Learning from an avatar video instructor

The role of gesture mimicry

Nicholas A. Vest[1], Emily R. Fyfe[2], Mitchell J. Nathan[1] and Martha W. Alibali[1]
[1] University of Wisconsin – Madison | [2] Indiana University

Teachers often produce gestures, and, in some cases, students mimic their teachers' gestures and adopt them into their own repertoires. However, little research has explored the role of gesture mimicry in technology-based learning contexts. In this research, we examined variations in the rate and form of students' gestures when learning from a computer-animated pedagogical avatar. Twenty-four middle school students received a lesson on polynomial multiplication from a gesturing avatar video instructor. After the lesson, students were asked to provide an explanation of what they learned. Students varied in their gesture rates, and some students produced gestures that were similar in form to the avatar's gestures. Students who produced gestures that aligned with the teacher's gestures scored higher than those who did not produce such gestures. These results suggest that middle school students' gestures play a key role when learning a mathematics lesson from an avatar instructor.

Keywords: gesture mimicry, avatar instructor, algebra, instructional technology

The actions we produce shape our learning from infancy to adulthood (Kontra, Goldin-Meadow, & Beilock, 2012). These actions can vary from actions on physical objects, such as reaching and picking up objects, to abstract movements, such as gestures (Libertus & Needham, 2010; Goldin-Meadow & Beilock, 2010). Gestures can be broadly defined as the movements people produce with their hands and arms while speaking. They are often spontaneous and can express information that is not expressed in the accompanying speech (McNeill, 1992). Importantly, both producing gestures and seeing others' gestures can have a profound effect on thinking and learning (Goldin-Meadow, 2005; Goldin-Meadow & Alibali, 2013; Hostetter, 2011). The purpose of this study is to examine variations

https://doi.org/10.1075/gest.18019.ves
Gesture 19:1 (2020), pp. 128–155. issn 1568-1475 | e-issn 1569-9773

in middle school students' gestures as they learn algebra from an avatar video instructor. Given advances in technology and the development of computer-animated instructors, we hope to shed light on the role of gestures within the learner-instructor dyad in this novel learning context.

Research suggests that learners' gestures are a critical factor in the learning environment. Indeed, variations in children's and adults' spontaneous gestures are often related to their problem-solving strategies as well as their likelihood of learning from instruction (e.g., Alibali, Spencer, Knox, & Kita, 2011; Church & Goldin-Meadow, 1986). For example, Church and Goldin-Meadow (1986) asked elementary school children to explain six conservation judgments (e.g., "Do the glasses have the same amount of water? How can you tell?"; Church & Goldin-Meadow, 1986). Children who often produced gestures that expressed task-relevant information that they did not express in speech (e.g., gesturing about width, but talking about height) were more likely to learn from a subsequent lesson than children who usually expressed similar information in gestures and speech. Instructing children to produce gesture can also improve learning. For example, Broaders et al. (2007) told children to gesture or not to gesture while explaining novel math problems. Children who were told to gesture often expressed correct strategies in their gestures, and they learned more from a subsequent mathematics lesson than children who were told not to gesture.

Research indicates that seeing others use gestures can also be beneficial. Numerous studies indicate that students learn more when the teacher uses speech and gesture rather than speech alone (e.g., Berch, Singleton, & Perry, 1995; Singer & Goldin-Meadow, 2005). For example, Ping and Goldin-Meadow (2008) found that first-grade children's performance on conservation tasks improved more when a human tutor used speech and gestures rather than speech alone. Similar benefits occur when the human tutor is video-recorded (e.g., Church, Ayman-Nolley, & Mahootian, 2004; Cook, Duffy, & Fenn, 2013). For example, Valenzeno et al. (2003) asked preschoolers to view a speech-only or a speech-plus-gesture videotaped lesson on the concept of symmetry. Children who viewed the speech-plus-gesture lesson scored higher on the post-lesson assessments than children who viewed the speech-only lesson.

There are a variety of reasons why seeing others use gestures may be beneficial (see Goldin-Meadow & Alibali, 2013). One potential reason is the learner's adoption or mimicry of task-relevant gestures. Mimicry in communication is defined as the convergence of verbal or nonverbal behavior across speakers (Parrill & Kimbara, 2006), and *gesture mimicry* occurs when participants in an interaction produce gestures that are similar in form. For example, Cook and Goldin-Meadow (2006) provided children instruction on solving mathematical equations (e.g., $3+4+5=3+\_\_$) with and without gesture. Children who saw the correct strategy in the instructor's gesture were more likely to reproduce that strategy in their own gestures than children who were not exposed to the strategy in the instructor's gesture. In turn, those children who mimicked the strategy in gesture were more likely to succeed on a posttest than children who did not mimic the strategy in gesture. Thus, gesture mimicry can play a key role in the learning environment.

Prior work has focused on mimicry of human teachers' gestures – either from live instructional lessons or from video-recorded lessons. However, with technological advances, increasing amounts of educational material are presented in the absence of a human teacher (e.g., educational apps, online lessons with avatars, etc.). The goal of the present study is to investigate variations in students' gestures in a novel, technology-based learning setting with an avatar video instructor, and to examine whether variations in students' gestures are associated with performance after the lesson. The avatar we used is a computer-animated embodied pedagogical agent that allows for complete control of both verbal and nonverbal behaviors (i.e., face, lip, and body movements). Avatars have been shown to support learning by directly communicating content to learners (e.g., Baylor, 2003; Cook, Friedman, Duggan, Cui, & Popescu, 2017; Lester, Converse, Stone, Kahler, & Barlow, 1997). Further, gestures from an avatar appear to be beneficial for students' learning (e.g., Buisine & Martin, 2007; Cook et al., 2017). For example, undergraduates learned more from an avatar that gestured and moved on the screen relative to an avatar who remained static on the screen (Lusk & Atkinson, 2007). Similarly, elementary school students learned more from a gesturing avatar instructor than from a non-gesturing avatar that used identical speech, eye gaze, head position, and lip movements (Cook et al., 2017).

