in a specific context and with a specific population. It is possible that other learning materials, for example those that rely more on visuospatial processing than ours and involve perception of dynamic stimuli, would benefit from collaboration. Indeed, our finding regarding representational gestures suggest that the more spatial and dynamic a perceptual task is, the more it might benefit from collaboration. Thus, exploring this venue in future research could yield new insights into where the boundary conditions are located that describe whether collaboration is effective.

Conclusion

Overall, the present research advances our understanding of when and how collaboration is helpful. A quasi-experiment on a perceptual training in an undergraduate chemistry course showed that nonverbal collaboration via gesturing yielded lower learning outcomes than individual learning. However, representational gestures contributed positively to the learning outcomes of collaborating students. These findings open new questions about the support of perceptual learning processes when they occur in collaborative contexts as well as about the role of representational gestures in such contexts.

References

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, *16*(3), 183–198. https://doi.org/10.1016/j.learninstruc.2006.03.001

Ainsworth, S. (2008). The educational value of multiple-representations when learning complex scientific concepts. In J. K. Gilbert, M. Reiner, & A. Nakama (Eds.), *Visualization: Theory and practice in science education* (pp. 191–208). Springer. https://doi.org/10.1007/978-1-4020-5267-5_9

Ainsworth, S., Bibby, P., & Wood, D. (1998). Analysing the costs and benefits of multi-representational learning environments. In M. W. van Someren, P. Reimann, H. P. A. Boshuizen, & T. de Jong (Eds.), *Learning with multiple representations*. Pergamon Press.

Airey, J., & Linder, C. (2009). A disciplinary discourse perspective on university science learning: Achieving fluency in a critical constellation of modes. *Journal of Research in Science Teaching*, *46*(1), 27–49. https://doi.org/10.1002/tea.20265

Alibali, M. W., & Nathan, M. J. (2007). Teachers' gestures as a means of scaffolding students' understanding: Evidence from an early algebra lesson. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences*. Erlbaum.

Alibali, M. W., Nathan, M. J., Wolfgram, M. S., Church, R. B., Jacobs, S. A., Johnson Martinez, C., & Knuth, E. J. (2014). How teachers link ideas in mathematics instruction Using speech and gesture: A corpus analysis. *Cognition and Instruction*, *32*(1), 65–100. https://doi.org/10.1080/07370008.2013.858161

Anderson, T. L., & Bodner, G. M. (2008). What can we do about 'Parker'? A case study of a good student who didn't 'get'organic chemistry. *Chemistry Education Research and Practice*, *9*(2), 93–101. https://doi.org/10.1039/B806223B

Ayres, P. (2018). Subjective measures of cognitive load: What can they reliably measure? In R. Z. Zheng (Ed.), *Cognitive load measurement and application* (pp. 9–28). Routledge.

Baron, J. (1978). The word-superiority effect: Perceptual learning from reading. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 8, pp. 131–166). Erlbaum.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Bauer, M., Kortuem, G., & Segall, Z. (1999). *"Where are you pointing at?" A study of remote collaboration in a wearable video conference system. In Proceedings of the Third International Symposium on Wearable Computers (ISWC'99)* (pp. 151–158). IEEE Computer Society.

Bekker, M. M., Olson, J. S., & Olson, G. M. (1995). Analysis of gestures in face-to-face design teams provides guidance for how to use groupware in design. In G. Olson & S. Schuon (Eds.), *Proceedings of the DIS*

1995 Conference on Designing Interactive Systems (pp. 157–166). ACM.

Bjork, R. A., & Bjork, E. L. (2020). Desirable difficulties in theory and practice. *Journal of Applied Research in Memory & Cognition*, *9*(4), 475–479. https://doi.org/10.1016/j.jarmac.2020.09.003

Braden, R. A., & Hortin, J. A. (1982). Identifying the theoretical foundations of visual literacy. *Journal of Visual, Verbal Languaging*, *2*(1), 37–42. https://doi.org/10.1080/23796529.1982.11674354

Broaders, S. C., Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology: General*, *136*(4), 539–550. https://doi.org/10.1037/0096-3445.136.4.539

Busey, T. A., & Vanderkolk, J. R. (2005). Behavioral and electrophysiological evidence for configural processing in fingerprint experts. *Vision Research*, *45*(4), 431–448. https://doi.org/10.1016/j.visres.2004.08.021

Carraher, D., & Schliemann, A. (2002). The transfer dilemma. *Journal of the Learning Sciences*, *11*(1), 1–24. https://doi.org/10.1207/S15327809JLS1101_1

Cassell, J., Stone, M., & Yan, H. (2000). Coordination and context-dependence in the generation of embodied conversation. In M. Elhadad (Ed.), *INLG'2000 Proceedings of the First International Conference on Natural Language Generation*. Association for Computational Linguistic.

