COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 49 (2025) e70055

ISSN: 1551-6709 online
DOI: 10.1111/cogs.70055

Action Predictions Facilitate Embodied Geometric Reasoning

Fangli Xia,[a] Mitchell J. Nathan,[a] Kelsey E. Schenck,[b] Michael I. Swart[c]

[a]*Department of Educational Psychology, University of Wisconsin-Madison*
[b]*Department of Teaching and Learning, Southern Methodist University*
[c]*Wisconsin Center for Education Research, University of Wisconsin-Madison*

Received 6 December 2023; received in revised form 17 February 2025; accepted 5 March 2025

Abstract

Task-relevant actions can facilitate mathematical thinking, even for complex topics, such as mathematical proof. We investigated whether such cognitive benefits also occur for action predictions. The action-cognition transduction (ACT) model posits a reciprocal relationship between movements and reasoning. Movements—imagined as well as real ones operating on real or imaginary objects—activate feedforward mechanisms for the plausible predicted outcomes of motor system planning, along with feedback from the effect actions have on the world. Thus, ACT posits cognitive influences for making action predictions regardless of whether those actions are performed. Using a two-by-two factorial design, we investigated how generating task-relevant *action predictions* or performing task-relevant *directed actions* influenced undergraduates' ($N = 127$) geometry proof performance. As predicted, making action predictions significantly enhanced participants' proof production. No evidence suggests that combining action predictions and directed actions provided additional benefits, supporting the claim that predicting and performing actions engage overlapping processes, as theorized by ACT. *Gestural replays*, reenactments of previously performed actions during explanations, were associated with significantly better insight and proof performance for both (actor-generated) predicted actions and (investigator-generated) directed actions. Prompting people to predict task-relevant actions enhances mathematical cognition, possibly through simulated actions of transformations on imagined mathematical objects, as revealed by increased production of speech describing mathematical operations and increased production of gestural replays. We discuss

Correspondence should be sent to Fangli Xia, Department of Educational Psychology, University of Wisconsin Madison, 1025 West Johnson Street, Madison, WI 53703, USA. E-mail: fxia24@wisc.edu

the theoretical implications of these findings regarding the influences of embodied simulation of movements on cognition, and the educational implications of facilitating mathematical reasoning through interventions prompting students to perform and imagine performing task-relevant body movements.

Keywords: Action-cognition transduction; Action prediction; Directed action; Embodied cognition; Gesture; Speech; Geometric reasoning; Predictive architecture

1. Introduction

What is the value to cognition of predicting one's future actions? While it is generally acknowledged that people's thoughts can control their actions, there is still much to unpack about the ways that one's actions—both enacted and imagined—can influence one's thoughts and subsequent learning. Addressing such matters may ultimately inform how human cognitive processes are shaped by people's intentions and interactions with the world.

Philosophers and scientists have proposed that cognitive processes operate within a *predictive architecture* poised to anticipate changes in a dynamic world (Clark, 2015; Glenberg, 1997). This anticipatory orientation toward future actions affects many aspects of one's behavior, from perception (Remillard, 2003; Schubotz & von Cramon, 2001) and motor control (Ashe, Lungu, Basford, & Lu, 2006; Wolpert, Doya, & Kawato, 2003) to cognitive capabilities such as imagery and action understanding (Jeannerod, 2001), decision-making (Butz & Pezzulo, 2008), and language learning (Glenberg & Gallese, 2012). One mechanism proposed to explain the role of automatic action prediction across a range of cognitive functions involves *mental simulations* of sensorimotor experiences to anticipate the outcomes of executed movements (Wolpert et al., 2003). For example, before stepping into a puddle, one might simulate how one's foot will land and the splash it would make, and adjust the movement accordingly. Simulations of motor behavior can be repurposed for cognitive processes such as imagination and reasoning (Barsalou, 2008; Grush, 2004).

The roles of predictive processing and action simulations in thinking motivate fundamental questions about cognition: What is the relationship between action, imagination, and reasoning? Can investigators deliberately trigger the mental simulation of actions as a way to influence cognitive performance? How can *action predictions* be leveraged to improve thinking and learning, and inform educational practices? Can we observe these effects in gesture and speech? This study aims to investigate these questions by prompting students to perform or to predict outcomes of actions related to mathematical concepts in the task domain of geometry proof. Geometry proof is central to advanced mathematics and science because it draws on people's logical and generalized reasoning about universal claims regarding the nature of space and shape. We engage in this inquiry within the framework of grounded and embodied cognition, which connects intellectual and body-based experiences, including the mental simulations of those experiences (Barsalou, 2008; Shapiro, 2019; Wilson, 2002).

2. Theoretical framework

Research in grounded and embodied cognition has shown that explicitly directing students to physically perform or mentally imagine (i.e., simulate) actions that are aligned with target concepts can enhance cognitive performance, including reading comprehension (Glenberg, Gutierrez, Levin, Japuntich, & Kaschak, 2004), insight problem-solving (Thomas & Lleras, 2007, 2009), and mathematical reasoning (Nathan et al., 2014). Conversely, interfering with task-relevant movements hinders anticipatory simulated actions and selectively impairs some forms of higher-order cognitive processing, such as inference-making (Nathan & Martinez, 2015) and conceptual reasoning (Yee, Chrysikou, Hoffman, & Thompson-Schill, 2013). These findings suggest that action simulation may be a critical constituent of the cognitive system.

2.1. Grounded and embodied cognitive processes utilize a predictive architecture

A core assumption of grounded and embodied cognition is that cognitive, perceptual, and motoric processes rest on a predictive architecture (Clark, 2012; Glenberg, 1997). People do not passively wait for input before acting, and only operate reactively. Instead, they continually anticipate what is to come in streams of sensory input and are poised to respond proactively (Clark, 2015; Gibson, 1977). This can occur implicitly, through habitual behavior, such as catching a ball without conscious thought, or explicitly, by planning goal-directed actions (e.g., Pezzulo, 2008).

A significant advantage of this predictive architecture is that it engages people in continuously *simulating* how their bodies will interact with the environment. These simulations allow people to predict plausible actions and their outcomes (feedforward models) and to rapidly correct any discrepancies between expected and actual outcomes (feedback models) (Jeannerod, 2001; Wolpert et al., 2003). For example, a person preparing to pour from a teapot will anticipate how much force to exert, then monitor and employ corrective feedback if it is unexpectedly empty and lighter than anticipated. These mental simulation mechanisms underlie action control and execution and contribute to complex cognitive functions such as reasoning, language comprehension, and action understanding and imitation (Barsalou, 2009; Grush, 2004).

These simulation processes can operate outside of conscious awareness (Barsalou, 2009; Jeannerod, 2001). They can be activated automatically, for example, when observing someone else performing an action (Grèzes et al., 2004) or when reading action words (Hauk et al., 2004). However, simulations can also be *deliberately* activated, for example, when imagining another person's feeling in a given situation (Ruby & Decety, 2004). This perspective stresses the role of consciously reactivating previous actions stored in memory during imagination and thinking (Decety & Ingvar, 1990). These unconscious and conscious simulations are not necessarily independent; they might rely on overlapping neural systems (Decety & Grèzes, 2006). Conscious processes with an explicit goal enable action simulations to better align with one's goals.

2.2. Simulated actions can influence cognition and learning

There is considerable evidence that action simulation that is spontaneously generated by a speaker occurs during intellectual activity. For example, children naturally physically touch and move each object during early counting skill development (Wynn, 1990) and eventually extend this to merely pointing to each element in object pairs when establishing one-to-one correspondence (Alibali & DiRusso, 1999). These actions also appear to convey information about the speaker's cognitive state, including information that speakers may not yet be able to verbalize (Goldin-Meadow, Alibali, & Church, 1993).

Engaging action simulations can also give rise to overt actions, such as gestures. *Gestures* are naturally occurring arm and hand movements that accompany speech (McNeill, 1992; 2008) and thought (Chu & Kita, 2016). Scholars have shown that co-speech gestures can reflect the semantic content of one's speech and thought (Kita & Özyürek, 2003; McNeill, 1992). Gestures can also indicate current and emerging understandings when knowledge is developing and transitioning (Church & Goldin-Meadow, 1986; Goldin-Meadow et al., 1993). The *Gesture as Simulated Action* (Hostetter & Alibali, 2008; 2019) framework proposes that spontaneous gestures arise when the level of activation of an individual's mental simulation of sensorimotor processes exceeds a gesture threshold, which is dependent on individual, social, and contextual factors. As evidence, speakers gesture at higher rates when speaking about spatial and motoric information, when describing or imagining objects that are manipulable, and when describing object and image transformations (Chu & Kita, 2008, 2011; Hostetter & Alibali, 2010; Hostetter & Skirving, 2011). Furthermore, the form of these gestures corresponds to the form of enacted or imagined transformations (Cook & Tanenhaus, 2009; Pouw, Wassenburg, Hostetter, De Koning, & Paas, 2020).

Gestures are themselves a special type of action in that they bring action information into one's encoding of a task (Goldin-Meadow & Beilock, 2010), and schematize the most relevant visual and physical attributes of that task (e.g., Kita, Alibali, & Chu, 2017). For example, children ($N = 90$; Novack, Congdon, Hemani-Lopez, & Goldin-Meadow, 2014) who were directed to perform an abstract gesture that was conceptually congruent with a mathematical operation for preserving the equivalence relation in arithmetic equations showed superior knowledge transfer over those who performed comparable physical actions on objects (physically moving number tiles) and those who performed gestures that mimed the physical actions (pretending to move the tiles). While learning was observed in each of the conditions involving physical actions, children who performed abstract gestures formed a deeper conceptual understanding that better supported generalization and transfer.

Abstract information can also be delivered through simulated actions performed outside of the immediate task context. For example, Beilock and Goldin-Meadow (2010) used weighted disks that either did or did not correlate with disk diameter. When the weight—size relationship was switched, task performance was significantly impaired, but only for participants who provided gesture-rich explanations of their earlier actions. Furthermore, the degree of impairment correlated with the frequency of explanatory gestures after performing the task. This suggests that engaging in *gestural replays* of actions from earlier trials contributed to participants' encoding of the mathematical principles of the task and subsequent task performance. Kita et al. (2017, p. 256) noted that this showed that performing gestural replays "exerted a stronger influence on how action-relevant information was mentally represented than actually performing the actions."

2.3. Directed actions: Interventions that influence cognition and learning

Gesture production, though often spontaneous, can also be directed or manipulated, which can have cognitive influences. Goldin-Meadow, Cook, and Mitchell (2009) showed that children instructed to produce a V-handshape accompanying the verbal articulation of a correct mathematical grouping strategy were more successful at later problem-solving than those directed to form a partially correct gesture and those in the no-gesture condition. Scholars refer to these *investigator-generated* simulated actions as *directed actions* (e.g., Pier et al., 2019). Investigator-generated directed actions typically involve telling participants to precisely follow a specific set of actions. Examples include telling students to make a specific handshape when manipulating symbols (e.g., Goldin-Meadow et al., 2009), and to mimic the behaviors of a digital avatar in an embodied video game (e.g., Nathan & Walkington, 2017). *Actor-generated* actions allow for greater flexibility from the participant, who is typically told to perform relevant actions in accordance with the task demands, without precisely specifying the form. Examples include learning benefits for children instructed simply to gesture while solving math problems without directing them to gesture in any specific manner (Broaders, Cook, Mitchell, & Goldin-Meadow, 2007).

Investigator-generated directed actions that are cognitively relevant to a target concept contribute to better recall (Cherdieu, Palombi, Gerber, Troccaz, & Rochet-Capellan, 2017; Stevanoni & Salmon, 2005), superior problem-solving performance (Goldin-Meadow et al., 2009), enhanced mathematical reasoning (Smith, King, & Hoyte, 2014), and understanding of complex ideas (Enyedy, Danish, & DeLiema, 2015; Zhang, Givvin, Sipple, Son, & Stigler, 2021). These findings also clarify that not all directed actions equally benefit cognition; rather, it is the actions that enact *conceptually relevant relations* that enhance cognitive performance (Brooks & Goldin-Meadow, 2016; Lindgren & Johnson-Glenberg, 2013; Walkington, Nathan, Wang, & Schenck, 2022).

Directed actions can benefit cognition even when they are *imagined*. For example, Glenberg et al. (2004) found that young children directed by investigators to imagine (i.e., simulate) performing actions on objects described in a text showed enhanced text memory and inference-making compared to those in the control conditions, benefitting low-income families (Glenberg, Brown, & Levin, 2007) and those raised in non-English-speaking households (Adams, Glenberg, & Restrepo, 2018).