Despite increasing research in this area, little is known about *students*' gestures while learning from a gesturing avatar. For example, it remains unclear whether students will mimic an avatar instructor's gestures, and if so, whether variability in this type of gesture mimicry is related to learning. There are reasons to expect variability in students' gestures in this novel context. For example, some students may view the avatar as a standard social agent and engage in typical gesture behavior, including reproducing task-relevant gestures produced by the avatar. Indeed, avatars are malleable and often display sociable characteristics and multimodal behaviors that humans display during face-to-face interactions (Cassell, Sullivan, Prevost, & Churchill, 2000). Further, eye gaze data suggests that adults often regard animated agents as legitimate conversational partners (Louwerse, Graesser, McNamara, & Lu, 2009).

However, other students may be less likely to respond to the avatar naturally or to adopt the avatar's gestures. Avatars are sometimes perceived as "too" human-like, and they may enter the "uncanny valley," in which they elicit feelings of eeriness or revulsion (Mori, MacDorman, & Kageki, 2012). For example, Black and colleagues (2009) assessed how nine children (ages 4 to 6) interacted with a human versus an animated avatar. They found that the children spoke more slowly, responded more slowly, and used fewer gestures when speaking to the avatar, relative to the human. However, in this study, the speech and gestures produced by the human and avatar were not identical. Thus, it remains unclear whether these students were mimicking the avatar's interaction style (and thus gesturing less because the avatar gestured less) or were changing their interaction style because they perceived the avatar differently from the human.

In the current study, we examined variability in students' gestures as they learned from an avatar video instructor. From a practical standpoint, it is important to study avatar instructors as they are becoming increasingly popular for communicating general information (e.g., reading news, presenting tourist information; Noma, Zhao, & Badler, 2000) and for providing targeted instruction in electronic learning environments (e.g., Adamo-Villani, Wilbur, Eccarius, & Abe-Harris, 2009; Lester et al., 1997). From a theoretical standpoint, research on avatar instructors can yield insight into learners' interpretations of animated agents and the role of gesture in technology-mediated environments (e.g., Cassell et al., 2000). In this study, we used an avatar video instructor to provide an algebra lesson on polynomial multiplication (e.g., $(2x+5)(x+2)$) to middle school students. This study is part of a larger project that developed an avatar instructor for facilitating learning of target algebra concepts in middle school (Anasingaraju et al., 2016; Popescu et al., 2014; Yeo et al., 2018). Algebra was selected as the target domain because it functions as a gatekeeper to future educational opportunities (e.g., Moses & Cobb, 2001), and it is a focal point of content standards in mathematics education (NGACBP, 2010). The target topic, polynomial multiplication, is foundational to higher mathematics and is introduced in beginning algebra classes.

In the current study, students completed an initial baseline assessment, which allowed us to measure their prior knowledge of the target content and assess variability in students' spontaneous use of task-relevant gestures, prior to encountering the avatar instructor. Students then received a brief video lesson, during which the avatar instructor provided instruction in speech and gesture using an area-based representation and an equation-based representation (see Figure 1). After viewing these lessons, students were asked to explain a target problem. Thus, students had the opportunity to view the avatar's gestures and then to produce gestures themselves. This design allowed us to address two specific aims. Our first aim was to assess *how frequently* students gestured after viewing the avatar's lesson and to assess whether students *mimicked* the gestures produced by the avatar. The second aim was to examine whether variations in gesture frequency or mimicry related to performance after the lesson. We hypothesized that students would vary in how frequently they gestured and whether they mimicked the avatar's gestures. Further, given that students' production of task-relevant gestures is often related to learning (e.g., Church & Goldin-Meadow, 1986; Cook & Goldin-Meadow, 2006), we hypothesized that students who mimicked the avatar's gestures would perform better on a post-lesson assessment than students who did not mimic the avatar's gestures.

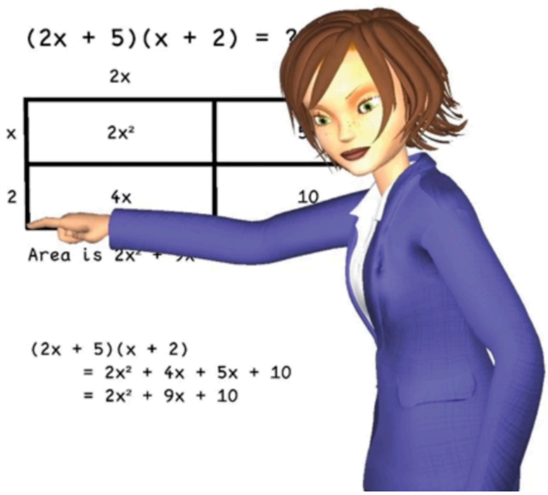


Figure 1. Screenshot of the video lesson with the area-based method on top and the symbolic equation-based method on bottom

Method

Participants

Participants were a convenience sample of 24 seventh- and eighth-grade students (16 seventh-graders; 8 eighth-graders) attending middle school in a mid-sized Midwestern city in the United States. The participants were predominantly White (75% White, 8% Asian, 4% Hispanic, 13% Other) and their mean age was 13.2 years (*min* = 11.5, *max* = 14.2). Thirty-eight percent were female. Students were recruited from three middle schools within a single school district. A district-approved email was sent to all seventh- and eighth-grade students inviting them to participate in a university-sponsored research project that would take place on the university's campus. All students were compensated $ 15 for their participation in a single study session.