Cen, H., Koedinger, K. R., & Junker, B. (2007). Is over practice necessary? - Improving learning efficiency with the Cognitive Tutor through educational data mining. In R. Lucken, K. R. Koedinger, & J. Greer (Eds.), *13th International Conference on Artificial Intelligence in Education* (pp. 511–518).

Chase, C. C., Shemwell, J. T., & Schwartz, D. L. (2010). Explaining across contrasting cases for deep understanding in science: An example using interactive simulations. In *Proceedings of the 9th International Conference of the Learning Sciences* (Vol. 1, pp. 153–160). International Society of the Learning Sciences.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*(1), 55–81. https://doi.org/10.1016/0010-0285(73)90004-2

Chi, M. T. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*(1), 73–105. https://doi.org/10.1111/j.1756-8765.2008.01005.x

Chi, M. T., Feltovitch, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*(2), 121–152. https://doi.org/10.1207/s15516709cog0502_2

Chin, J., & Schooler, J. (2008). Why do words hurt? Content, process, and criterion shift accounts of verbal overshadowing. *The European Journal of Cognitive Psychology*, *20*(3), 396–413. https://doi.org/10.1080/09541440701728623

Chu, M., & Kita, S. (2008). Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the motor strategy. *Journal of Experimental Psychology: General*, *137*(4), 706–723. https://doi.org/10.1037/a0013157

Churchill, E. F., Trevor, J., Bly, S., Nelson, L., & Cubranic, D. (2000). Anchored conversations: chatting in the context of a document. In R. Grinter, T. Rodden, P. Aoki, E. Cutrel, R. Jeffries, & G. Olson (Eds.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 454 – 461). Association for Computing Machinery.

Ciborra, C. C., & Olson, M. H. (1988). Encountering electronic work groups: A transaction costs perspective. *Office Technology and People*, *4*(4), 285–298. https://doi.org/10.1108/eb022667

Cienki, A. (2005). Image schemas and gestures. In B. Hampe (Ed.), *From perception to meaning: Image schemas in cognitive linguistics* (pp. 421–441). Mouton de Gruyter. https://doi.org/10.1515/9783110197532.5.421

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on Socially Shared Cognition*, (pp. 127–149). American Psychological Association. https://doi.org/10.1037/10096-006

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

Cook, S. W., & Fenn, K. M. (2017). The function of gesture in learning and memory. In B. Church, S. A. Kelly, & M. W. Alibali (Eds.), *Why gesture? How the hands function in speaking, thinking and communicating* (pp. 129–153). John Benjamins Publishing Company.

Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*, *106*(2), 1047–1058. https://doi.org/10.1016/j.cognition.2007.04.010

Cope, A. C., Bezemer, J., Kneebone, R., & Lingard, L. (2015). 'You see?' Teaching and learning how to interpret visual cues during surgery. *Medical Education*, *49*(11), 1103–1116. https://doi.org/10.1111/medu.12780

Cress, U. (2008). The need for considering multilevel analysis in CSCL research—An appeal for the use of more advanced statistical methods. *International Journal of Computer-Supported Collaborative Learning*, *3*(1), 69–84. https://doi.org/10.1007/s11412-007-9032-2

Curby, K. M., & Gauthier, I. (2010). To the trained eye: Perceptual expertise alters visual processing. *Topics in Cognitive Science*, *2*(2), 189–201. https://doi.org/10.1111/j.1756-8765.2009.01058.x