2.4. Action-cognition transduction as a mechanism linking movement and thought

The literature discussed so far has shown that physical actions, simulated actions (i.e., gestures), and imagined actions can influence cognition. How, then, do actions affect cognition? Further, how does such an account apply to simulated actions on imaginary objects that take place during mathematical activity and explanatory gestures generated outside of that activity? One proposed mechanism is *action-cognition transduction* (ACT; Nathan, 2017), which hypothesizes that goal-directed actions and explanatory gestures of prior actions affect cognitive processes through the reciprocal exchange of energy and information between motor and cognitive processes, each shaping the state of the other. ACT operates in a manner that is analogous to the transduction between physical devices (such as the reciprocity between motors and generators), electromagnetic devices (LED lights and optical sensors), and physiological systems.

The continuous interplay between acting in the world and predicting the effects of those actions on the world illustrates the close integration and reciprocity between sensorimotor and cognitive processes (Neisser, 1976/1987). This interplay is modeled by the HMOSAIC architecture for motor control (e.g., Haruno, Wolpert, & Kawato, 2003; Wolpert & Kawato, 1998). HMOSAIC regulates motor control for goal-directed behavior in dynamic and uncertain environments, adopting a predictive stance to control and monitor actions and anticipate their effects on the body and surroundings. HMOSAIC has contributed to models of complex cognitive behaviors, social interactions (Wolpert et al., 2003), and language comprehension (Glenberg & Gallese, 2012).

Building on the HMOSAIC architecture for motor control, the ACT account posits that actions on real and imagined entities activate feedforward (predictive) and feedback (responsive) signals that are associated with important relationships and behaviors of the entities. To explain how gestures on imagined objects can contribute to cognition, Kita et al. (2017) proposed the gesture-for-conceptualization hypothesis. They provided evidence showing how one's gestures *schematize* information generated via task-relevant simulated actions, such as gestures that accompany people's explanations of their analytic reasoning and problem-solving. Examples of complex cognitive phenomena that have been described by ACT include inference-making while learning from text (Nathan & Martinez, 2015) and the production of mathematically valid proofs (Nathan & Walkington, 2017).

For example, Walkington et al. (2022) showed that the directed actions for players of an embodied video game for promoting geometric reasoning improved high school students' ($N = 85$) mathematical reasoning, *especially* when players' explanations reinstated actions from gameplay using gestural replays. They found that students' gestures schematized the transformational operations they performed on the imagined mathematical objects, emphasizing invariant properties that were crucial for geometric reasoning and proof generation. In their theoretical account, an action starts with a goal structure—an intention to change the state of the world, coupled with a set of action plans to achieve that goal, subject to the preconditions of the current context and the actor. Even before the action commences, each action plan forms its own motor program for how to proceed and how it may affect the world. To achieve this, the HMOSAIC architecture generates multiple, paired predictor-controller modules. Each predictor-controller module (i.e., module_i) is a simpler motor program that guides (via controller c_i) the particulars of the intended motor behavior (e.g., solving the reverse kinematics equations for each of the joints along the arm and hand to match a specific target action) and anticipates (via predictor p_i) the set of likely next states of the motor system. Performing an action generates an afferent copy that transmits the information about the actual movements performed.

By this account, each p_i-c_i module provides, in essence, a *simulation* of the many plausible future states of the world, W_i, that are compatible with the current action plan and the prior state of the world. The active p_i-c_i modules are in continuous competition with one another.

The array of currently active p_i-c_i pairs forms a set of plausible inferences about the world. The system rewards p_i-c_i modules that are most successful at guiding the motor system within the dynamic environment in real time, selecting the best-suited simulation of the goal-directed action sequence and updating the participant's record of the cognitive state of their world model as altered by their actions.

Comparable benefits for imagining actions and performing specific actions as directed by investigators (i.e., not spontaneously generated by actors) may occur because imagining performing an action activates many of the same motor system processes as actually performing the action (Jeannerod, 2001). Experimental interventions demonstrating the benefits of investigator-generated directed actions and investigator-prompted imagined actions support hypotheses about the cognitive influences of simulated actions, providing "some of the *strongest* evidence for grounded and embodied cognition" (Nathan, 2021, p. 65; also see Barsalou, 2008). This suggests that triggering actor-generated conceptually congruent simulated actions—initiated by the speakers—alongside investigator-generated directed actions may interfere with or enhance cognitive performance, or offer evidence that imagined and actual actions use overlapping processes.

Walkington et al. (2022) showed that this same mechanism can explain how explanatory gestural replays can, in like fashion, influence cognition. This account draws on the schematizing function of gestures that highlight specific properties of actions and concepts they simulate. As with physical actions, the goal structure for generating an explanatory gesture as part of an explanation generates a set of action plans that include gestural replays. To faithfully perform a planned action, each intended gestural replay generates an afferent copy that elicits a corresponding set of predictor-controller modules. However, unlike the modules generated for directed actions, explanatory gestures need not match a specific movement sequence. Rather, they allow for any of a broad set of action sequences that enact *idealized forms* or abstractions that are congruent with the semantic and spatial information of the explanation. The schematizing role of gestural replays highlights idealized and invariant qualities of the mathematical objects and relations they are meant to convey, often more economically, as when students use hands to explain transformations they had performed earlier with their arms.

Here, we extend the ACT mechanism to account for action prediction as well. In this extension, actors prompted to predict a future action sequence, generate an action goal structure intended to be congruent with the spatial properties of a mathematical conjecture such as *The diagonals of a rectangle always have the same length*. As with the explanatory gestural replays, an action prediction generates an afferent copy that elicits a set of predictor-controller modules that simulates possible future actions. Predicted actions operate like the simulated actions of one's explanatory gestures, schematizing the idealized properties of the imagined mathematical objects and highlighting the invariant relations that may conform to or violate the conjecture being evaluated. The information generated by schematized simulated actions is then activated and brought to conscious awareness where it activates relevant lexical items (naming parts of objects) and associated relational information (e.g., geometric congruence). The most highly activated information contributes to the analytical processes involved in operational thinking and logical deduction to support one's proof and justifications, such as transformational proof or proof by contradiction.

The ACT model offers a number of testable claims about the influence of embodied interventions on mathematical reasoning. One claim is that cognitively relevant, goal-directed actions can influence and possibly benefit one's intellectual performance. As our review shows, this is an area with significant empirical support, though it is rarely applied to advanced topics in mathematics. A second claim, which is **novel**, is that cognitively relevant *action predictions* can also incur cognitive benefits. A third claim, also **novel**, is that some of the same processes that manage transduction between cognition and physical actions on tangible objects also apply to action predictions regarding imaginary objects, such as those used in mathematical activity. We investigate these claims in the task domain of secondary and post-secondary level geometry proofs.

2.5. Embodied interventions for mathematical reasoning and proof

Despite the documented influences that actions can have on mathematical cognition, relatively little attention has been paid to their role in influencing reasoning in advanced mathematical topics such as secondary and post-secondary geometry proofs. *Mathematical proof* is a fundamental disciplinary method that uses logical argumentation to demonstrate the truth of a mathematical statement beyond any doubt (Rav, 1999). *Geometric proofs* are generalized statements about universal properties and relations of space and shape, and are necessary for generating and communicating mathematical knowledge (Stylianides, 2007). However, students struggle to provide generalized mathematical ideas and construct convincing mathematical arguments (Dreyfus, 1999; Healy & Hoyles, 2000). Often, students erroneously conclude that a universal statement is true based only on specific examples (Knuth, Choppin, & Bieda, 2009), perceptual features (Jones, 2000), or authoritative sources, such as a textbook or teacher.

Educators recognize it is crucial to expand the concept of proof to allow for a wider range of argumentation beyond the traditional two-column format (Knuth, 2002; Harel & Sowder, 1998). Harel and Sowder's (2007) *transformational proof scheme* permits less formal proof types while still addressing all three necessary criteria of a mathematically valid deductive proof: (1) It must be *general*, showing the argument must be true for all possible cases; (2) involve *operational thinking*, where a prover progresses through goal-directed mental operations; and (3) exhibit *logical inference*, with conclusions drawn from valid premises. Together, these criteria require students to form arguments that evaluate the truth of mathematical conjectures based on firm deductive reasoning rather than "surface perceptions, symbol manipulation, and proof rituals" (Harel & Sowder, 1998, p. 244).

Empirical studies of proof production among mathematicians and math teachers have revealed that proof "is a richly embodied practice" (Marghetis, Edwards, & Núñez, 2014, p. 243; Schenck, Walkington, & Nathan, 2022; Walkington, Chelule, Woods, & Nathan, 2019). People often spontaneously produce gestures during verbal proofs (Nathan et al., 2021; Pier et al., 2019). For example, Nathan et al. (2021) found that *dynamic depictive gestures* (Garcia & Infante, 2012) most strongly predicted students' production of mathematically valid proofs by using their bodies to explore the invariant properties of geometric objects. Examples of these *dynamic* *depictive gestures* included simulated actions like rotating, dilating, and skewing imaginary objects.

Embodied interventions have shown promise for improving geometry proof performance. Nathan et al. (2014) found that performing cognitively relevant directed actions led to significantly more correct intuitions (snap judgments) and key mathematical insights, even when participants were not consciously aware of the cognitive relevance of those directed actions. This suggests an implicit effect of actions on geometric reasoning. When the cognitive relevance of the actions was explicitly signaled using pedagogical prompts, participants' proof production enhanced further, indicating that additional benefits can be incurred when the unconscious processes are explicitly brought into conscious awareness.

Nathan and Walkington (2017) further explored the influences of directed actions on proof performance. They designed a motion-capture video game called *The Hidden Village* that prompted high school-level geometry students to perform directed actions that mimicked behaviors of on-screen avatars. Directed actions were designed to be either cognitively relevant, capturing the key conceptual relations underlying geometry conjectures, or irrelevant. When participants were informed that these actions were relevant to the conjectures, they produced more dynamic depictive gestures and improved their mathematical insights and valid proof production. These findings suggest that cognitively relevant directed actions can enhance geometric reasoning, and this effect may be influenced by gestures produced during students' explanations.

Walkington et al. (2022) demonstrated further benefits of directed actions for enhancing proof performance through ACT by varying within-subjects the cognitive relevance of directed actions for high school students who played *The Hidden Village*. Although performing relevant directed actions did not *directly* cause students to produce more gestures or improve their reasoning, participants who produced gestures, including gestural replays of previous cognitively relevant actions, during their explanations showed significantly greater insight and higher proof performance. These findings suggest that the presence of gestural replays derived from relevant directed actions *moderated* the effect of those actions on proof performance. This may be because gestures schematize conceptual relations (Kita et al., 2017), enabling students to better utilize the conceptual information embodied in the relevant actions.

To date, empirical findings across multiple laboratory- and classroom-based studies of high school and college students engaged in geometric reasoning and proof production during single and multi-session investigations provide evidence that supports claims about the grounded and embodied nature of mathematical reasoning and the potential for action-based interventions to improve the mathematical validity of students' geometry proof practices (Nathan, Walkington, & Swart, 2022). The results of these studies suggest that for advanced areas of mathematics such as geometry proof, interventions designed to elicit cognitively relevant directed actions can enhance one's cognitive processing. The effect of these actions appears to be mediated or moderated by gestural and verbal explanatory processes. Proof performance can benefit from interventions that foster simulated actions on imagined mathematical entities and engage both verbal and nonverbal processes. Additionally, bringing unconscious, action-based forms of knowledge into conscious awareness can provide further cognitive benefits.

2.6. Reproducibility of findings of embodied cognition

While embodied interventions have shown tremendous promise, replication of empirical findings is a hallmark of a maturing area of scientific inquiry. Thus, it has been concerning that a number of peer-reviewed studies demonstrating significant positive effects of embodied cognition that employed rigorous research design methods and data analysis techniques failed to show significant effects when scholars attempted to replicate them. One particularly well-studied phenomenon is the Action Compatibility Effect (Glenberg & Kaschak, 2002), which demonstrated faster sensibility judgments for sentences describing movements (either physical or abstract) that are compatible with the direction of the response action (e.g., pulling a lever toward or away from one's body). The effect, demonstrated in several variations across multiple laboratories (e.g., Borreggine & Kaschak, 2006; Bub & Masson, 2010; Zwaan & Taylor, 2006), is typically interpreted as providing evidence that motor processing influences language processing. However, a formal replication effort (Kaschak et al., 2018; Morey et al., 2022) that included several investigators from the original studies failed to achieve replication.