Design and procedure

Each student participated in a single one-on-one session that lasted approximately 45 minutes. See Figure 2 for a schematic of the study design. First, students completed a pretest to assess their prior knowledge of polynomial multiplication, as well as relevant background knowledge. Each item was presented one at a time on an interactive smart board. Students could write out their answers on the board using their fingers, and they had access to additional functions as well (e.g., erase, click to proceed, etc.). Students were told that the problems were intended to assess what they already knew, and that it was okay to be unsure.

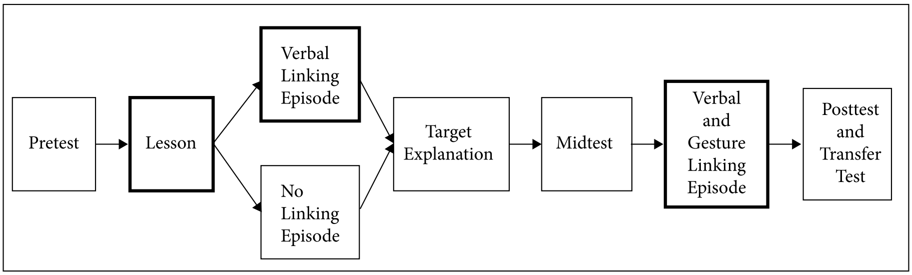


Figure 2. Schematic of the study design; segments of the procedure that involved the avatar teacher are indicated with bold outline

Next, students viewed a lesson on the interactive smart board presented by an avatar video instructor. The avatar instructor provided key information in both speech and gesture. The lesson focused on multiplying binomials using a target problem: $(2x+5)(x+2)$. The avatar instructor described an area-based method first, followed by an equation-based method. Below is an excerpt of the lesson script that includes both speech and gestures:

> … Suppose we want to multiply 2x+5 times x+2. We can model this multiplication using a rectangle. If the length is 2x+5 [points with flat palm to the length of the rectangle] and the width is x+2 [points with flat palm to the width of the rectangle], then the area of this rectangle [points with flat palm to the center of the rectangle] is the product of 2x+5 and x+2. We can break up the length into two parts: 2x and 5 [points with index finger to 5]. And we can break the width into two parts: x and 2 [points with index finger to 2] …. We can find the area of each smaller rectangle [circles four smaller rectangles] and add them all together to find the area of the large rectangle. So, we have the length 2x times the width x [points to 2x then drags point to x with index finger] is $2x^2$. The length 2x times the width 2 [points to 2x then drags point to 2 with index finger] is 4x. The length 5 times the width x [points to 5 then drags point to x with index finger] is 5x. The length 5 times the width 2 [points to 5 then drags point to 2 with

> index finger] is 10. Adding the areas of these smaller rectangles together, we have $2x^2 + 4x + 5x + 10$ [points with flat palm to whole equation] …

Half of the students were then randomly assigned to view a brief verbal "linking episode" in which the avatar instructor used speech to delineate the correspondences between the two representations (e.g., "$2x + 5$ in the equation corresponds to the length $2x + 5$ in the rectangle"). This episode included some beat gestures but did not include any task-relevant gestures. The other half of the students were randomly assigned to not view this episode and to move on to the next activity. This manipulation was included to address a research question unrelated to the focus of the present study, namely, whether exposure to multiple linking episodes (as opposed to a single linking episode) affected the quality of students' connection-making and their problem-solving performance (see Fyfe et al., 2017, for results relevant to this research question). Critically, because this linking episode did not include any task-relevant gestures, we did not expect participants' experience of this linking episode to affect their behavior relevant to the current research questions. However, given that some students were exposed to this additional information, we included condition as a control variable in our analyses.

After the lesson, all students engaged in an explanation of the target problem and solved two problems in a "midtest". The target problem for the explanation and the two midtest problems were displayed on the interactive smart board. The purpose of the explanation episode was to assess students' use of gesture when explaining a target problem, and the purpose of the midtest was to assess their ability to use information from the lesson to solve problems. After the explanation and midtest, all students viewed a brief verbal-and-gesture linking episode in which the avatar video instructor used speech and gesture to delineate the correspondences between the two representations. This linking episode was included to address a different research question related to students' connection-making; we did not expect it to influence behavior relevant to the current research questions, given that it occurred after the target explanation, and given that it was identical for all participants. After viewing this verbal-and-gesture linking episode, all students completed the posttest and transfer test, which were also displayed on the smart board. The purpose of these tests was similar to that of the midtest – namely, to assess students' abilities to use information from the lesson to solve polynomial multiplication problems, and to transfer that information to whole numbers, which were not the focus of the lesson. Throughout the session, students were encouraged to think aloud, so we could gain a richer account of their thought processes (Ericsson & Simon, 1993). Although the lesson was delivered via the avatar video instructor, a human experimenter was present to facilitate and explain the study protocol.

Materials

Materials, coding schemes, and stimuli are publicly available through the Open Science Framework at https://osf.io/9jg3v/?view_only=e88ee41e90fb4910b52baf e5b2abd3a0

Pretest

The pretest included six items that were displayed one at a time on the smart board (see Table 1 for examples). The first two items were *background knowledge items* that tapped students' ability to operate on variables (i.e., $x+x$, $x \times x$). The next three items were *background knowledge items* that tapped students' ability to calculate the areas of rectangles. The sixth item was a target *solve item* that tapped students' prior knowledge of how to solve a polynomial multiplication problem [i.e., $(x+2)(x+1)$].

Lesson

The lesson was presented by the avatar teacher. It included a brief introduction to calculating the area of a rectangle, and then proceeded to focus on multiplying binomials using the target problem: $(2x+5)(x+2)$. First, the avatar teacher explained how to solve the target problem using an area-based method, and then the avatar teacher explained how to do so using an equation-based method. See Figure 1 for a screenshot of the lesson.