Curby, K. M., Glazek, K., & Gauthier, I. (2009). A visual short-term memory advantage for objects of expertise. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 94–107. https://doi.org/10.1037/0096-1523.35.1.94

de Croock, M. B. M., Van Merrienboër, J. J. G., & Paas, F. G. W. C. (1998). High versus low contextual interference in simulation-based training of troubleshooting skills: Effects on transfer performance and invested mental effort. *Computers in Human Behavior*, *14*(2), 249–267. https://doi.org/10.1016/S0747-5632(98)00005-3

DeLiema, D., Enyedy, N., Steen, F., & Danish, J. A. (2021). Integrating viewpoint and space: How lamination across gesture, body movement, language, and material resources shapes learning. *Cognition and Instruction*, *39*(3), 328–365. https://doi.org/10.1080/07370008.2021.1928133

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*(2), 107–117. https://doi.org/10.1037/0096-3445.115.2.107

Dillenbourg, P., Baker, M., Blaye, A., & O'Malley, C. (1996). The evolution of research on collaborative learning. In P. Reimann & H. Spada (Eds.), *Learning in humans and machines: Towards an interdisciplinary learning science* (pp. 189–211). Elsevier/Pergamon Press.

diSessa, A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, *10*(2 & 3), 105–226. https://doi.org/10.1080/07370008.1985.9649008

Dreyfus, H., & Dreyfus, S. E. (1986). Five steps from novice to expert. In H. Dreyfus & S. E. Dreyfus (Eds.), *Mind over machine: The power of human intuition and expertise in the era of the computer* (pp. 16–51). The Free Press.

Emmorey, K., & Casey, S. (2001). Gesture, thought and spatial language. *Gesture*, *13*(1), 35–50. https://doi.org/10.1075/gest.1.1.04emm

Flood, V. J. (2018). Multimodal revoicing as an interactional mechanism for connecting scientific and everyday concepts. *Human Development*, *61*(3), 145–173. https://doi.org/10.1159/000488693

Fodor, J. A. (1985). Precis of the modularity of mind. *Behavioral and Brain Sciences*, *8*(1), 1–5. https://doi.org/10.1017/S0140525X0001921X

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Fussell, S. R., Kraut, R. E., & Siegel, J. (2000). Coordination of communication: Effects of shared visual context on collaborative work. In W. Kellogg & S. Whittaker (Eds.), *Proceedings of the CSCW 2000 Conference on Computer Supported Cooperative Work* (pp. 21–30). ACM.

Fussell, S. R., Setlock, L. D., Yang, J., Ou, J., Mauer, E., & Kramer, A. D. (2004). Gestures over video streams to support remote collaboration on physical tasks. *Human-Computer Interaction*, *19*(3), 273–309. https://doi.org/10.1207/s15327051hci1903_3

Gibson, E. J. (1969). *Principles of perceptual learning and development*. Prentice Hall.

Gibson, E. J. (2000). Perceptual learning in development: Some basic concepts. *Ecological Psychology*, *12*(4), 295–302. https://doi.org/10.1207/S15326969ECO1204_04

Gilbert, J. K. (2005). Visualization: A metacognitive skill in science and science education. In J. K. Gilbert (Ed.), *Visualization: Theory and practice in science education* (pp. 9–27). Springer. https://doi.org/10.1007/1-4020-3613-2_2

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Harvard University Press.

Goldin-Meadow, S., & Alibali, M. W. (2013). Gesture's role in speaking, learning, and creating language. *Annual Review of Psychology*, *64*(1), 257–283. https://doi.org/10.1146/annurev-psych-113011-143802

Goldstone, R. L., & Barsalou, L. W. (1998). Reuniting perception and conception. *Cognition*, *65*(2–3), 231–262. https://doi.org/10.1016/S0010-0277(97)00047-4

Goldstone, R. L., Landy, D., & Son, J. Y. (2008). A well grounded education: The role of perception in science and mathematics. In A. Glenberg, M. DeVega, & A. Graesser (Eds.), *Symbols, embodiment, and meaning* (pp. 327–355). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199217274.003.0016

Goodwin, C. (1994). Professional vision. *American Anthropologist*, *96*(3), 606–633. https://doi.org/10.1525/aa.1994.96.3.02a00100

Graham, J. A., & Heywood, S. (1975). The effects of elimination of hand gestures and of verbal codability on speech performance. *European Journal of Social Psychology*, *5*(2), 189–195. https://doi.org/10.1002/ejsp.2420050204

Greeno, J. G., & Hall, R. P. (1997). Practicing representation. *Phi Delta Kappan*, *78*(5), 361–367.