There have also been investigations of the influences of gestures on cognition that have reported negative or null findings. In one study of mathematics learning of equations, children directed to mimic gestures that instantiated a successful strategy saw no advantage on an immediate posttest and performed significantly worse on a 4-week follow-up test than children who were directed to produce eye movements consistent with the gesture strategy, or to mimic the strategy in speech (Byrd, McNeil, D'Mello, & Cook, 2014). In another study (Yeo, Ledesma, Nathan, Alibali, & Church, 2017), children learned less from video-taped lessons that used gestures aimed at building conceptual links between symbolic equations and graphs than when gestures were not displayed. These negative findings, though reportedly rare (e.g., Cook, 2018; Hostetter, 2011; Roth, 2001), demonstrate that gestures do not always benefit learning and raise the possibility that any observed benefits are likely to depend upon mediators and moderators. Because of mixed findings in this area of inquiry, our investigation uses a rigorous methodology, driven by clear, theoretically motivated hypotheses and careful thought to the experimental controls and reliability of coding and analyses, including those testing for mediating and moderating roles of gestures on performance.

3. Research questions and hypotheses

The current study takes a novel approach to foster conscious simulations of actions by prompting students to predict the outcomes of task-relevant actions on mathematical entities. We investigate how such actor-generated action predictions affect geometric reasoning when made alone, and when accompanied by investigator-generated, task-relevant directed actions.

Our first research question (RQ1) asks: *Does prompting students to predict actions affect their mathematical insight and geometry proof performance*? Theoretical accounts posit that explicit anticipation activates the conscious simulations of actions (Decety & Ingvar, 1990; Pezzulo, 2008). The ACT account (Nathan, 2017) suggests that explicit action prediction influences cognition by generating goal structures congruent with mathematical tasks and simulating future actions that schematize important relations of mathematical objects. These schematized simulated actions are then activated and brought to conscious awareness, supporting analytical processes such as operational thinking and logical deduction that contribute to valid transformational proofs. Thus, we hypothesize (H1, the *action prediction hypothesis*) that prompting students to predict actions related to mathematical objects can activate simulations of object-related actions, thereby enhancing mathematical insight (Prediction 1a) and mathematically valid proof performance (Prediction 1b). This hypothesis tests the main effect of action prediction on these outcomes.

Our second research question (RQ2) asks: *Does performing directed actions enhance students' mathematical insight and geometry proof performance*? Prior research found that students who performed cognitively relevant directed actions showed superior mathematical insight and proof performance to those who performed irrelevant actions (Nathan & Walkington, 2017; Nathan et al., 2014). Thus, we hypothesize (H2, the *directed action hypothesis*) that participants who perform relevant directed actions, whether or not they make action predictions, will show superior mathematical insight (Prediction 2a) and mathematically valid proof performance (Prediction 2b) compared to students who do not perform mathematically relevant directed actions. This hypothesis tests the main effect of mathematically relevant directed action.

The third research question explores the interaction effect between action prediction and directed action. RQ3 asks: What is the combined influence of making both action predictions and directed actions on students' mathematical insight and geometry proof performance? Some theoretical accounts (e.g., Clark, 2015; Wolpert & Flanagan, 2001) suggest that engaging in any goal-directed actions automatically triggers motor prediction mechanisms, simulating the plausible sensorimotor consequences of those actions. This implies that performing directed actions, even without explicit prediction prompts, involves automatic action predictions and unconscious simulations. In contrast, when people are explicitly prompted to predict future actions (i.e., action prediction), they engage in conscious simulations of those actions. Both unconscious and conscious simulations are assumed to rely on some of the same neural systems (Decety & Grèzes, 2006). Thus, we propose the overlapping processes hypothesis (H3), which predicts that combining directed actions and action predictions will not impair performance or yield additional benefits for mathematical insight (Prediction 3a) and valid proof performance (Prediction 3b). Further, since cognitively relevant actions enhance mathematical reasoning (e.g., Nathan et al., 2022), those combining action predictions and directed actions will likely outperform control group participants who do neither in both insight (Prediction 3c) and proof (Prediction 3d).

Our final research question examines the mediating or moderating effects of gestures, which are not mutually exclusive. RQ4 asks: *How does gesture production during subsequent explanations influence the effect of making action predictions, performing directed actions, and their combination on mathematical insight and proof performance?* Previous studies exploring the *mediating* role of gesture suggest that task-relevant actions can alter gesture frequency or type during later explanations, which, in turn, benefits reasoning and performance (Cook & Goldin-Meadow, 2006; Goldin-Meadow et al., 2009). If the *gesture as a mediator hypothesis (H4-1)* holds, we expect to see an increase in dynamic (Prediction 4-1a) and nondynamic gesture production (Prediction 4-1b) when participants make action

Table 1
Summary of research questions and hypotheses

		Dependent measures		
		Insight	Proof	Gesture
RQ1:	H1: Action prediction hypothesis (Main effect)	√	√	
RQ2	H2: Directed action hypothesis (Main effect)	√	√	
RQ3	H3: Overlapping processes hypothesis (Interaction effect)	–	–	
RQ4	H4-1: Gesture as a mediator hypothesis			√
	H4-2: Gestural replay as a moderator hypothesis	√	√	

Note. √ means that an increase in the likelihood of generating correct insight, valid proof, or gesture is expected. – means that an increase in the likelihood of generating correct insight or valid proof is not expected.

predictions, perform directed actions, or do the combination, compared to the control condition. Other studies report that task-relevant actions can enhance mathematical reasoning *only* when gestural replays are produced during subsequent explanations (Beilock & Goldin-Meadow, 2010; Walkington et al., 2022). If the *gestural replay as a moderator hypothesis (H4-2)* holds, participants who produce gestural replays of earlier action predictions, directed actions, or their combination will show superior insight (Prediction 4-2a) and higher rates of mathematically valid proofs (Prediction 4-2b) compared to those who do not produce gestural replays. The hypotheses and research questions are summarized in Table 1.

4. Methods

4.1. Participants

Participants were adult students ($N = 127$; power analysis is reported below) recruited from a large university in the Midwestern United States; 38 were first-year students, 16 second-year, 30 third-year, 40 fourth-year, 1 graduate student, and 2 identified “other.” Self-reported gender indicated 63.0% female, 37.0% male participants. Ninety-one students self-identified as White/Non-Hispanic, 24 as Asian, 8 as biracial, 2 as Hispanic, and 2 as African American. All participants were English-speaking students, though 22 reported that English was not their first language. Participants had diverse mathematical coursework backgrounds: 43 participants had not taken Calculus I, 59 had completed or were enrolling in Calculus I or II at the time of the study, and 25 had taken or were taking a mathematics class beyond Calculus II. Each participant received a $20 digital gift card as remuneration. Some participants also received extra credit for completing write-ups about their participation if offered by their instructors.

4.2. Experimental design

4.2.1. Materials

4.2.1.1. Conjectures: Eight geometry conjectures were presented to participants. Each conjecture presented a mathematical assertion that focused on the general properties of two-dimensional geometric objects and was selected from a variety of secondary mathematics

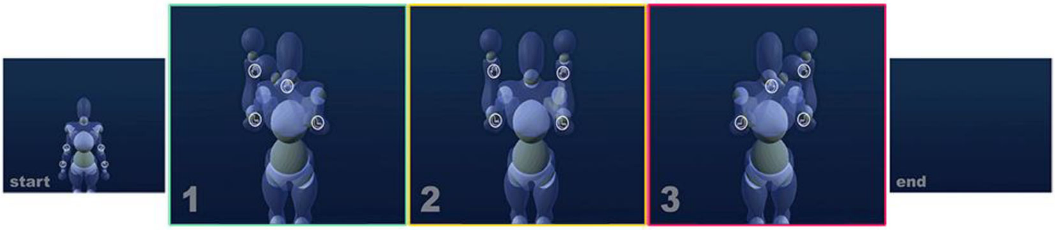

Fig. 1. An example of directed actions.
Note. Directed actions for the Rectangle Diagonals conjecture, which reads, "The diagonals of a rectangle always have the same length." These directed actions include a sequence of three poses designed to convey a key insight related to this conjecture—that the diagonals must be congruent because they each are the hypotenuses of left and right triangles with equal arm lengths.

textbooks. Supplementary Appendix A shows the text of each conjecture, an indication of whether the conjecture was always true or could be false, examples of a coded correct mathematical insight for that conjecture, and examples of its mathematically valid proof.

4.2.1.2. Directed actions and action predictions: Participants were randomly assigned to conditions in the 2 (directed actions or not) × 2 (action predictions or not) factorial design. Directed actions consisted of a series of three upper-body poses that, when performed sequentially, enacted cognitively relevant mathematical properties related to one of the eight conjectures (see an example in Fig. 1). These directed actions were derived from prior observations of the spontaneous hand and arm gestures made by students during successful proof production (Nathan & Walkington, 2017). For example, participants were asked to evaluate the truth of the conjecture, "*The diagonals of a rectangle always have the same length*," and to mimic the movements of the digital avatar, thereby performing relevant directed actions. These motions are intended to convey a key mathematical insight that the diagonals must be congruent because they each are the hypotenuse of a triangle with congruent base and height lengths. These directed action sequences were created using *The Hidden Village* pose editor (Nathan & Swart, 2021) and then turned into animated GIFs to convey continuous movements (see Fig. 1).

Action predictions consisted of partial sequences of two directed actions of a digital avatar, accompanied by prompts to predict and perform the movements the digital avatar could generate for each conjecture (see Fig. 2).

4.2.1.3. Software: The original study protocol was designed for in-person participation in a controlled laboratory setting. In light of the COVID-19 pandemic and the need for social distancing mandated by campus-wide restrictions, this experiment was conducted online via *Blackboard Collaborate Ultra* (*BBCU*), a university-approved, secure platform for collecting human-subject research data and proctored by a single experimenter. Similar to other video platforms like Zoom, BBCU is a real-time video conferencing tool that allows the moderator

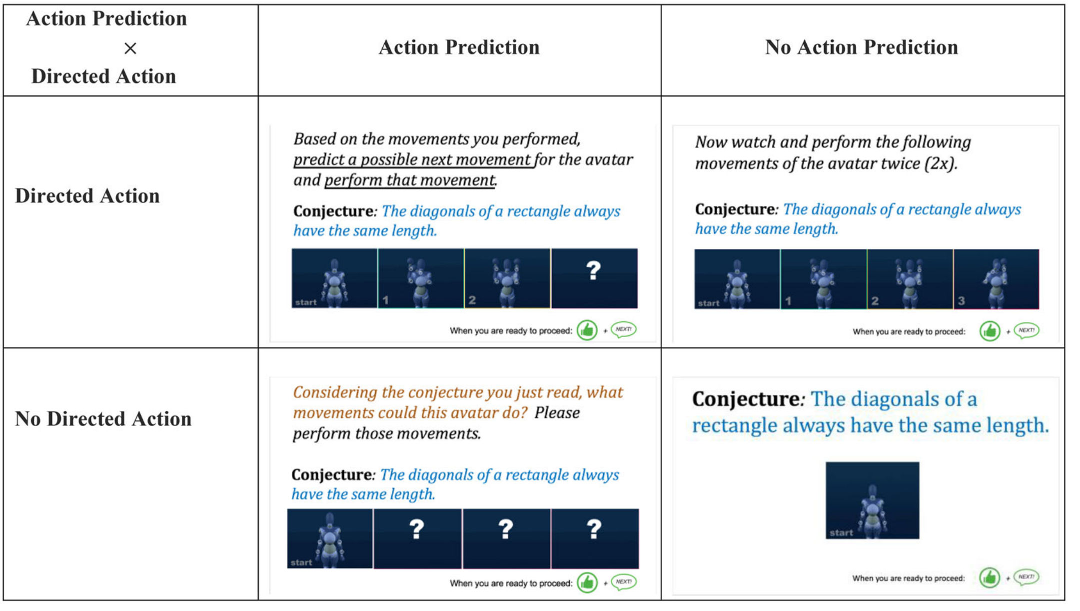

Fig. 2. 2 by 2 factorial design.

(e.g., a researcher) to share and advance presentation files, interact with participants, and observe their speech and movements in near full-screen. Sessions were video recorded with *BBCU* and *Debut* (NCH Software, 2020), a licensed screen-recording software that recorded participants' verbal responses, movements, and interactions with the researchers.

4.2.2. Measures

We measured spatial reasoning, which has been shown to predict geometry proof performance (Nathan et al., 2021) and gesture production (Hostetter & Alibali, 2007). In addition, we included exploratory measures such as general geometry knowledge, geometry interest, gender, and level of math coursework to examine their potential influence on study outcomes.