Explanation

After the lesson from the avatar teacher, students were shown the instructional problem from the lesson on the interactive smart board. It included the problem statement, $(2x+5)(x+2)$, an area-based model, and an equation-based model (see Figure 3). Students were asked to explain how to solve the target problem. Specifically, they were told: "Here is the same problem you just learned about. Imagine that another student is seeing this example for the first time. Can you explain how to solve this problem?"

Midtest

After the explanation, students answered two items, which we refer to as midtest items (see Table 1). These items were displayed one at a time on the smart board. The first item was a polynomial multiplication *solve* item that students could solve

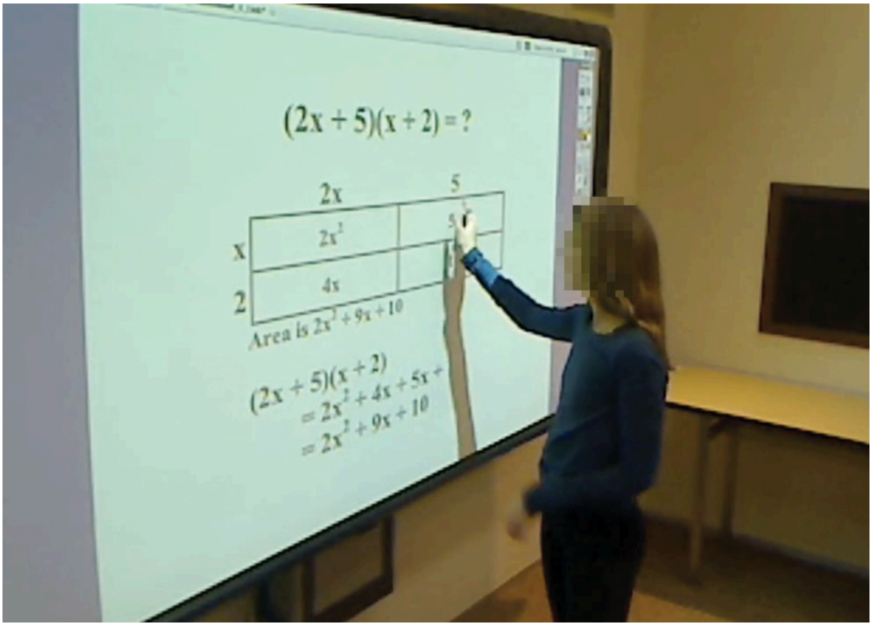


Figure 3. Screenshot of participant explaining target problem during explanation phase

using any method of choice. The second item was a *link* item that tapped students' understanding of the correspondence between the area-based representation and the equation-based representation.

Posttest

The posttest included five items presented one at a time on the smart board (see Table 1 for examples). Two were polynomial multiplication *solve* items that students were asked to solve using a particular method. Three were *link* items.

Transfer test

After the posttest, students also solved two transfer items presented on the smart board, which assessed whether students could apply what they learned about multiplying expressions with variables to multiplying whole numbers (see Figure 4).

Table 1. Example items on pretest, midtest, and posttest assessments with example correct responses

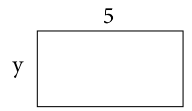
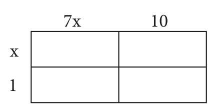
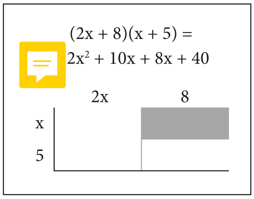
Item type	Example item	Instructions	Possible responses
Pretest item types			
Background knowledge of variable operation	(X)(X)	Simplify the expression.	Correct: 2x Incorrect: 1, 1x, x, 2, x^2
Background knowledge of area calculation		Find the area of the rectangle.	Correct: 5y Incorrect: 5*y, 12, 10 + 2y
Solve item	(X + 2)(X + 1)	Simplify the expression by multiplying the terms x plus 2 and x plus 1.	Correct: $x^2 + 3x + 2$ $x^2 + 2x + 1x + 2$ $1x + 2x + 2 + x^2$ Incorrect: 2x+3, 2x*1x, 3x,3xx
Midtest and posttest item types			
Solve item	(6x + 3)(y + 7)	Simplify the expression by multiplying the terms 6x plus 3 and y plus 7.	Correct: 6xy + 42x + 3y + 21 Incorrect: 6xy + 10, 21*6xy
Link item from equation to rectangles	 (7x + 10)(x + 1) = $(7x^2 + \underline{7x} + 10x + \underline{10}$	The underlined terms in the equation represent the area of two rectangles. Which ones?	Correct: Select bottom left and bottom right Incorrect: Select top left and bottom right

Table 1. *(continued)*

Link item from rectangles to equations		Circle the term in the equation that represents the area of the shaded rectangle.	Correct: 8x Incorrect: 40, 8x+40, 10x

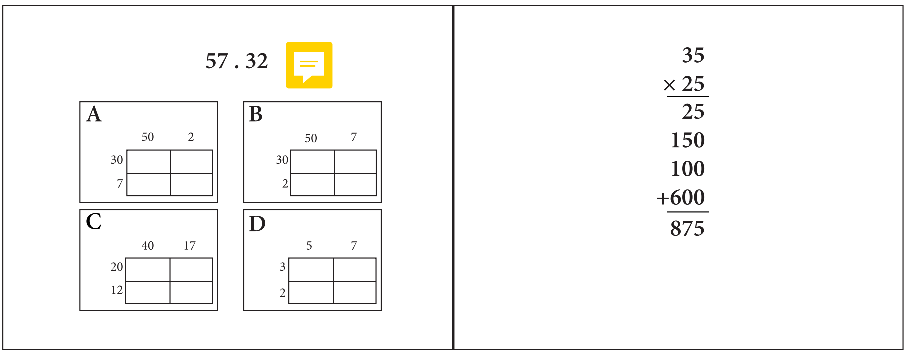


Figure 4. Transfer items
Note. For the item on the left, students were asked to select which area model(s) corresponded to multiplying 57 times 32. For the item on the right, students were asked to decide if the method shown was a correct method for solving 35 times 25.