Guzdial, M. (1997). Information ecology of collaborations in educational settings: Influence of tool. In R. Hall, N. Miyake, & N. Enyedi (Eds.), *Proceedings of the 2nd Conference on Computer Supported Collaborative Learning* (CSCL'97) (pp. 83–90). International Society of the Learning Sciences.

Hakkarainen, K., Paavola, S., Kangas, K., & Seitamaa-Hakkarainen, P. (2013). Sociocultural perspectives on collaborative learning: Towards collaborative knowledge creation. In C. E. Hmelo-Silver, C. A. Chinn, C. K. K. Chan, & A. O'Donnell (Eds.), *The international handbook of collaborative learning* (pp. 57–73). Routledge.

Hostetter, A. B., & Alibali, M. W. (2007). Raise your hand if you're spatial: Relations between verbal and spatial skills and gesture production. *Gesture*, *7*(1), 73–95. https://doi.org/10.1075/gest.7.1.05hos

Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review*, *15*(3), 495–514. https://doi.org/10.3758/PBR.15.3.495

Hostetter, A. B., & Boncoddo, R. (2017). Gestures highlight perceptual-motor representations in thinking. In B. Church & S. A. Kelly, M. W. Alibali (Eds.), *Why gesture? How the hands function in speaking, thinking and communicating* (pp. 155–174). John Benjamins Publishing Company. https://doi.org/10.1075/gs.7.08hos

Kellman, P. J., & Garrigan, P. (2009). Perceptual learning and human expertise. *Physics of Life Reviews*, *6*(2), 53–84. https://doi.org/10.1016/j.plrev.2008.12.001

Kellman, P. J., & Massey, C. M. (2013). Perceptual learning, cognition, and expertise. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 558, pp. 117–165). Elsevier Academic Press.

Kellman, P. J., Massey, C. M., & Son, J. Y. (2010). Perceptual learning modules in mathematics: Enhancing students' pattern recognition, structure extraction, and fluency. *Topics in Cognitive Science*, *2*(2), 285–305. https://doi.org/10.1111/j.1756-8765.2009.01053.x

Kellman, P. J., Massey, C. M., Roth, Z., Burke, T., Zucker, J., Saw, A., & Wise, J. (2008). Perceptual learning and the technology of expertise: Studies in fraction learning and algebra. *Pragmatics & Cognition*, *16*(2), 356–405. https://doi.org/10.1075/p&c.16.2.07kel

Kirschner, F., Paas, F., & Kirschner, P. A. (2009a). A cognitive load approach to collaborative learning: United brains for complex tasks. *Educational Psychology Review*, *21*(1), 31–42. https://doi.org/10.1007/s10648-008-9095-2

Kirschner, F., Paas, F., & Kirschner, P. A. (2009b). Individual and group-based learning from complex cognitive tasks: Effects on retention and transfer efficiency. *Computers in Human Behavior*, *25*(2), 306–314. https://doi.org/10.1016/j.chb.2008.12.008

Kirschner, F., Paas, F., & Kirschner, P. A. (2010). Task complexity as a driver for collaborative learning efficiency: The collective working-memory effect. *Applied Cognitive Psychology*, *25*(4), 615–624. https://doi.org/10.1002/acp.1730

Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 162–185). Cambridge University Press. https://doi.org/10.1017/CBO9780511620850.011

Koedinger, K. R., & Mathan, S. (2004). Distinguishing qualitatively different kinds of learning using log files and learning curves. In J. C. Lester, R. M. Vicari, & F. Paraguacu (Eds.), *Intelligent tutoring systems 2004 log analysis workshop* (pp. 39–46). Springer.

Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The knowledge-learning-instruction framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive Science*, *36*(5), 757–798. https://doi.org/10.1111/j.1551-6709.2012.01245.x

Kozma, R., & Russell, J. (2005). Students becoming chemists: Developing representational competence. In J. Gilbert (Ed.), *Visualization in science education* (pp. 121–145). Springer. https://doi.org/10.1007/1-4020-3613-2_8

Krauss, R. M., Chen, Y., & Gottesman, R. F. (2000). Lexical gestures and lexical access: A process model. In D. McNeill (Ed.), *Language and gesture* (pp. 261–283). Cambridge University Press. https://doi.org/10.1017/CBO9780511620850.017

Kundel, H. L., Nodine, C. F., Conant, E. F., & Weinstein, S. P. (2007). Holistic component of image perception in mammogram interpretation: Gaze-tracking study. *Radiology*, *242*(2), 396–402. https://doi.org/10.1148/radiol.2422051997

Massey, C. M., Kellman, P. J., Roth, Z., & Burke, T. (2011). Perceptual learning and adaptive learning technology - developing new approaches to mathematics learning in the classroom. In N. L. Stein & S. W. Raudenbush (Eds.), *Developmental cognitive science goes to school* (pp. 235–249). Routledge.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.004

Mayer, R. E. (2009). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 31–48). Cambridge University Press.

McNeill, D. (1992). Guide to gesture classification, transcription, and distribution. In D. McNeill (Ed.), *Hand and mind. What gestures reveal about thought* (pp. 75–104). University of Chicago Press.

Meyer, A. S., Belke, E., Telling, A. L., & Humphreys, G. W. (2007). Early activation of object names in visual search. *Psychonomic Bulletin & Review*, *14*(4), 710–716. https://doi.org/10.3758/BF03196826

Moores, E., Laiti, L., & Chelazzi, L. (2003). Associative knowledge controls deployment of visual selective attention. *Nature Neuroscience*, *6*(2), 182–189. https://doi.org/10.1038/nn996

Muller, M. (2014). Curiosity, creativity, and surprise as analytic tools: Grounded theory method. In J. S. Olson & W. A. Kellogg (Eds.), *Ways of knowing in HCI* (pp. 25–48). Springer. https://doi.org/10.1007/978-1-4939-0378-8_2

Nathan, M. J. (2008). An embodied cognition perspective on symbols, grounding, and instructional gesture. In M. de Vega, A. Glenberg, & A. Graesser (Eds.), *Symbols and embodiment: Debates on meaning and cognition* (pp. 375–396). Oxford University Press.

Nathan, M. J., Eilam, B., & Kim, S. (2007). To disagree, we must also agree: How intersubjectivity structures and perpetuates discourse in a mathematics classroom. *Journal of the Learning Sciences*, *16*(4), 523–563. https://doi.org/10.1080/10508400701525238

Nathan, M. J., Williams-Pierce, C., Walkington, C., Abrahamson, D., Ottmar, E., Soto, H., & Alibali, M. W. (2021). *Embodied design for mathematical imagination and cognition*. Rapid Community Reports Workshop Outcomes. https://repository.isls.org/handle/1/6852

Nokes-Malach, T. J., Meade, M. L., & Morrow, D. G. (2012). The effect of expertise on collaborative problem solving. *Thinking & Reasoning*, *18*(1), 32–58. https://doi.org/10.1080/13546783.2011.642206

Nokes-Malach, T. J., Richey, J. E., & Gadgil, S. (2015). When is it better to learn together? Insights from research on collaborative learning. *Educational Psychology Review*, *27*(4), 645–656. https://doi.org/10.1007/s10648-015-9312-8

Northedge, A. (2002). Organizing excursions into specialist discourse communities: A sociocultural account of university teaching. In G. Wells & G. Claxton (Eds.), *Learning for life in the 21st century* (pp. 252–264). Blackwell Publishing Ltd. https://doi.org/10.1002/9780470753545.ch19

Northedge, A. (2003). Enabling participation in academic discourse. *Teaching in Higher Education*, *8*(2), 169–180. https://doi.org/10.1080/1356251032000052429

NRC. (2006). *Learning to think spatially*. National Academies Press.

Park, O., & Lee, J. (2003). Adaptive instructional systems. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 651–658). Erlbaum.

Paas, F. G. W. C., & Van Merrienboër, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology*, *86*(1), 122–133. http://www.redi-bw.de/db/ebsco.php/search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1994-31418-001&site=ehost-live

Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg and Kuse mental rotations test: Different versions and factors that affect performance. *Brain and Cognition*, *28*(1), 39–58. https://doi.org/10.1006/brcg.1995.1032

Postman, N., & Wiengartner, C. (1971). Languaging. In N. Postman & C. Wiengartner (Eds.), *Teaching as a subversive activity* (pp. 98–132). Penguin Education.