4.2.2.1. Spatial reasoning: Participants' spatial reasoning was assessed using the *Spatial Reasoning Instrument* (SRI; Ramful, Lowrie, & Logan, 2017) with 15 multiple-choice questions. This instrument originally consisted of 30 multiple-choice items based on three constructs (with 10 items per construct): mental rotation, spatial orientation, and spatial visualization. A set of 10 items was analyzed for each of the three constructs using factor analysis and ranked from highest to lowest based on factor loadings. The items with the top five factor loadings for each construct were used in this study. Scores were computed by giving 1 point for each correct answer and 0 points for each incorrect answer. Total scores were summed across all three constructs for a final score out of 15. We used total scores as a covariate in our analysis since evidence has revealed the existence of a common genetic network that underlies all spatial abilities (Malanchini et al., 2020). The original coefficient of reliability (i.e., Cronbach's alpha) for the full SRI was .85, while the reliability for our sample was 0.66.

4.2.2.2. General geometry knowledge: Participants' knowledge about geometry was assessed with 11 multiple-choice questions excerpted from the *Diagnostic Geometry Assessment* (DGA) (Masters, 2010). The DGA measures students' knowledge of three subdomains of geometry content: shape properties, transformations, and geometric measures. The original DGA measure consists of 14 items on shape properties, 10 on transformations, and 11 on geometric measurement. We selected 11 items (4 items on shape properties, 4 items on transformations, and 3 items on geometric measurement) that closely matched the content for our study. For each item in these three constructs, scores were computed by giving 1 point for each correct answer and 0.5 points for each partially correct answer. Total scores were summed across all three constructs for a final score out of 11. The whole survey had a reliability of 0.67 for our sample.

4.2.2.3. Geometry interest: We assessed participants' individual interest levels in geometry using Linnenbrink-Garcia et al. (2010) individual interest scale. This scale contains eight items. Participants rated each item from 1 (*not at all true*) to 5 (*very true*). Total scores were summed across all eight items, and the average score for each participant was used in our analysis.

4.2.2.4. Demographic information: A Qualtrics survey obtained self-reported demographic information, including participants' grade level, gender, languages spoken, race, and ethnicity. Additional information was collected about the last math course completed and current math course enrollment for each participant.

4.2.3. Procedures

Participants who had signed consent forms were randomly assigned into one of four groups using a 2×2 between-subjects design: performing *directed actions* (yes/no) and generating *action predictions* (yes/no), resulting in four conditions (see Fig. 2): (1) participants in the directed action condition ($n = 30$) were presented with the digital avatar first in its baseline pose and then as it animated a sequence of three poses. Participants were prompted to mimic the sequence of directed actions without any prompt to make predictions; (2) participants in the action prediction condition ($n = 37$) saw the avatar in its baseline pose, but were not exposed to any directed actions. They were prompted to predict the movements that the digital avatar could make to enact the geometric transformation of each conjecture, given the prompt, "*Considering the conjecture you just read, what movements could this avatar do? Please perform those movements.*"; (3) participants in the Combined condition ($n = 30$) saw the avatar in its baseline pose and partial sequence of two directed actions and the "?" symbol, and were prompted to mimic the partial sequence and then predict a "possible" third movement; (4) participants in the Control condition ($n = 30$) saw the digital avatar in its baseline pose and received no directed actions and no prompts to make action predictions.

Participants individually logged into an online experimental session via BBCU. The experiment began with instructions that were read aloud by the experimenter as participants followed along in the textual and visual instructions on the presentation slides. Experimenters facilitated participation by advancing presentation slides, qualifying instructions, clarifying questions, and prompting participants to provide clear explanations for their responses to each conjecture. Participants were instructed to stand approximately four to six feet away from their laptop, facing the web camera and the experimenter. They then were introduced to a practice-trial example conjecture to familiarize themselves with each component of their participation. After completing the practice trial, participants completed the eight conjecture tasks. For each conjecture, participants were asked to read the conjecture (see an example in Fig. 1), and, depending on condition, prompted to perform directed actions, make action predictions, do both, or not prompted to perform either, and to then answer prompts to evaluate the veracity of each statement (i.e., *true* or *false*) and provide verbal justification for their evaluation. Each participant completed a set of eight conjectures, for a total of 1016 responses. After giving videotaped responses to the eight conjectures, each participant was asked to complete the online surveys about spatial reasoning, general geometry knowledge, individual interest in geometry, and demographic information. Upon completion, participants received their compensation and were given debriefing information about the goals of the experiment.

4.3. Data coding

Videos of the experimental sessions were transcribed using the Transana video analysis software (Woods & Fassnacht, 2012). Full transcripts were segmented for each of the eight conjectures using timestamps. One participant's six responses for six conjectures were dropped from the dataset from a control condition due to an experimenter's error. This resulted in 1010 video clips being coded (127 participants $\times 8$ conjectures -6 conjectures). Transcripts were coded for three main outcomes: *insight*, *proof*, and *gestures* (described below). Supplementary Appendix A provides specific coding criteria for each conjecture and Supplementary Appendix B provides coding criteria for gestures. Inter-rater reliability is reported below.

4.3.1. Insight

Insight was coded based on evidence of participants' understanding of key mathematical ideas, or "gist," for each conjecture. Each transcript that demonstrated initial correct mathematical insight for the conjecture was coded as 1. All other responses, including those that shifted from incorrect to correct insights, were coded as zero (0).

4.3.2. Mathematically valid proof

Following Harel and Sowder's (2007) criterion for deductive proofs, participants' verbal responses were independently coded for three defining characteristics: (1) *generality* of the conjecture across the class of mathematical objects under consideration; (2) use of *operational thinking*, a systematic progression establishing a goal structure and anticipation of outcomes resulting from proposed transformations; and (3) exhibit a chain of *logical inference* with conclusions following from valid premises. Each verbalized response, including speech and gesture, was coded as 1 only if it met all three defining criteria, and zero (0) otherwise.

4.3.3. Gesture

Spontaneously generated gestures produced during verbal explanations were coded for whether they were *representational* or not (1/0). *Representational gestures* were defined as gestures that depict semantic content, either literally or metaphorically, by virtue of handshape or motion (Alibali, Heath, & Myers, 2001). Any gestures made that were not representational (e.g., beat gestures) were coded as zero (0).

If gestures were representational, they were subsequently coded as either *nondynamic gestures* or *dynamic depictive gestures*. *Nondynamic gestures* reflect only static properties of the mathematical entities or ideas they are depicting, such as forming a shape with one's hand, tracing its perimeter, or pointing to its interior angles. In contrast, *dynamic depictive gestures* must enact motion-based transformations of a mathematical object (Garcia & Infante, 2012). For example, gestures for skewing sides of the parallelogram or dilating the size of a triangle would be each coded as a dynamic depictive. A transcript was coded as *dynamic* if it included at least one dynamic depictive gesture and as *nondynamic* if it included no dynamic gestures.

4.3.4. Gestural replay

Operationally, *gestural replays* are nonverbal re-enactments of earlier cognitively relevant actions that are repeated during participants' explanations. Spontaneous gestures made during participants' explanations that could be matched to earlier actions on the same conjecture were coded as *gestural replays* (1/0), including *exact replays* (same actions made by same body parts) and *corresponding replays* that produced the same actions with different body parts. For example, for a conjecture "*The opposite angles of two lines that intersect each other are always the same*," participants were directed to cross two arms to form a set of vertical angles (i.e., cognitively relevant actions). We coded subsequent spontaneous gestures as gestural replays if participants repeated the same actions with the same body parts like arms, or if they used different body parts, such as fingers, to form vertical angles, thus conveying the same mathematical concept. Movements like using a finger to trace an angle were coded as gestures but not as gestural replays. Coded gestural replays were further classified as either *nondynamic* or *dynamic depictive gestures* (0/1) as described earlier.

4.3.5. Inter-rater reliability

Inter-rater reliability was established by having a second coder evaluate 203 (20%) of the video clips. The second coder was not involved in the original coding process. Cohen's kappa (κ) was 0.93 for insight, 0.96 for proof, 0.97 for nondynamic gestures, and 0.95 for dynamic depictive gestures. Furthermore, the second coder coded 156 (20%) of the video clips that may include gestural replays of earlier actions. Cohen's kappa (κ) was 0.90 for overall gestural replays, 0.88 for nondynamic gestural replays, and 0.95 for dynamic gestural replays. Additionally, we calculated the Shaffer's *rho* (ρ) for each code (see Eagan et al., 2017; online: https://app.calcrho.org/) to test the reliability of the inter-rater reliability given the small sample size using the common κ threshold of 0.65 (Cohen, 1960). The Shaffer's ρ

for each inter-rater reliability for each code was 0.00, which indicates acceptable type I error rates (<0.05) and sufficient sample size to estimate inter-rater reliability at the 0.65 threshold.

4.4. Analysis

4.4.1. Power analysis

At the planning stage, a power analysis was conducted a priori to determine the appropriate sample size to detect the effect of the treatment variables (directed action, action prediction, Combined, and Control) on proof performance. We used G*Power's (Faul, Erdfelder, Buchner, & Lang, 2009) ANOVA for Repeated Measures-Between with four groups and a power level of $\beta = 0.80$ and $\alpha = 0.05$. Based on previous data (Nathan et al., 2014), correlations between a participant solving repeated geometry proofs were estimated at 0.6 for an estimated effect size of $f = 0.31$ of the effect of relevant directed actions on insight performance. To the best of our knowledge, no prior studies have directly examined the effect of action predictions on geometric reasoning, leaving us without an existing effect size for reference. Thus, to ensure our design could detect an effect size smaller than $f = 0.31$, we conducted a power analysis using a medium effect size of $f = 0.25$ (i.e., $d = 0.5$), which returned a minimum sample of 120. Our final sample of 127 participants exceeds the minimum criteria, providing sufficient statistical power for our study.

It is important to note, however, that the statistical model used for data analysis—mixed-effects logistic regression models that we will describe below—differed from this ANOVA-based method used for power analysis. At the planning stage, G*Power's ANOVA-based calculation was chosen because it was the most accessible and suitable tool available for approximating the sample size. We updated our analysis approach based on recommendations from quantitative experts, which provided a more nuanced method for addressing the data structure and research questions. While the assumptions of the ANOVA-based power analysis differ from those of the mixed-effects logistic regression models ultimately used, the ANOVA-based approach provided a reasonable approximation at the planning stage to guide our sample size determination. This limitation is explicitly acknowledged here to ensure transparency and reproducibility.

4.4.2. Data analysis

We used mixed-effects logistic regression models for repeated observations of solving conjectures nested within individual students, using the *glmer* function in the R software package *lme4* (Bates, Maechler, Bolker, & Walker, 2014). For all analyses, *Participant ID* was included as a random effect, and the *Conjecture* being solved was included as an additional random effect since it was sampled from a larger pool of conjectures. Results of using *Conjecture* as a fixed effect did not change the significance levels in any of our models. To ensure that treating *Conjecture* as a random effect offers the best model, we tested models that included random slope for action prediction, directed action, and their interaction across *Conjecture*. Results did not change the significance level for all models. Furthermore, all tested models with random slopes encountered singularity issues, indicating near-zero variance for the random slopes. To further assess the necessity of including random slopes, we compared models with and without random slopes across *Conjecture* for both insight and proof outcomes using *anova()* function. Results showed no significant improvement in model fit with the inclusion of random slopes, suggesting that random slopes across *Conjecture* do not meaningfully contribute to the model. For transparency, the results of the random slope models are publicly available on the Open Science Framework.

For Research Questions 1–3, the outcome variables were accuracy of *insight* and *mathematically valid proof* (see Table 1). Following the 2×2 factorial design, we initially fitted mixed-effects logistic regression models that included treatment conditions (i.e., action prediction and directed action) and their interaction (action prediction × directed action) as the primary predictor, along with all student characteristics as covariates (i.e., gender, language, students' most advanced completed mathematics course, spatial reasoning score, general geometry knowledge, and interest in geometry).

To evaluate the contribution of these covariates, we conducted robustness checks and determined best-fit model selection using the *anova()* function, which compared models by testing for significant reductions in deviance using a chi-square distribution. Results showed that including gender, language, general geometry knowledge, interest in geometry did not significantly improve the model fit ($\chi^2 = 9.08$, $df = 4$, $p = .06$ for *insight* models, and $\chi^2 = 1.49$, $df = 4$, $p = .83$ for *proof* models). Based on these findings, we retained two statistically significant covariates in the simplified models to enhance parsimony: students' most advanced completed mathematics course and spatial reasoning. The latter was also retained as a theoretically driven covariate, given its established role in predicting geometry proof performance. In the main text, we reported the results of the simplified models, while the full results from models including all covariates are provided in Supplementary Appendix C for transparency.