Coding

Scoring

Items on each assessment were scored as correct or incorrect based on students' written answers and on the verbal think-aloud reports they provided while solving. Table 1 provides sample correct and incorrect responses. For *solve* items, responses were counted as correct if they included the presence and addition of the correct terms. Individual terms did not need to be in a particular order nor did they need to be simplified. For example, for the problem $(x+2)(x+1)$, both of the following represent possible correct answers: $x^2+2x+x+2$ and x^2+3x+2. For *link* items, responses were counted as correct if the correct term was circled or the correct rectangle was shaded. For the *transfer* items, responses were counted as correct if the student provided the exact correct answer (see Figure 4).

Gesture

We used ELAN, a professional software for annotating and coding video data, to code and quantify students' gestures during the pretest and during their explanations of the target problem (ELAN, 2017). We first assessed *gesture frequency*. Each video was coded frame-by-frame for the presence of a student gesture (for the duration of the pretest and for the duration of the target explanation). Coders marked each gesture by noting the start and end time at which the gesture occurred. Each gesture was then classified into one of three categories according to McNeill's (1992) taxonomy. *Pointing* gestures are gestures that indicate locations or inscriptions with an extended finger or hand (e.g., indicating a number on the smart board by pointing directly to that number with the index finger). *Iconic* gestures are gestures that bear a close formal relationship to the semantic content of speech via hand shape or trajectory (e.g., indicating a rectangle by tracing it in the air). *Beat* gestures are simple movements that do not present semantic content but often align with the prosody or discourse structure of speech (e.g., moving hand up and down as one talks).

Second, coders also identified the referents of the gestures (e.g., a pointing gesture might indicate the 5 or the $2x$ or a rectangle). The purpose of mapping gestures to referents was to code for *gesture mimicry* during the target explanation (i.e., to compare gesture patterns between students and the avatar instructor). For example, the avatar gestured with an index finger to the 5 in the rectangle; if a student did so, it would be considered a mimicked gesture because the avatar produced that same gesture. However, the avatar did not point to the 10 in the final equation; if the student did so, it would not be considered a mimicked gesture, because the avatar did not produce that same gesture. In all, the avatar produced 16 unique gestures or gesture patterns. Gesture patterns were sequences of gestures produced in quick temporal order (e.g., point to the 2, then drag point to the 2x). Coders assigned each student a score ranging from 0 to 16 indicating how many gestures or gesture patterns they mimicked.

A similar analysis was done at pretest to code for *task-relevant gestures* that were similar in form to the avatar's gestures. The goal was to obtain a baseline measure of students' spontaneous use of gestures that were similar to the avatar's gestures. For example, at one point in the lesson, the avatar teacher pointed to the 2 on the length of the rectangle and dragged that point to the 2x on the width of the rectangle. One pretest item required students to calculate the area of a rectangle with length 5 and width y; if a student pointed to the length 5 and then dragged the point to the width y, this would be considered a task-relevant gesture similar in form to one of the avatar's gestures. Across all items, coders marked each gesture that was similar in form to one of the sixteen gestures or gesture patterns produced by the avatar.

Two researchers independently coded each student's gestures. Inter-rater agreement was high for the number of gestures produced during the target explanation, $ICC(3,2) = .99$, $p < .001$, the number of gestures produced during the pretest, $ICC(3,2) = .97$, $p < .001$, the classification of gestures as pointing, iconic, or beat, kappa = .85, the number of gesture patterns that were mimicked during the target explanation, $ICC(3,2) = .81$, $p < .001$, and the number of task-relevant gestures at pretest that were similar in form to the avatar's gestures, $ICC(3,2) = .85$, $p < .001$. Discrepancies were resolved by discussion between coders.

Results

Pretest

On average, students solved 3.4 problems correctly on the pretest (out of 6, $SD = 1.5$). Students did moderately well on the five background knowledge items, though performance varied by item (proportion correct on the background knowledge items ranged from .42 to .88) and by student (number of background knowledge items solved correctly out of 5 ranged from 0 to 5). Only one student (out of 24) correctly answered the polynomial multiplication *solve* item: $(x+2)(x+1)$. The two most common errors on that problem were to add the two x's and add the two integers to get $2x+3$, or to incorrectly combine terms within parenthesis to get $2x * 1x$. Thus, the students were best characterized as novices in the domain.

On average, students took 212.1 seconds to complete the pretest ($SD = 66.4$, $range = 125$ to 351); thus, it took about 3.5 minutes. Students produced an average of 10.5 gestures as they spoke aloud about their thinking, which included time to process the problems and execute their solution strategies ($SD = 9.7$, $range = 0$ to 37). The average rate of gesture per minute during the pretest was 2.84 (about one gesture every twenty seconds, $SD = 2.21$, $range = 0.00$ to 7.59). We examined students' gestures at pretest to derive a "baseline" measure for gestures that were similar in form to the gestures that students would later see produced by the teacher avatar during the lesson. On average, students produced 1 gesture that was similar in form to one of the teacher avatar's gestures ($SD = 1.3$, $range = 0$ to 4), though the distribution was skewed, with exactly 50% of the sample never producing a gesture that was similar in form to the avatar's gestures. Most of these task-relevant gestures that were similar in form to the avatar's gestures were produced on symbolic equation-based problems (68%) rather than on rectangle area problems (32%). Indeed, the most common of these gestures was on the polynomial multiplication problem, $(x+2)(x+1)$, and involved pointing to the first x and then dragging the point to the other *x* or pointing to the 2 and dragging the point to the 1. Similar gestures were produced by the avatar during the lesson when describing the equation-based method of multiplication.