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, *22*(3), 341–365. https://doi.org/10.1017/S0140525X99002022

Rau, M. A. (2017). Conditions for the effectiveness of multiple visual representations in enhancing STEM learning. *Educational Psychology Review*, *29*(4), 717–761. https://doi.org/10.1007/s10648-016-9365-3

Rau, M. A. (2018). Sequencing support for sense making and perceptual induction of connections among multiple visual representations. *Journal of Educational Psychology*, *110*(6), 811–833. https://doi.org/10.1037/edu0000229

Rau, M. A., & Patel, P. (2018). A Collaboration Script for Nonverbal Communication Enhances Perceptual Fluency with Visual Representations. In J. Kay & R. Luckin (Eds.), *Rethinking learning in the digital age: Making the learning sciences count (ICLS) 2018* (pp. 272–279). International Society of the Learning Sciences.

Rau, M. A., & Wu, S. P. W. (2018). Support for sense-making processes and inductive processes in connection-making among multiple visual representations. *Cognition and Instruction*, *36*(4), 361–395. https://doi.org/10.1080/07370008.2018.1494179

Rau, M. A., Aleven, V., & Rummel, N. (2017a). Making connections between multiple graphical representations of fractions: Conceptual understanding facilitates perceptual fluency, but not vice versa. *Instructional Science*, *45*(3), 331–357. https://doi.org/10.1007/s11251-017-9403-7

Rau, M. A., Aleven, V., & Rummel, N. (2017b). Supporting students in making sense of connections and in becoming perceptually fluent in making connections among multiple graphical representations. *Journal of Educational Psychology*, *109*(3), 355–373. https://doi.org/10.1037/edu0000145

Rau, M. A., Bowman, H. E., & Moore, J. W. (2017). Intelligent technology-support for collaborative connection-making among multiple visual representations in chemistry. *Computers & Education*, *109*(C), 38–55. https://doi.org/10.1016/j.compedu.2017.02.006

Raudenbush, S. W., & Bryk, A. S. (Eds.). (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Sage.

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science*, *7*(4), 226–231. https://doi.org/10.1111/j.1467-9280.1996.tb00364.x

Richler, J. J., Wong, Y. K., & Gauthier, I. (2011). Perceptual expertise as a shift from strategic interference to automatic holistic processing. *Current Directions in Psychological Science*, *20*(2), 129–134. https://doi.org/10.1177/0963721411402472

Richman, H. B., Gobet, F., Staszewski, J. J., & Simon, H. A. (1996). Perceptual and memory processes in the acquisition of expert performance: The EPAM model. In K. A. Ericsson (Ed.), *The road to excellence? The acquisition of expert performance in the arts and sciences, sports and games* (pp. 167–187). Erlbaum Associates.

Rimé, B., Schiaratura, L., Hupet, M., & Ghysselinckx, A. (1984). Effects of relative immobilization on the speaker's nonverbal behavior and on the dialogue imagery level. *Motivation and Emotion*, *8*(4), 311–325. https://doi.org/10.1007/BF00991870

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *Journal of the Learning Sciences*, *2*(3), 235–276. https://doi.org/10.1207/s15327809jls0203_1

Rossion, B., Collins, D., Goffaux, V., & Curran, T. (2007). Long-term expertise with artificial objects increases visual competition with early face categorization processes. *Journal of Cognitive Neuroscience*, *19*(3), 543–555. https://doi.org/10.1162/jocn.2007.19.3.543

Rossion, B., Gauthier, I., Goffaux, V., Tarr, M. J., & Crommelinck, M. (2002). Expertise training with novel objects leads to left-lateralized facelike electrophysiological responses. *Psychological Science*, *13*(3), 250–257. https://doi.org/10.1111/1467-9280.00446

Roth, W. M. (2014). Science language wanted alive: Through the dialectical/dialogical lens of Vygotsky and the Bakhtin circle. *Journal of Research in Science Teaching*, *51*(8), 1049–1083. https://doi.org/10.1002/tea.21158