The interaction term was tested in both full and simplified models and was found to be nonsignificant for both insight and mathematically valid proof. These results indicate that the study did not find evidence that combining action predictions and directed actions provided additional outcome benefits, aligning with the overlapping processes hypothesis (H3). Given the nonsignificance of the interaction term and its theoretical alignment with H3, we excluded the interaction item to re-examine the main effects of action prediction and directed action for RQ1 and RQ2. This approach maintained the integrity of the 2×2 factorial design and avoided collapsing groups.

For Research Question 4, we planned to conduct mediator and moderator analyses to examine how spontaneous gestures produced during subsequent explanations influenced the effects of making action predictions, performing directed actions, or their combination on insight and proof performance. We first examined the mediating role of gesture. We used models with occurrences of dynamic depictive or nondynamic gestures as the dependent variables, with treatment conditions as the main predictors and students' two characteristics as covariates. If there was a significant association between treatment conditions and gestures, we planned to then use gesture as a mediator variable to see if it would significantly predict insight and proof performance.

Next, we planned to examine the moderating role of gestural replays of earlier actions, with data from three treatment conditions (i.e., directed action, action prediction, and Combined) because there was no source of gestural replays in the Control group. Mixed-effects logistic

Table 2
Descriptive statistics for participants in four groups including standard deviation (SD)

	Control (n=30)	Combined (n=30)	Directed action (n=30)	Action prediction (n=37)	Total (N=127)
% native English speakers	83.33%	80.00%	83.33%	83.78%	83.68%
% Ethnicity—White	76.67%	66.67%	80.00%	64.86%	71.65%
% Ethnicity—Asian	13.31%	26.67%	16.67%	18.92%	18.90%
% Ethnicity—Other	10.00%	6.67%	3.33%	16.21%	9.45%
% Below Calculus I	40.00%	30.00%	23.33%	40.54%	33.86%
% Calculus I or II	40.00%	43.33%	60.00%	43.24%	46.47%
% Beyond Calculus II	20.00%	26.67%	16.67%	16.22%	19.70%
Average geometry interest scores (SD)	2.98 (0.91)	3.12 (0.92)	2.92 (0.94)	3.05 (0.89)	3.02 (0.92)
Average geometry test scores (SD)	10.06 (1.02)	10.36 (1.02)	10.08 (0.99)	10.05 (0.86)	10.13 (0.98)
Average spatial scores (SD)	12.33 (2.14)	12.60 (2.31)	12.60 (2.32)	12.19 (1.95)	12.42 (2.18)
Likelihood of correct insight	64.58%	76.25%	67.92%	70.61%	69.88%
Likelihood of correct proof	29.48%	35.83%	30.00%	35.47%	32.67%
Likelihood of nondynamic gesture	37.92%	38.33%	34.58%	42.90%	38.68%
Likelihood of dynamic depictive gesture	32.92%	31.67%	36.67%	37.5%	34.84%

regression models included gestural replay, treatment condition, and their interaction term as fixed effects, and students' two characteristics (spatial reasoning and most advanced completed mathematics course) were retained as covariates. We further used the *avg_comparisons* function to estimate the average marginal effects of gestural replay on insight and proof performance within each condition, while controlling for covariates.

Results from all models are reported as odds ratios (OR) (raw coefficients exponentiated) since the dependent variables in all models are dichotomous. The presentation of the results is organized by research questions. The dataset from this study is publicly available via the Open Science Framework at: https://osf.io/5xup3/.

5. Results

Descriptive statistics for demographic and performance measures (insight, proof, dynamic depictive gesture, and nondynamic gesture) are listed in Table 2. Overall, 69.88% of participants recognized key mathematical insights and 32.67% constructed a mathematically valid, generalized proof across eight conjectures. For gestures, 38.68% of participants produced at least one nondynamic gesture, and 34.84% produced at least one dynamic depictive gesture across eight conjectures.

5.1. RQ 1: Main effects of action predictions on insight and proof performance

Regression results for RQ1 are given in Table 3. For mathematical insights, Model 1 demonstrates that completing at least one math course above Calculus II and high spatial reasoning significantly increased the relative odds of generating correct insights ($OR = 4.55$,

Table 3
Estimates (and standard errors) for main effect of action prediction, directed action, and their interaction on insight and proof (RQs 1–3)

	Main effect		Interaction effect	
	Model 1: Insight	Model 2: Proof	Model 3: Insight	Model 4: Proof
Random effects: (Std. Dev.)				
Participant ID	0.65	0.45	0.66	0.45
Conjecture	0.72	1.26	0.72	1.26
Fixed effects:				
(Intercept)	−0.60 (0.64)	−3.31 (0.75)***	−0.59 (0.66)	−3.34 (0.76)***
Directed action[a]	0.17 (0.20)	−0.05 (0.18)	0.13 (0.28)	0.02 (0.26)
Action prediction[b]	0.35 (0.20)	0.38 (0.18)*	0.31 (0.27)	0.44 (0.25)
Math courses: Above Calculus II[c]	1.52 (0.32)***	0.49 (0.23)*	1.51 (0.32)***	0.49 (0.24)*
Spatial	0.09 (0.05)*	0.16 (0.05)***	0.09 (0.05)*	0.16 (0.05)***
Directed action: Action prediction			0.07 (0.39)	−0.12 (0.35)

[a]No directed action condition is the reference group.
[b]No action prediction condition is the reference group.
[c]Taking math courses that are not beyond Calculus II is the reference group.
$^{*}p < .05$ $^{***}p < .001$.

$p < .001$; $OR = 1.10$; $p = .050$, respectively). When controlling for prior math courses, spatial reasoning, and performing directed actions, action prediction was not significantly associated with the generation of mathematical insight ($OR = 1.42$, $p = .072$). Model 2 in Table 3 shows the regression results for constructing a mathematically valid proof. We see that participants who had taken at least one math course above Calculus II ($OR = 1.62$, $p = .041$) or had higher spatial reasoning ($OR = 1.18$, $p < .001$) were more likely to generate a mathematically valid proof. Prompting participants to make action prediction was significantly associated with higher odds of generating a mathematically valid proof ($OR = 1.47$, $p = .033$), compared to conditions without action prediction.

In support of H1, these results reveal that action prediction had a significant main effect on proof performance (Prediction 1b), though not for mathematical insight (Prediction 1a), even when controlling for producing directed actions.

5.2. RQ 2: Main effects of directed actions on insight and proof performance

Performing relevant directed actions was not a significant predictor of insight (Table 3, Model 1; $OR = 1.18$; $p = .397$) or proof production (Model 2; $OR = 0.95$; $p = .787$), when controlling for math course taken ($OR = 1.62$, $p = .040$), spatial ability ($OR = 1.18$, $p < .001$), and making action prediction ($OR = 1.47$, $p = .033$). Thus, there was no evidence to suggest that performing directed actions had a direct effect on insight or proof performance, which does not provide support for Predictions 2a and 2b.

5.3. RQ 3: Interaction effect between action predictions and directed actions on insight and proof performance

The interaction effects between making action predictions and performing directed actions on students' mathematical insight and geometry proof performance are reported in Table 3.

Table 4
Estimates (and standard errors) for combined group on insight and proof compared to action prediction, directed action, and control groups (RQ3: Exploratory analysis)

	Model 1: Insight	Model 2: Proof
Random effects: (Std. Dev.)		
Participant ID	0.68	0.44
Conjecture	0.72	1.26
Fixed effects:		
(Intercept)	−0.06 (0.66)	−3.00 (0.76)***
Combined	(ref.)	(ref.)
Directed action	−0.38 (0.29)	−0.32 (0.26)
Control	−0.52 (0.29)	−0.33 (0.26)
Action prediction	−0.20 (0.27)	0.11 (0.24)
Math courses: Above Calculus II[a]	1.51 (0.31)***	0.49 (0.24)*
Spatial	0.09 (0.05)*	0.16 (0.05)***

[a]Taking math courses that are not beyond Calculus II is the reference group.
$^{*}p < .05$ $^{***}p < .001$.

Taking a math course above Calculus II and showing high spatial ability were both significant predictors of mathematical insight (Model 3; $OR = 4.53$, $p < .001$; $OR = 1.10$, $p = .050$, respectively) and proof performance (Model 4; $OR = 1.64$, $p = .039$; $OR = 1.18$, $p < .001$, respectively). However, the analysis did not find evidence of a statistically significant interaction effect between action prediction and directed action on insight ($OR = 1.07$; $p = .858$) or proof performance ($OR = .88$; $p = .738$), supporting Predictions 3a and 3b.

5.3.1. Exploratory analysis

Although the interaction analysis did not find evidence of a significant interaction effect, participants in the Combined groups—who engaged in both action prediction and directed action—exhibited the highest percentages of correct insight and valid proof compared to the other three groups (i.e., action prediction alone, directed action alone, and Control) (see Table 2). To investigate whether the Combined group's performance was significantly superior to that of the other groups, post hoc comparisons (see Table 4) were conducted. The results show that the Combined group did not achieve significant advantages over the action prediction or directed action groups individually, suggesting that this study did not detect additive effects from combining action prediction and directed action. Participants in the Combined group had superior insight ($OR = 1.68$, $p = .071$; Model 1) compared to the Control group, although this difference was not significant. Overall, these findings support the overlapping processes hypothesis as predicted by ACT, suggesting that simulating actions and performing actions involves overlapping cognitive processes, as theorized by the ACT framework.

5.4. Research question 4: The role of spontaneous gesture

RQ4 explores how spontaneous gestures produced during subsequent explanations influence the relationship between geometric reasoning and action prediction, directed

Table 5
Estimates (and standard errors) for producing dynamic depictive and nondynamic gestures in action prediction, directed action, and combined groups compared to control group (RQ4: Mediation analysis of gesture)

	Model 1: Dynamic gestures	Model 2: Nondynamic gestures
Random effects: (Std. Dev.)		
Participant ID	0.82	0.35
Conjecture	1.29	0.10
Fixed effects:		
Intercept	−2.68 (0.85)**	0.48 (0.61)
Control	(ref.)	(ref.)
Combined	−0.20 (0.32)	−0.01 (0.23)
Directed action	0.16 (0.31)	−0.21 (0.23)
Action prediction	0.27 (0.30)	0.18 (0.21)
Math courses: Above Calculus II[a]	0.32 (0.30)	−0.05 (0.21)
Spatial	0.14 (0.06)*	−0.08 (0.04)*

[a]Taking math courses that are not beyond Calculus II is the reference group.
*$p < .05$; **$p < .01$.

action, or their combination. We first examined the mediating role of gesture. If gesture production were a *mediator*, we would expect increased gesture production in the action prediction, directed action, or Combined conditions. However, the regression results (see Table 5) show that neither condition led to a significant increase in dynamic or nondynamic depictive gesture production compared to Control condition. We found that spatial reasoning ability was positively associated with the production of dynamic depictive gestures ($OR = 1.14$, $p = .016$) and negatively associated with the production of nondynamic depictive gestures ($OR = .91$, $p = .030$). The lack of significant effects suggest that gesture is not a mediator for insight or proof. Therefore, further mediation analysis is unnecessary.

We then fit the models where gesture serves as a *moderator*. Based on Walkington et al. (2022), we focused on the role of gestural replays—gestures that re-enacted the prior predicted and or directed actions. Separate models were fitted for each type of gestural replay (overall, dynamic gestural, and nondynamic), using *avg_comparions* function to estimate the average marginal effects of each gestural replay type on insight and proof performance. Results indicate that when participants produced *dynamic gestural replays*, making action predictions or performing directed actions was significantly associated with insight and proof production (see Table 6).

For insight, making dynamic gestural replays in action prediction condition was associated with an 18.95% increase in odds ratios for generating correct insight ($p = .001$), compared to those who did not make dynamic gestural replays. Neither the directed action condition (6.61% *increase* in odds ratios; $p = .338$) nor the Combined condition (2.06% *decrease* in odds ratios; $p = .770$) showed a significant benefit for insight when making dynamic gestural replays of earlier actions.

Table 6
Results of the moderating effect of dynamic gestural replay on insight and proof (RQ4: Moderation analysis of gesture)

	Directed action		Action prediction		Combined	
	Insight	Proof	Insight	Proof	Insight	Proof
OR	1.07	1.19	1.20	1.48	0.98	1.09
Diff. (%)	6.61	17.66	18.95	39.17	−2.06	9.31
Std. error	0.069	0.067	0.058	0.066	0.070	0.072
p-value	.338	.008**	.001***	<.001***	.770	.193

Abbreviations: Diff., difference in OR compared to those who did not make dynamic gestural replays; OR, odds ratio.
**$p < .01$, ***$p < .001$.