Explanation

Following the pretest and lesson, students were asked to explain how to solve a polynomial multiplication problem. Explanations varied on several dimensions. Consider the three explanations presented below, which were selected to highlight these variations. The text in brackets represents the students' gestures, and italicized text in brackets represents a gesture that mimicked one produced during the lesson by the avatar instructor.

Student 1: "You take 2x times x [*points to 2x, then drags point to x*] which is $2x^2$ [points to $2x^2$ in area model], you take 5 times x [*points to* 5, *then drags point to x*] which is 5x [points to 5x in area model], then you take 2x times 2 [*points to 2x, then drags point to* 2] which is 4x [points to 4x in area model], and then 5 times 2 [*points to* 5, *then drags point to* 2] which is 10 [points to 10 in area model]. You can also do it this way [points to equation below] where you take 2x times x [*points to 2x, then drags point to x*] which is $2x^2$, and then 2x times 2 [*points to 2x, then drags point to* 2] which is 4x, and then you can take 5 times x [*points to* 5, *then drags point to x*] which is 5x, then 5 times 2 is 10 [*points to* 5, *then drags point to* 2], then you just add them all together [points to equation below]."

Student 2: "So you do 2x times x [*points to 2x, then drags point to x*], which is $2x^2$ [points to $2x^2$ in area model]. Then 5 times x [*points to* 5, *then drags point to x*] which is 5x [points to 5x in area model]. Then 2x times 2 [*points to 2x, then drags to point* 2] which is 4x [points to 4x in area model] and 5 times 2 [*points to* 5, *then drags point to* 2] which is 10 [points to 10 in area model]. Then you add them together to get the area [*points to whole equation*]."

Student 3: "Okay, so basically what you would do is multiply each number [points to terms in equation] by every other [points to terms in equation] that's in the different set [swoops across equation]. So 2x times x is 2x squared. 2x times 2 is 4x. So both of these [points to terms in equation] are one side so you don't have to multiply these [points to terms in equation]. Then you do 5 [points to 5] times x, which is 5x. And 5 times 2, which is 10. Then [points to entire equation] you would take all those answers together and simplify them. So 4x and 5x is 9x. And then 10 and $2x^2$. So you get $2x^2$ plus 9x plus 10."

As highlighted in these examples, students' explanations varied in length, in gesture rate, and in whether the gestures mimicked those of the avatar. For example, Student 1 mimicked 8 of the avatar's gestures, Student 2 mimicked 5 of the avatar's gestures, and Student 3 did not mimic any of the avatar's gestures.

Students' explanations were 57.5 seconds in length on average ($SD = 45.3$, $range = 9.5$ to 247.5). Thus, they were about one minute long, but ranged from only a few seconds to over 4 minutes. During their explanations, students gestured approximately 28.7 times on average ($SD = 21.8$, $range = 4$ to 108), yielding an average gesture rate of 32.14 gestures per minute (about one gesture every two seconds, $SD = 16.24$, $range = 12.04$ to 68.45). Most gestures were pointing gestures that referred to an element on the screen (84% of all gestures), though some were iconic gestures (7%) or beat gestures (9%).

During the lesson, the avatar produced 16 unique gesture patterns. On average, students mimicked 3.0 of these gesture patterns in their explanations ($SD = 2.5$, $range = 0$ to 8, see Figure 5 for full distribution). In addition to examining the frequency of gesture mimicry, we were also interested in whether students *ever* mimicked the avatar's gestures. In our sample, 20% ($n = 5$) of students never mimicked the avatar's gestures and were thus categorized as non-mimickers. The remaining 80% ($n = 19$) of students mimicked at least one of the avatar's gestures and were thus categorized as mimickers. Most of the mimicked gestures were gestures to the area-based representation (76%) rather than to the equation-based representation (24%). Indeed, the most frequently mimicked gesture was pointing to the length $2x$ on the rectangle and dragging the point to the width x. These mimicked gestures provided during the target explanation were quite different from the task-relevant gestures that students spontaneously produced at pretest. For example, at pretest, only four students used a dragging gesture from length to width across any of the four rectangle area problems. However, during the explanation, thirteen students used a dragging gesture on the rectangle. Also, at pretest, ten students used a dragging gesture across the parentheses of an equation problem, but only one student did so during the target explanation. This lack of alignment between students' spontaneous gestures at pretest and students' gestures after viewing the avatar lessons suggests that students were indeed mimicking the avatar's gestures and adopting them into their own repertoires, rather than producing such gestures spontaneously.

Importantly, non-mimickers still produced gestures – they just did not use gestures like those produced by the avatar. In fact, gesture rates for non-mimickers ($M = 24.1$, $SD = 10.4$) did not differ significantly from gesture rates for mimickers