Roth, W. M., & McGinn, M. K. (1997). Graphing: Cognitive ability or practice? *Science Education*, *81*(1), 91–106. https://doi.org/10.1002/(SICI)1098-237X(199701)81:1<91::AID-SCE5>3.0.CO;2-X

Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 49–69). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.005

Schnotz, W. (2014). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 72–103). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.006

Schönborn, K. J., & Anderson, T. R. (2006). The importance of visual literacy in the education of biochemists. *Biochemistry and Molecular Biology Education*, *34*(2), 94–102. https://doi.org/10.1002/bmb.2006.49403402094

Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, *22*(1), 36–71. https://doi.org/10.1016/0010-0285(90)90003-M

Schooler, J. W., Fiore, S., & Brandimonte, M. (1997). At a loss From words: Verbal overshadowing of perceptual memories. In D. L. Medin

(Ed.), *Psychology of Learning and Motivation: Advances in Research and Theory*, (Vol. 37, pp. 291–340). Academic Press. https://doi.org/10.1016/S0079-7421(08)60505-8

Schwartz, D. L. (1995). The emergence of abstract representations in dyad problem solving. *Journal of the Learning Sciences*, *4*(3), 321–354. https://doi.org/10.1207/s15327809jls0403_3

Schwartz, D. L., & Black, J. B. (1996). Analog imagery in mental model reasoning: Depictive models. *Cognitive Psychology*, *30*(2), 154–219. https://doi.org/10.1006/cogp.1996.0006

Scott, L. S., Tanaka, J. W., Sheinberg, D. L., & Curran, T. (2006). A reevaluation of the electrophysiological correlates of expert object processing. *Journal of Cognitive Neuroscience*, *18*(9), 1453–1465. https://doi.org/10.1162/jocn.2006.18.9.1453

Siegel, S. (2006). Which properties are represented in perception. In T. S. Gendler & J. Hawthorne (Eds.), *Perceptual experience* (pp. 481–503). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199289769.003.0015

Singer, M. (2017). The function of gesture in mathematical and scientific discourse in the classroom. In B. Church, S. A. Kelly, & M. W. Alibali (Eds.), *Why gesture? How the hands function in speaking, thinking and communicating* (pp. 317–329). John Benjamins Publishing Company.

Sinha, T., & Kapur, M. (2021). When problem solving followed by instruction works: Evidence for productive failure. *Review of Educational Research*, *91*(5), 761–798. https://doi.org/10.3102/00346543211019105

Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, prediction, and search* (2nd ed.). MIT Press.

Stamper, J. C., Koedinger, K. R., & McLaughlin, E. A. (2013). A comparison of model Selection metrics in DataShop. In S. K. D'Mello, R. A. Calvo, & A. Olney (Eds.), *Proceedings of the 6th International Conference on Educational Data Mining (EDM 2013)* (pp. 284–287). International Educational Data Mining Society.

Stephens, D. (1983). *Hemispheric language dominance and gesture hand preference*. University of Chicago.

Stevens, R., & Hall, R. (1998). Disciplined perception: Learning to see in technoscience. In M. Lampert & M. Blunk (Eds.), *Talking mathematics in school: Studies of teaching and learning* (pp. 107–149). Cambridge University Press. https://doi.org/10.1017/CBO9780511571251.007

Stieff, M. (2007). Mental rotation and diagrammatic reasoning in science. *Learning and Instruction*, *17*(2), 219–234. https://doi.org/10.1016/j.learninstruc.2007.01.012

Stieff, M. (2013). Sex differences in the mental rotation of chemistry representations. *Journal of Chemical Education*, *90*(2), 165–170. https://doi.org/10.1021/ed300499t

Stokes, D. (2021a). Attention and cognitive influence on perception. In *Thinking and perceiving: On the malleability of the mind* (pp. 115–140). Routledge. https://doi.org/10.4324/9781315189895-6

Stokes, D. (2021b). Perceptual expertise I: Mental architecture. In *Thinking and perceiving: On the malleability of the mind* (pp. 141–172). Routledge. https://doi.org/10.4324/9781315189895-7

Stokes, D. (2021c). Thought and perception: distinctions, similarities, and relations. In *Thinking and perceiving: On the malleability of the mind* (pp. 10–43). Routledge. https://doi.org/10.4324/9781315189895-2

Strickland, A. M., Kraft, A., & Bhattacharyya, G. (2010). What happens when representations fail to represent? Graduate students' mental models of organic chemistry diagrams. *Chemistry Education Research and Practice*, *11*(4), 293–301. https://doi.org/10.1039/C0RP90009E

Suthers, D., & Xu, J. (2002). Kükäkükä: An online environment for artifact-centered discourse. In D. Lassner (Ed.), *Eleventh World Wide Web Conference (WWW 2002)* (pp. 472–480). ACM.