For proof, making dynamic gestural replays had a significant effect on constructing a valid mathematically proof for both action prediction (39.17% *increase* in odds ratios; $p < .001$) and directed action (17.66% *increase* in odds ratios; $p = .008$) conditions, but not the Combined condition ($p = .193$).

Furthermore, when we compared the action prediction and directed action conditions among all participants who engaged in dynamic gestural replay, we found that participants in action prediction condition were 3.38 times more likely to generate a mathematically valid proof compared to those in directed action condition ($OR = 3.38$, $p = .002$), indicating a stronger positive effect of gestural replay in the action prediction condition.

Thus, the moderating role of gestural replays, while beneficial overall, is even stronger when *predicting* actions than actually *performing* actions, providing a novel contribution for how embodied processes influence cognition. These results demonstrate generally beneficial effects of gestural replays for those generating action predictions for both proof and insight, and those performing directed actions for just proof, but not when they were performed together (Combined). This finding supports the *gestural replay as a moderator hypothesis (H4-2)* and provides support for Predictions 4-2a and 2b.

Table 7 summarizes the findings and evidence supporting or countering each prediction in relation to each question.

5.5. *Post hoc analyses: Effects of action predictions on operational speech production*

The results above show that prompting people to predict task-relevant actions was a significant contributor to making correct mathematical insights and formulating valid mathematical proofs. Of the two, proof is by far the more critical academic skill (National Council of Teachers of Mathematics, 2000; Rav, 1999; Stylianides, 2007). Our data reveal that proof performance saw broad benefits from making action predictions. While this set of empirical findings is consistent with the predictions derived from the theorized reciprocity described by ACT, this begs a further question: What might these data reveal about *how* action predictions are beneficial for constructing mathematically valid proofs? To address this emergent question, we conducted a set of post hoc analyses. To narrow this investigation, post hoc analyses

Table 7
Summary of evidence supporting or countering each prediction for each research question

Hypothesis			Support for predicted differences on dependent variables ($p < .05$)	
H1	*Action prediction hypothesis (main effect)*: Action predictions enhance		Insight m	Proof ✓
H2	*Directed action hypothesis (main effect)*: Directed actions enhance		Insight ×	Proof ×
H3	*Overlapping processes hypothesis (interaction effect)*:	Combined = Action prediction or directed action[a]	Insight ✓	Proof ✓
		Combined > Control[b]	Insight m	Proof ×
H4	1. *Gesture as a mediator hypothesis*: Action prediction, directed action, or combined increases		Dynamic gesture ×	Nondynamic gesture ×
	2. *Gestural replay as a moderator hypothesis*:	Action predictions enhance	Insight ✓	Proof ✓
		Directed actions enhance	Insight ×	Proof ✓

Note. ✓: Supports the hypothesis; m: Marginal support; ×: Does not support the hypothesis.
[a]The Combined group does not yield additional benefits compared to action prediction or directed action alone.
[b]The Combined group outperforms the Control group.

examined the relative contributions of each of the three essential criteria used for assessing mathematically valid proof production (Harel & Sowder, 2007): generalization, operational thinking, and logic inference.

Logistic regression results are shown in Table 8. Model 2 shows that even when controlling for spatial ability score ($OR = 1.16$, $p = .006$), participants who were prompted to make

Table 8
Results of estimate (and standard error) for generating generality, operational thinking, logic inference: Comparing action prediction, directed action, combined groups to control group (post hoc analysis)

	Model 1: Generality	Model 2: Operational thinking	Model 3: Logic inference
Random effects: (Std. Dev.)			
Participant ID	0.68	0.60	0.60
Conjecture	0.82	2.03	0.62
Fixed effects:			
(Intercept)	−0.90 (0.67)	−3.13 (0.10)**	−0.62 (0.61)
Control	(ref.)	(ref.)	(ref.)
Directed action	−0.26 (0.27)	0.50 (0.30)	−0.31 (0.26)
Combined	0.32 (0.28)	0.75 (0.30)*	0.42 (0.27)
Action prediction	−0.08 (0.26)	1.02 (0.29)***	0.19 (0.25)
Math courses: Above Calculus II[a]	1.05 (0.27)***	0.06 (0.28)	1.10 (0.27)***
Spatial	0.07 (0.05)	0.15 (0.05)**	0.07 (0.05)

[a]Taking math courses that are not beyond Calculus II is the reference group.
$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

action predictions with directed actions (i.e., Combined; $OR = 2.13$, $p = .012$) and those prompted to make action predictions without directed actions ($OR = 2.78$, $p < .001$) significantly increased the relative odds of verbalizing operational thoughts. This was instantiated in speech that described a progressive sequence of goal-directed actions performed on imagined mathematical objects (Harel & Sowder, 2007). Assignment to the action prediction conditions was not statistically related to students' generalizations and logical inferences. These results suggest that making action predictions, with or without directed actions, facilitates geometric proof by encouraging the production of simulated actions that execute goal-directed operations (such as object manipulations and transformations) on imagined mathematical objects. This finding, coupled with the important role of gestural replays in proof production (H4-2), suggests that action prediction promotes the mental simulation of geometric transformations that engage both verbal and nonverbal processes. We explore this idea further in the discussion section.

6. Discussion

Our actions, whether real or imagined, can affect our thinking. Body-based interventions that are emblematic and enactive of geometric concepts offer new ways for learners to access information beyond the abstract formalisms, such as two-column proofs, typically employed in educational settings (Abrahamson et al., 2020). This study investigated the role of actions and action predictions in mathematical reasoning by prompting students to predict future bodily actions that enact properties and transformations of imagined geometric objects. The purpose of this intervention was to leverage the affordances of a motor control system that anticipates plausible outcomes of planned actions (Wolpert et al., 2003) by generating feedforward (proactive) signals and monitoring the feedback (reactive) of those signals on one's body and the world. The ACT model theorizes that the enactment and monitoring of actual and imagined goal-directed actions activate feedforward and feedback processes can induce task-relevant cognitive states that facilitate task performance through the mechanism of action-cognition transduction (Nathan, 2017). By predicting future bodily actions, participants' production of mathematically valid geometry proofs and mathematical insights was significantly improved. The ACT model further posits that the same mechanism applies when people replay prior actions and action simulations during subsequent explanations, thus offering an embodied account of the moderating role of explanatory gestures to support cognitive processes (Beilock & Goldin-Meadow, 2010; Walkington et al., 2022). Further investigation revealed that making action predictions specifically induced greater amounts of operational speech, one of the necessary criteria for determining mathematically valid deductive proofs.

After a brief summary of the main findings and novel contributions of this study, we explore their implications and discuss the limitations of the current investigation along with suggestions for future research.

6.1. Summary of findings

In addressing RQ1, we found that prompting participants to predict and perform action predictions can benefit the reasoning involved in an advanced topic of mathematical cognition, that of proof production for geometry. Our examination of RQ2 showed that participants' mimicry of directed actions did not directly lead to superior mathematical insights or proofs. Investigating RQ3 revealed that participants who performed directed actions and made action predictions did not perform significantly better than those who only performed directed actions or made action predictions. This finding favors the *overlapping processes hypothesis*, the position that some of the same processes that incur benefits to cognition through action on physical objects are also engaged when *predicting* actions operating on imaginary mathematical objects. In pursuit of RQ4, we found, contrary to our hypothesis (H4-1), that none of the conditions led to more gesture production. This finding suggests that gesture by itself is not a mediator of the observed cognitive behaviors that followed from action predictions and directed actions. However, in support of the gestural replay as a moderator hypothesis (H4-2), the analysis showed that participants achieved superior insights and proofs when their explanations included dynamic gestural replays of prior actions, regardless of whether these were investigator-generated (as with directed actions) or actor-generated (as with action predictions). In those cases of explanatory gestural replays, action prediction participants significantly outperformed directed action. Furthermore, a post hoc analysis indicated a possible causal process by revealing that participants who were prompted to make action predictions (those in both the Combined and action prediction only conditions) were significantly more likely to generate operational speech during their explanations describing the mathematical operations they would perform on mathematical objects.

In sum, our investigation potentially offers two novel contributions to the understanding of human mathematical cognition. The first is identifying circumstances in which prompts to imagine performing actions situated within a mathematical context can confer substantial benefits for participants' mathematical reasoning. The second potential contribution lends empirical support to the claim derived from the ACT model (Nathan, 2017; 2021) that simulated actions performed on imaginary mathematical objects are likely to draw on some of the same processes as actually performing the corresponding physical actions, and therefore, incur some of the same cognitive benefits observed with action production (e.g., Nathan et al., 2022). We next explore interpretations of these findings.

6.2. *Possible interpretations and explanations of findings*

Examining our findings within the ACT framework leads us to consider two possible interpretations, as illustrated in Fig. 3. One account is that the prediction process engages mental processes that simulate transformations on imagined geometric objects, as revealed by participants' heightened operational speech and increased gestural replays of earlier actions. An alternative account is that action predictions directly engage operational thinking in ways that directed actions do not. This operational thinking, manifested by operational speech that describes actions on mathematical entities, elicits highly appropriate explanatory gestures that schematize the most relevant attributes of the task (e.g., Kita et al., 2017). These gestures are replayed as simulated actions (Hostetter & Alibali, 2008), which contributes to superior insights and mathematically valid proof production. In the first account, action prediction drives simulation processes that lead to greater operational speech and dynamic

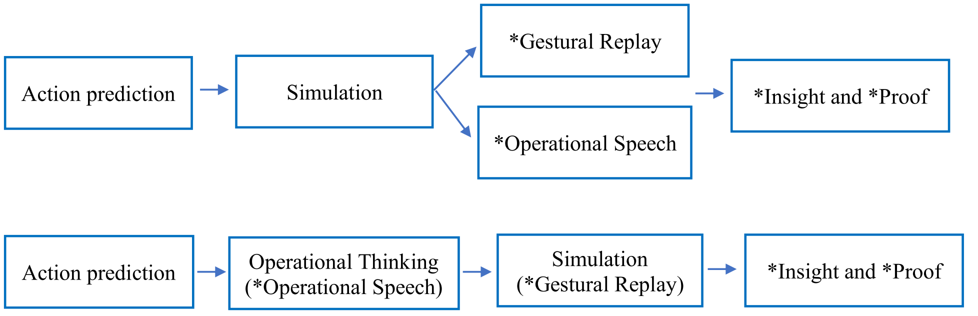

Fig. 3. Two accounts illustrating how action predictions might contribute to mathematical insight and proof performance.
Note. The * indicates that the event is observable and measurable.

gestural replays. In the second account, action prediction drives operational thinking, which then engages simulation processes that facilitate the re-enactment of dynamic gestural replays. Both accounts suggest that action prediction may facilitate mental simulations of actions on imagined mathematical objects that engage both verbal and nonverbal processes that contribute to enhanced proof production. Future research would be needed to disentangle the temporal differences of these two accounts.

Although conditions that prompted participants to make action predictions, with or without directed actions, led to improved proof performance, the odds ratios of generating valid proofs were significantly higher when participants made action predictions *without* performing directed actions. There are several possible reasons for this. One possibility is that making actor-generated action predictions promotes a form of multimodal self-explanation that helps bring subconscious ways of knowing into consciousness to promote better reasoning. Participants spontaneously self-explained what they enacted. Previous research using action-based interventions (e.g., Nathan et al., 2014; Thomas & Lleras, 2007) found that those who benefited from actions typically report being unaware of the actions' relevance to the reasoning task, but that explicit attention to their relevance further improved cognitive performance (Nathan et al., 2014; Nathan & Walkington, 2017). Several scholars (e.g., Aleven & Koedinger, 2002; Bielaczyc, Pirolli, & Brown, 1995; Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Heijltjes, Van Gog, Leppink, & Paas, 2015; McNamara, 2004) have suggested that the benefits of self-explanation may arise by raising conscious awareness of otherwise unconscious and automated processes. Thus, we may speculate at this time that inviting participants to predict future actions may facilitate proof production by engaging participants in multimodal (gestural and verbal) self-explanations that help integrate task-relevant unconscious processes, such as perception, and conscious processes, such as operational thinking, in service of their mathematical reasoning.