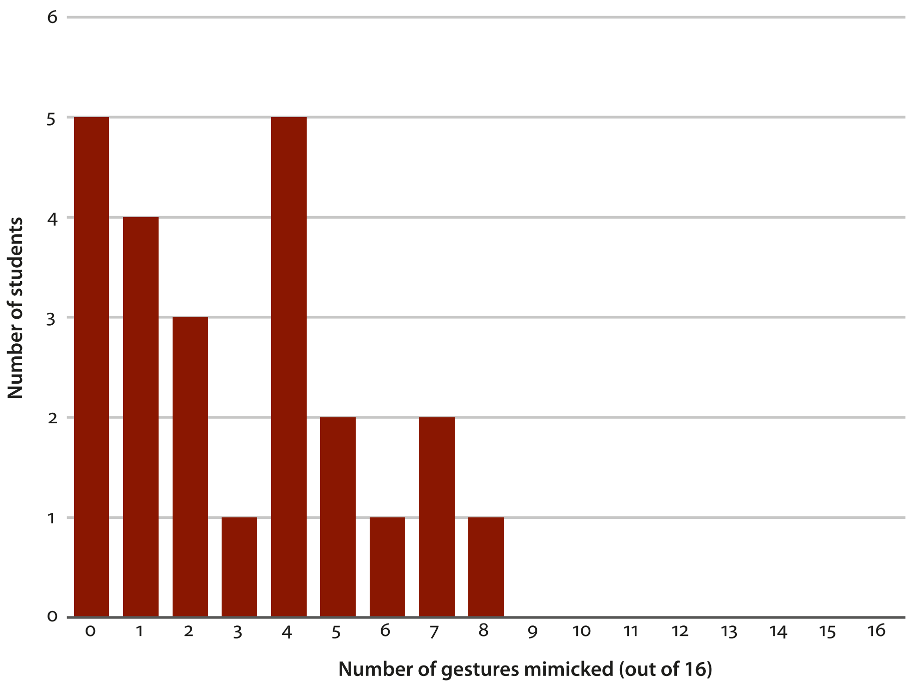


Figure 5. Distribution of gestures mimicked

($M=34.2$, $SD=17.0$), $t(22)=-1.25$, $p=0.22$. Further, although the number of non-mimickers was small, these students were not reliably different from mimickers on any available metric except mimicry. Non-mimickers and mimickers did not differ reliably in age ($M=13.2$ vs. 13.2 years), $t(22)=0.04$, $p=.97$, in pretest scores ($M=3.8$ vs 3.3, out of 6), $t(22)=0.71$, $p=.49$, in explanation length ($M=48.0$ vs. 60.0 seconds), $t(22)=-0.52$, $p=.61$, in percent female (20% vs. 42%, Fisher's Exact $p=.62$), or in percent white (60% vs. 78%, Fisher's Exact $p=.57$). Further, students' gesture production was not reliably associated with their prior knowledge or background characteristics. Gesture rates during the explanation were not significantly correlated with pretest scores, $r(22)=-.31$, $p=.14$, or with age, $r(22)=-.02$, $p=.94$. Similarly, frequency of gesture mimicry was not significantly correlated with pretest scores, $r(22)=-.24$, $p=.26$, or with age, $r(22)=.12$, $p=.59$.

Learning from the lesson

After the explanation, students answered two midtest items, viewed a brief instructional episode that was the same for all students, and then completed the posttest (5 items) and transfer test (2 items). The data for each measure (midtest, posttest, and transfer test) are plotted separately in Figure 6; however, for simplicity, we combined scores on the midtest, posttest and transfer items into a single score for analysis (Cronbach's alpha = .80). This comprehensive *post-lesson assessment score* allowed us to evaluate whether variations in students' gestures during the explanation related to performance. Overall, scores were high, with students solving an average of 6.3 problems correctly out of 9 ($SD = 2.4$). Most students demonstrated some learning. At pretest, only one student (4% of the sample) solved a polynomial multiplication problem correctly, but at posttest, 17 of the 24 students (71%) did so.

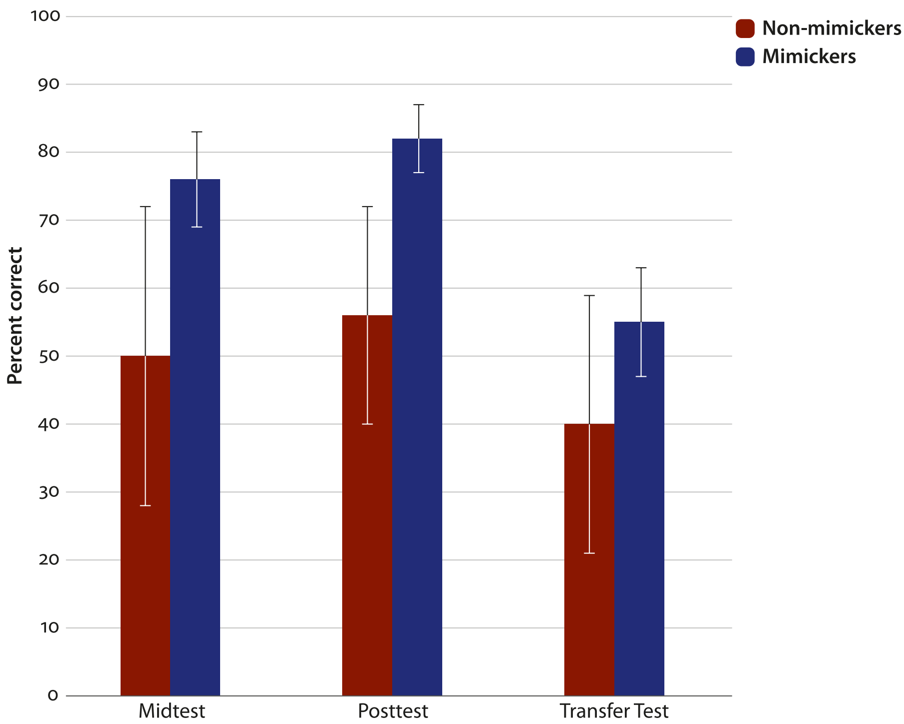


Figure 6. Raw scores on midtest, posttest, and transfer test items for mimickers and non-mimickers; error bars reflect standard errors