Taber, K. (2002). Scaffolding learning in chemistry. In K. Taber (Ed.), *Chemical misconceptions - prevention, diagnosis and cure* (Vol. 1, pp. 67–83). Royal Society of Chemistry.

Taber, K. S. (2014). The significance of implicit knowledge for learning and teaching chemistry. *Chemistry Education Research and Practice*, *15*(4), 447–461. https://doi.org/10.1039/C4RP00124A

Tanaka, J. W., & Curran, T. (2001). A neural basis for expert object recognition. *Psychological Science*, *12*(1), 43–47. https://doi.org/10.1111/1467-9280.00308

Tang, J. C., & Leifer, L. J. (1988). A framework for understanding the workspace activity of design teams. In I. Greif (Ed.), *Proceedings of the CSCW 1988 Conference on Computer Supported Cooperative Work* (pp. 244–249). ACM.

Telling, A. L., Kumar, S., Meyer, A. S., & Humphreys, G. W. (2010). Electrophysiological evidence of semantic interference in visual search. *Journal of Cognitive Neuroscience*, *22*(10), 2212–2225. https://doi.org/10.1162/jocn.2009.21348

Vogt, S., & Magnussen, S. (2007). Expertise in pictorial perception: Eye-movement patterns and visual memory in artists and laymen. *Perception*, *36*(1), 91–100. https://doi.org/10.1068/p5262

Vygotsky, L. S. (1962). *Thought and language*. Wiley. https://doi.org/10.1037/11193-000

Wertsch, J. V. (1997). Properties of mediated action. In J. V. Wertsch (Ed.), *Mind as action* (pp. 23–72). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195117530.003.0002

Wertsch, J. V., & Kazak, S. (2011). Saying more than you know in instructional settings. In T. Koschmann (Ed.), *Theories of learning and studies of instructional practice* (pp. 153–166). Springer. https://doi.org/10.1007/978-1-4419-7582-9_9

Wise, A. F., & Schwarz, B. B. (2017). Visions of CSCL: Eight provocations for the future of the field. *International Journal of Computer-Supported Collaborative Learning*, *12*(4), 423–467. https://doi.org/10.1007/s11412-017-9267-5

Wise, J. A., Kubose, T., Chang, N., Russell, A., & Kellman, P. J. (2000). Perceptual learning modules in mathematics and science instruction. In P. Hoffman & D. Lemke (Eds.), *Teaching and learning in a network world* (pp. 169–176). IOS Press.

Wu, S. P., Corr, J., & Rau, M. A. (2019). How instructors frame students' interactions with educational technologies can enhance or reduce learning with multiple representations. *Computers & Education*, *128*, 199–213. https://doi.org/10.1016/j.compedu.2018.09.012

Wu, H. K., & Shah, P. (2004). Exploring visuospatial thinking in chemistry learning. *Science Education*, *88*(3), 465–492. https://doi.org/10.1002/sce.10126

Wylie, R., Koedinger, K. R., & Mitamura, T. (2009). Is self-explanation always better? The effects of adding self-explanation prompts to an English grammar tutor. In N. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 1300–1305).

Yamane, D. (1996). Collaboration and its discontents: Steps toward overcoming barriers to successful group projects. *Teaching Sociology*, *24*(4), 378–383. https://doi.org/10.2307/1318875

Zhen, A., Van Hedger, S., Heald, S., Goldin-Meadow, S., & Tian, X. (2019). Manual directional gestures facilitate cross-modal perceptual learning. *Cognition*, *187*(1), 178–187. https://doi.org/10.1016/j.cognition.2019.03.004

Received March 21, 2022
Revision received August 3, 2022
Accepted August 9, 2022 ■