We have been asked to speculate how it might be that conscious processes operating at "higher levels" of cognition might shape "lower-level" unconscious processes. Although we recognize it is commonplace to use the metaphorical language of "lower" and "higher" to denote apparent hierarchical relationships in cognition, we are cautious about applying these relationships within complex systems. As a counterpoint to this framing, Nathan (2021) suggests adopting the language of supervenience relations to describe the mutual effects that processes operating at different scales have on one another. In the current context, we propose that conscious cognitive processes such as operational thinking may supervene on the associated unconscious sensorimotor processes, emerging from their behaviors, but also be supervenient to, and place constraints on the functioning of those sensorimotor processes. For example, directing one's attention through self-directed explanations (Aleven & Koedinger, 2002) and the schematization role of gestures (Kita et al., 2017) enables conscious cognitive processes to influence what information gets activated and is, therefore, most likely to get processed unconsciously. Forming cognitively relevant gestures of geometric forms and relationships may also constitute a form of priming that also can influence unconscious perceptual and motoric processes that contribute to one's mathematical insight and reasoning. Thus, the effects across levels can be mutual.

Another possibility for how action predictions may incur greater benefits for mathematical reasoning than performing cognitively relevant directed actions is that *actor-generated* actions that correspond to participants' predictions may be more meaningful and potent than mimicking investigator-generated directed actions. Enactivist theoretical accounts (e.g., Abrahamson & Sánchez-García, 2016; O'Regan, & Noë, 2001) may play a role in explaining this phenomenon.

A **novel contribution** from our study is that predicting ways of completing a partial directed action sequence that one had performed (i.e., the Combined condition) did not show additional cognitive benefits when compared to *predicting* future actions without any cues about the nature of the actions (action prediction) or *performing* the entire directed action sequence. This result suggests that simulated actions elicited by action predictions may activate some of the same overlapping processes as actually performing cognitively relevant actions, leading to redundant, rather than additive, contributions. One alternative explanation for these findings is that performing both action predictions and directed actions in the Combined condition may be more complicated than performing either alone, and therefore, fail to yield any benefit, even if these are not dependent on overlapping processes. However, this explanation seems implausible. This is because the Combined group actually did better than action prediction group or directed action group, even though that difference is not significant (see Section 5.3.1). Furthermore, participants in the Combined group prompted to make action predictions with directed actions and those prompted to make action predictions without directed actions had significantly greater relative odds of verbalizing operational thoughts. There does not seem to be any evidence that combining action predictions and directed actions created any relevant complications that burdened participant performance. Thus, it seems unlikely that any added complication of performing both plausibly accounts for these findings. Still, future studies may be valuable to address this.

6.3. Limitations and future directions

These findings should be interpreted in light of several limitations of the current design. Participants were observed within a remotely conducted experimental setting on a limited set of tasks, those common to secondary and post-secondary geometry courses, with an emphasis on planar, two-dimensional objects and relations. We must be cautious not to overgeneralize these findings to vastly dissimilar domains and contexts until these effects can be demonstrated more widely. Action prediction prompting also followed a specific protocol, prompted by visual stimuli for predicting the movements of a digital avatar (Fig. 1). Other ways of eliciting action predictions might lead to different behaviors and outcomes. Finally, there is some ambiguity (illustrated in Fig. 3) as to whether action predictions directly generate mental simulations that lead to heightened operational speech and gestural replays, or if the intervention activated operational speech and gestural replays directly, contributing to the kinds of inferential reasoning that supports generalizable proofs. Having identified these, we can consider ways new process-level data collection activities in future work can help to disentangle these competing interpretations.

6.4. Conclusions

This study is one of very few that used a randomized controlled experimental design to explore the role of action predictions on cognition. We chose an area of advanced mathematical reasoning—geometry proof performance—to further test claims of the potential of embodied learning. The evidence supports novel findings, derived from the ACT model of embodied cognition, that prompting people to engage in task-relevant action predictions benefits their mathematical reasoning, as demonstrated by superior mathematical insights and mathematically valid proof production, and that these simulated actions performed on imaginary mathematical objects are likely to draw on some of the same processes as actually performing the corresponding physical actions. We also found that the moderating role of gestural replays, while overall helpful, benefits action predictions even more strongly than actually performing actions, for outcome measures of both mathematically valid proof and mathematical insight.

The findings of this theoretically motivated study offer potentially valuable insights for cognitive scientists, instructional designers, and educational practitioners. For cognitive scientists, our findings contribute to the growing body of evidence that cognition is embodied, even in an area of advanced mathematical reasoning typically regarded as "abstract." By engaging embodiment through action prediction, we extend this account by demonstrating the potential that mental simulation of sensorimotor processes may play in embodied cognition (e.g., Barsalou, 2009), seemingly by activating some of the same cognitive resources attributable to actions performed on physical objects. We offer transduction as a possible mechanism for understanding how action and cognition influence one another. Thus, these findings provide some provisional support to causal claims of the role that actual and imagined body movements play in cognitive processes (Shapiro, 2019) and in supporting claims about the reciprocity between intellectual and body-based processes, as theorized in ACT (Nathan, 2017). In addition to our outcome measures, process-level data suggests that prompting people to generate action predictions may enhance their mathematical cognition by eliciting greater operational thinking and richer multimodal explanations that help integrate nonverbal and verbal task-relevant processes.

For instructional designers, this work offers ways to support the development of students' skills for mathematical proof and justification that may complement current instructional methods. Traditional approaches to secondary and post-secondary geometry instruction, such as the two-column proof, focus on constructing arguments based on logical deduction that are constructed through a sequence of static assertions about geometric properties along with supporting formal justifications. In this study, we observed that students also gain insight and produce mathematically valid proofs when they explore the veracity of mathematical conjectures using dynamic transformations of the mathematical objects under investigation. While transformational proof practices can be supported using physical manipulatives, digital geometry systems, and augmented, virtual, and extended reality systems (Walkington et al., 2024), the findings reported here demonstrate that students can also achieve this type of reasoning by engaging body-based resources to simulate these geometric transformations. Thus, there is educational value in designing for embodied interactions of both real and imagined mathematical objects.

Related to our suggestions for instructional designs, educational practitioners who interact directly with students may witness some of the benefits of embodied simulations by fostering a classroom climate that encourages movement, prediction, and mathematical imagination. In addition to the potential enhancements in the quality of students' reasoning, teachers attending to students' explanatory gestures may obtain valuable information about students' nonverbal ways of knowing (Alibali & Nathan, 2018) that can benefit teachers' formative assessment practices to inform their real-time instruction and summative assessment of students' understanding of the universal properties of shape and space. Overall, these findings may offer cognitive scientists, instructional designers, and educational practitioners novel pathways for understanding, fostering, and assessing mathematical reasoning by attending to embodied learning processes.

References

Abrahamson, D., Nathan, M. J., Williams-Pierce, C., Walkington, C., Ottmar, E. R., Soto, H., & Alibali, M. W. (2020). The future of embodied design for mathematics teaching and learning. *Frontiers in Education*, *5*, 147.

Abrahamson, D., & Sánchez-García, R. (2016). Learning is moving in new ways: The ecological dynamics of mathematics education. *Journal of the Learning Sciences*, *25*(2), 203–239.

Adams, A. M., Glenberg, A. M., & Restrepo, M. A. (2018). Moved by reading in a Spanish-speaking, dual language learner population. *Language, Speech, and Hearing Services in Schools*, *49*(3), 582–594.

Aleven, V. A., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science*, *26*(2), 147–179.

Alibali, M. W., & DiRusso, A. A. (1999). The function of gesture in learning to count: More than keeping track. *Cognitive Development*, *14*(1), 37–56.

Alibali, M. W., & Nathan, M. J. (2018). Embodied cognition in learning and teaching: Action, observation, and imagination. In F. Fischer, S. Goldman, C. Hmelo-Silver, & P. Riemann (Eds.), *International handbook of the learning sciences* (pp. 75–85). New York: Routledge/Taylor & Francis. https://doi.org/10.4324/9781315617572-8

Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language*, *44*(2), 169–188.

Ashe, J., Lungu, O. V., Basford, A. T., & Lu, X. (2006). Cortical control of motor sequences. *Current Opinion in Neurobiology*, *16*(2), 213–221.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*(1), 617–645.

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1281–1289.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7. Retrieved from http://CRAN.R-project.org/package=lme4

Beilock, S. L., & Goldin-Meadow, S. (2010). Gesture changes thought by grounding it in action. *Psychological Science*, *21*(11), 1605–1610.

Bielaczyc, K., Pirolli, P. L., & Brown, A. L. (1995). Training in self-explanation and self-regulation strategies: Investigating the effects of knowledge acquisition activities on problem solving. *Cognition and Instruction*, *13*(2), 221–252.

Borreggine, K. L., & Kaschak, M. P. (2006). The action–sentence compatibility effect: It's all in the timing. *Cognitive Science*, *30*(6), 1097–1112.

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.

Brooks, N., & Goldin-Meadow, S. (2016). Moving to learn: How guiding the hands can set the stage for learning. *Cognitive Science*, *40*(7), 1831–1849.

Butz, M. V., & Pezzulo, G. (2008). Benefits of anticipations in cognitive agents. In *The challenge of anticipation: A unifying framework for the analysis and design of artificial cognitive systems* (pp. 45–62). Berlin, Heidelberg: Springer Berlin Heidelberg.

Bub, D. N., & Masson, M. E. J. (2010). On the nature of hand-action representations evoked during written sentence comprehension. *Cognition*, *116*, 394–408.

Byrd, C. E., McNeil, N. M., D'Mello, S. K., & Cook, S. W. (2014). Gesturing may not always make learning last. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 1982–1987). Austin, TX: Cognitive Science Society.

Cherdieu, M., Palombi, O., Gerber, S., Troccaz, J., & Rochet-Capellan, A. (2017). Make gestures to learn: Reproducing gestures improves the learning of anatomical knowledge more than just seeing gestures. *Frontiers in Psychology*, *8*, 1689.

Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*(2), 145–182.

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.

Chu, M., & Kita, S. (2011). The nature of gestures' beneficial role in spatial problem solving. *Journal of Experimental Psychology: General*, *140*(1), 102.

Chu, M., & Kita, S. (2016). Co-thought and co-speech gestures are generated by the same action generation process. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(2), 257–270.

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition*, *23*(1), 43–71.

Clark, A. (2012). Dreaming the whole cat: Generative models, predictive processing, and the enactivist conception of perceptual experience. *Mind*, *121*(483), 753–771.

Clark, A. (2015). *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46.

Cook, S. W. (2018). Enhancing learning with hand gestures: Potential mechanisms. In K. D. Federmeier (Ed.), *Psychology of learning and motivation* (pp. 107–133). Amsterdam, NL: Academic Press. https://doi.org/10.1016/bs.plm.2018.10.001

Cook, S. W., & Goldin-Meadow, S. (2006). The role of gesture in learning: Do children use their hands to change their minds? *Journal of Cognition and Development*, *7*(2), 211–232.

Cook, S. W., & Tanenhaus, M. K. (2009). Embodied communication: Speakers' gestures affect listeners' actions. *Cognition*, *113*(1), 98–104.

Decety, J., & Ingvar, D. H. (1990). Brain structures participating in mental simulation of motor behavior: A neuropsychological interpretation. *Acta Psychologica*, *73*(1), 13–34.

Decety, J., & Grèzes, J. (2006). The power of simulation: imagining one's own and other's behavior. *Brain Research*, *1079*(1), 4–14.

Dreyfus, T. (1999). Why Johnny can't prove. *Educational Studies in Mathematics*, *38*(1), 85–109.

Eagan, B. R., Rogers, B., Serlin, R., Ruis, A. R., Arastoopour Irgens, G., & Shaffer, D. W. (2017). Can we rely on IRR? Testing the assumptions of inter-rater reliability. In *International Conference on Computer Supported Collaborative Learning*.

Enyedy, N., Danish, J. A., & DeLiema, D. (2015). Constructing liminal blends in a collaborative augmented-reality learning environment. *International Journal of Computer-Supported Collaborative Learning*, *10*(1), 7–34.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160.

Garcia, N., & Infante, N. E. (2012). Gestures as facilitators to proficient mental modelers. In L. R. Van Zoest, J.-J. Lo, & J. L. Kratky (Eds.), *Proceedings of the 34th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 289–295).

Gibson, J. J. (1977). *The Theory of Affordances: The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, *20*(1), 1–19.

Glenberg, A. M., Brown, M., & Levin, J. R. (2007). Enhancing comprehension in small reading groups using a manipulation strategy. *Contemporary Educational Psychology*, *32*(3), 389–399.

Glenberg, A. M., & Gallese, V. (2012). Action-based language: A theory of language acquisition, comprehension, and production. *Cortex*, *48*(7), 905–922.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*(3), 558–565.

Glenberg, A. M., Gutierrez, T., Levin, J. R., Japuntich, S., & Kaschak, M. P. (2004). Activity and imagined activity can enhance young children's reading comprehension. *Journal of Educational Psychology*, *96*(3), 424.

Goldin-Meadow, S., Alibali, M. W., & Church, R. B. (1993). Transitions in concept acquisition: Using the hand to read the mind. *Psychological Review*, *100*(2), 279.

Goldin-Meadow, S., & Beilock, S. L. (2010). Action's influence on thought: The case of gesture. *Perspectives on Psychological Science*, *5*(6), 664–674.

Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009). Gesturing gives children new ideas about math. *Psychological Science*, *20*(3), 267–272.

Grèzes, J., Frith, C. D., & Passingham, R. E. (2004). Inferring false beliefs from the actions of oneself and others: an fMRI study. *Neuroimage*, *21*(2), 744–750.

Grush, R. (2004). The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Sciences*, *27*(3), 377–396.

Harel, G., & Sowder, L. (1998). Students' proof schemes. In E. Dubinsky, A. Schoenfeld, & J. Kaput (Eds.), *Research on collegiate mathematics education* (pp. 234–283, Vol. III). Providence, RI: American Mathematical Society.

Harel, G., & Sowder, L. (2007). Toward comprehensive perspectives on the learning and teaching of proof. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*. Reston, VA: National Council of Teachers of Mathematics. 805–842.

Haruno, M., Wolpert, D. M., & Kawato, M. (2003). Hierarchical MOSAIC for movement generation. In *International Congress Series* (pp. 575–590, Vol. 1250). Elsevier.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*(2), 301–307.

Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education*, *31*(4), 396–428.

Heijltjes, A., Van Gog, T., Leppink, J., & Paas, F. (2015). Unraveling the effects of critical thinking instructions, practice, and self-explanation on students' reasoning performance. *Instructional Science*, *43*, 487–506.

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin*, *137*(2), 297.

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.

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

Hostetter, A. B., & Alibali, M. W. (2019). Gesture as simulated action: Revisiting the framework. *Psychonomic Bulletin & Review*, *26*(3), 721–752.

Hostetter, A. B., & Alibali, M. W. (2010). Language, gesture, action! A test of the Gesture as Simulated Action framework. *Journal of Memory and Language*, *63*(2), 245–257.

Hostetter, A. B., & Skirving, C. J. (2011). The effect of visual vs. verbal stimuli on gesture production. *Journal of Nonverbal Behavior*, *35*, 205–223.

Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *Neuroimage*, *14*(1), S103–S109.

Jones, K. (2000). Providing a foundation for deductive reasoning: Students' interpretations when using dynamic geometry software and their evolving mathematical explanations. *Educational Studies in Mathematics*, *44*(1), 55–85.

Kaschak, M., Zwaan, R. A., Glenberg, A., Morey, R. D., Ibanez, A., Gianelli, C., & Haaf, J. M. (2018). Action-sentence compatibility effect (ACE) pre-registered replication. Retrieved from osf. io/ynbwu.

Kita, S., Alibali, M. W., & Chu, M. (2017). How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological Review*, *124*(3), 245.

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, *48*(1), 16–32.

Knuth, E. J. (2002). Secondary school mathematics teachers' conceptions of proof. *Journal for Research in Mathematics Education*, *33*(5), 379–405.

Knuth, E., Choppin, J., & Bieda, K. (2009). Middle school students' production of mathematical justifications. In D. Stylianou, M. Blanton, & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K–16 perspective* (pp. 153–170). New York: Routledge.

Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by embodiment: Six precepts for research on embodied learning and mixed reality. *Educational Researcher*, *42*(8), 445–452.

Linnenbrink-Garcia, L., Durik, A., Conley, A., Barron, K., Tauer, J., Karabenick, S., & Harackiewicz, J. (2010). Measuring situational interest in academic domains. *Educational Psychological Measurement*, *70*, 647–671.

Malanchini, M., Rimfeld, K., Shakeshaft, N. G., McMillan, A., Schofield, K. L., Rodic, M., Rossi, V., Kovas, Y., Dale, P. S., Tucker-Drob, E. M., & Plomin, R. (2020). Evidence for a unitary structure of spatial cognition beyond general intelligence. *NPJ Science of Learning*, *5*(1), 1–13.

Marghetis, T., Edwards, L. D., & Núñez, R. (2014). More than mere handwaving: Gesture and embodiment in expert mathematical proof. In L. D. Edwards, F. Ferrara, & D. Moore-Russo (Eds.), *Emerging perspectives on gesture and embodiment in mathematics* (pp. 227–246). Charlotte, NC: Information Age Publishing.

Masters, J. (2010). *Diagnostic geometry assessment project: Item characteristics*. Chestnut Hill, MA: Technology and Assessment Study Collaborative, Lynch School of Education, Boston College.

McNamara, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes*, *38*(1), 1–30.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.

McNeill, D. (2008). *Gesture and thought*. University of Chicago Press.

Morey, R. D., Kaschak, M. P., Díez-Álamo, A. M., Glenberg, A. M., Zwaan, R. A., Lakens, D., Ibáñez, A., Garcia, A., Gianelli, C., Jones, J. L., Madden, J., Alifano, F., Bergen, B., Bloxsom, N. G., Bub, D. N., Cai, Z. G., Chartier, C. R., Chatterjee, A., Conwell, E., Cook, S. W., Davis, J. D., Evers, E., Girard, S., Harter, D., Hartung, F., Herrera, E., Huettig, F., Humphries, S., Juanchich, M., Kühne, K., Lu, S., Lynes, T., Masson, M.

E., Ostarek, M., Pessers, S., Reglin, R., Steegen, S., Thiessen, E. D., Thomas, L. E., Trott, S., Vandekerckhove, J., Vanpaemel, W., Vlachou, M., Williams, K., & Ziv-Crispel, N. (2022). A pre-registered, multi-lab non-replication of the action-sentence compatibility effect (ACE). *Psychonomic Bulletin & Review*, *29*(2), 613–626.

Nathan, M. J. (2017). One function of gesture is to make new ideas. In R. B. Church, M. W. Alibali, & S. D. Kelly (Eds.), *Why gesture? How the hands function in speaking, thinking and communicating* (pp. 175–196). Amsterdam, The Netherlands: John Benjamins Publishing Company.

Nathan, M. J. (2021). *Foundations of embodied learning: A paradigm for education*. Routledge.

Nathan, M. J., & Martinez, C. V. (2015). Gesture as model enactment: The role of gesture in mental model construction and inference making when learning from text. *Learning: Research and Practice*, *1*(1), 4–37.

Nathan, M. J., Schenck, K. E., Vinsonhaler, R., Michaelis, J. E., Swart, M. I., & Walkington, C. (2021). Embodied geometric reasoning: Dynamic gestures during intuition, insight, and proof. *Journal of Educational Psychology*, *113*(5), 929.

Nathan, M. J., & Swart, M. I. (2021). Materialist epistemology lends design wings: Educational design as an embodied process. *Educational Technology Research and Development*, *69*, 1925–1954.

Nathan, M. J., & Walkington, C. (2017). Grounded and embodied mathematical cognition: Promoting mathematical insight and proof using action and language. *Cognitive Research: Principles and Implications*, *2*(1), 1–20.

Nathan, M. J., Walkington, C., Boncoddo, R., Pier, E., Williams, C. C., & Alibali, M. W. (2014). Actions speak louder with words: The roles of action and pedagogical language for grounding mathematical proof. *Learning and Instruction*, *33*, 182–193.

Nathan, M. J., Walkington, C., & Swart, M. I. (2022). Designs for grounded and embodied mathematical learning. In *Proceedings of the 16th International Conference of the Learning Sciences-ICLS 2022* (pp. 179–186). International Society of the Learning Sciences.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

NCH Software. (2020). Version 6.16. Retrieved from https://www.nchsoftware.com/capture/index.html.

Neisser, U. (1976/1987). *Cognition and reality*. New York: W. H. Freeman.

Novack, M. A., Congdon, E. L., Hemani-Lopez, N., & Goldin-Meadow, S. (2014). From action to abstraction: Using the hands to learn math. *Psychological Science*, *25*(4), 903–910.

O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, *24*(5), 939–973.

Pezzulo, G. (2008). Coordinating with the future: The anticipatory nature of representation. *Minds and Machines*, *18*(2), 179–225.

Pier, E. L., Walkington, C., Clinton, V., Boncoddo, R., Williams-Pierce, C., Alibali, M. W., & Nathan, M. J. (2019). Embodied truths: How dynamic gestures and speech contribute to mathematical proof practices. *Contemporary Educational Psychology*, *58*, 44–57.

Pouw, W., Wassenburg, S. I., Hostetter, A. B., De Koning, B. B., & Paas, F. (2020). Does gesture strengthen sensorimotor knowledge of objects? The case of the size-weight illusion. *Psychological Research*, *84*, 966–980.

Ramful, A., Lowrie, T., & Logan, T. (2017). Measurement of spatial ability: Construction and validation of the spatial reasoning instrument for middle school students. *Journal of Psychoeducational Assessment*, *35*(7), 709–727.

Rav, Y. (1999). Why do we prove theorems? *Philosophia Mathematica*, *7*(1), 5–41.

Remillard, G. (2003). Pure perceptual-based sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(4), 581.

Roth, W. M. (2001). Gestures: Their role in teaching and learning. *Review of Educational Research*, *71*(3), 365–392.

Ruby, P., & Decety, J. (2004). How would you feel versus how do you think she would feel? A neuroimaging study of perspective-taking with social emotions. *Journal of Cognitive Neuroscience*, *16*(6), 988–999.

Schenck, K. E., Walkington, C., & Nathan, M. J. (2022). Groups that move together, prove together: Collaborative gestures and gesture attitudes among teachers performing embodied geometry. In S. Macrine & J. Fugate (Eds.), *Movement matters: How embodied cognition informs teaching and learning* (pp. 131–145). MIT Press.

Schubotz, R. I., & von Cramon, D. Y. (2001). Functional organization of the lateral premotor cortex: FMRI reveals different regions activated by anticipation of object properties, location and speed. *Cognitive Brain Research*, *11*(1), 97–112.

Shapiro, L. (2019). *Embodied cognition*. Routledge.

Smith, C. P., King, B., & Hoyte, J. (2014). Learning angles through movement: Critical actions for developing understanding in an embodied activity. *Journal of Mathematical Behavior*, *36*, 95–108.

Stevanoni, E., & Salmon, K. (2005). Giving memory a hand: Instructing children to gesture enhances their event recall. *Journal of Nonverbal Behavior*, *29*(4), 217–233.

Stylianides, A. J. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education*, *38*(3), 289–321.

Thomas, L. E., & Lleras, A. (2007). Moving eyes and moving thought: On the spatial compatibility between eye movements and cognition. *Psychonomic Bulletin & Review*, *14*(4), 663–668.

Thomas, L. E., & Lleras, A. (2009). Swinging into thought: Directed movement guides insight in problem solving. *Psychonomic Bulletin & Review*, *16*(4), 719–723.

Walkington, C., Chelule, G., Woods, D., & Nathan, M. J. (2019). Collaborative gesture as a case of extended mathematical cognition. *Journal of Mathematical Behavior*, *55*, 100683.

Walkington, C., Nathan, M. J., Wang, M., & Schenck, K. (2022). The effect of cognitive relevance of directed actions on mathematical reasoning. *Cognitive Science*, *46*(9), e13180.

Walkington, C., Nathan, M. J., Washington, J., Hunnicutt, J., Darwin, T., Daughrity, L., & Schenck, K. (2024). Comparing learning geometry using a tablet to head-mounted display augmented reality: How and when dimensionality matters. *Education and Information Technologies*, 1–30. https://doi.org/10.1007/s10639-024-13008-z

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*, 625–636.

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology*, *11*(18), R729–R732.

Wolpert, D. M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, *11*(7–8), 1317–1329.

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 593–602.

Woods, D., & Fassnacht, C. (2012). Transana (Version 2.52) [Computer software]. Retrieved from http://transana.com

Wynn, K. (1990). Children's understanding of counting. *Cognition*, *36*(2), 155–193.

Yee, E., Chrysikou, E. G., Hoffman, E., & Thompson-Schill, S. L. (2013). Manual experience shapes object representations. *Psychological Science*, *24*(6), 909–919.

Yeo, A., Ledesma, I., Nathan, M. J., Alibali, M. W., & Church, R. B. (2017). Teachers' gestures and students' learning: Sometimes "hands off" is better. *Cognitive Research: Principles and Implications*, *2*, 1–11.

Zhang, I., Givvin, K. B., Sipple, J. M., Son, J. Y., & Stigler, J. W. (2021). Instructed hand movements affect students' learning of an abstract concept from video. *Cognitive Science*, *45*(2), e12940.

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General*, *135*, 1–11.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix A
Appendix B
Appendix C