To test whether gesture mimicry was associated with performance after the lesson, we conducted an ANCOVA with gesture mimicry (mimickers vs. non-mimickers) as a between-subjects variable and total post-lesson assessment scores (out of 9) as the dependent variable. We included four covariates in the model: accuracy on the pretest (out of 6), use of task-relevant gestures at pretest that were similar in form to the avatar's gestures (yes vs. no), gesture rate during the explanation (per minute), and linking condition (exposed to verbal link vs. not exposed, see Method). As shown in Figure 7, there was a significant main effect of gesture mimicry, $F(1,18)=6.77$, $p=0.018$, $\eta_p^2=0.27$. Students who mimicked the avatar's gestures had higher total scores (*Estimated Marginal Mean* = 7.0 out of 9, *SE* = 0.5) than students who did not (*Estimated Marginal Mean* = 3.8, *SE* = 1.1). None of the four covariates were statistically significant predictors of post-lesson assessment scores, *p's* > 0.05, including overall gesture rate during the explanation, $F(1,18)=0.55$, $p=0.470$, $\eta_p^2=0.03$. See the Appendix for full model statistics.

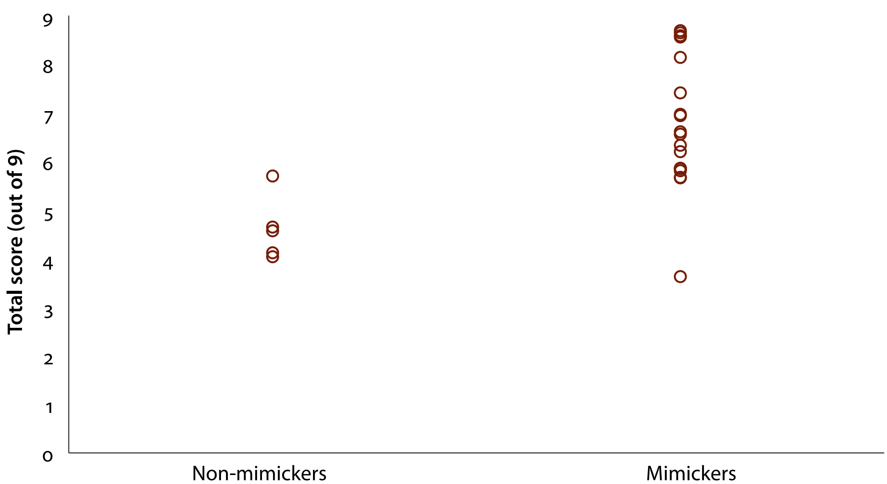


Figure 7. Individual students' total assessment scores as a function of gesture mimicry
Note. Each point represents an individual student. Scores are unstandardized predicted values from an ANCOVA model adjusted for four covariates: accuracy on pretest (out of 6), use of task-relevant gestures at pretest that are similar in form to avatar's gestures (yes vs. no), gesture rate during explanation (per minute), and condition (exposed to verbal link vs. not exposed).

We also explored whether post-lesson assessment scores were associated with the *amount* of gesture mimicry, rather than the *presence* of gesture mimicry. We carried out a linear regression with the frequency of mimicked gestures (out of 16 possible) as the independent variable and total scores (out of 9) as the dependent variable. We included the same four covariates as in the preceding analysis: accuracy on pretest, use of task-relevant gestures at pretest, gesture rate during explanation, and linking condition. In this model, the relation between gesture mimicry and total scores was positive, but not statistically significant, $\beta=.35$, $p=.204$. Further, if we consider only students who mimicked at least one gesture and split them into low-mimickers (i.e., 1, 2, or 3 mimicked gestures, $n=8$) versus high-mimickers (i.e., 4 or more mimicked gestures, $n = 11$), we find no effect of gesture mimicry on students' total scores, after controlling for the four covariates, $F(1, 13) = 0.01$, $p = 0.946$, $\eta_p^2 = 0.00$.

Discussion

In the current study, we examined middle school students' gestures as they learned from an algebra lesson provided by a gesturing avatar video instructor. We hypothesized that students would vary in how frequently they gestured and in whether they mimicked the avatar's gestures. Further, we hypothesized that students who mimicked the avatar's gestures would perform better on the post-lesson assessments than students who did not mimic the avatar's gestures. The results from a convenience sample of 24 middle school students supported these hypotheses. After viewing a brief lesson from the gesturing avatar, many students produced gestures that were similar in form to the avatar's gestures in their own explanations of the target material.

Moreover, those students performed better on the post-lesson assessments than students who did not mimic the avatar's gestures, though it is important to view these findings with caution, given the small sample size. Importantly, students who *did not* mimic the avatar's gestures (non-mimickers) were not statistically different from students who *did* mimic the avatar's gestures (mimickers) on a variety of metrics, including how frequently they gestured during the explanation, the length of their explanations, their pretest scores, and demographic characteristics. These findings suggest that gesture mimicry differentiated students in a unique way that related to their performance, and they warrant further research on gesture mimicry and learning.

A large body of research has investigated the benefits of gesture in a variety of contexts (see Kelly, Church, & Alibali, 2017), including contexts in which teachers incorporate gesture into lessons (e.g., Valenzeno et al., 2003), students spontaneously gesture while problem solving (e.g., Cook & Goldin-Meadow, 2006), and students are directed to gesture during a lesson (e.g., Broaders et al., 2007). A few studies have focused more explicitly on the *relations* between teachers' and students' gestures, suggesting that *gesture mimicry* may be particularly important (e.g., Cook, Mitchell, & Goldin-Meadow, 2008; de Nooijer, Van Gog, Paas, & Zwaan, 2013). Our findings contribute to this literature in at least two key ways: (1) by providing evidence that suggests students mimicked the gestures produced by an avatar instructor, and (2) by showing that gesture mimicry was associated with scores on a post-lesson assessment of polynomial multiplication, an important topic in early algebra.

An increasing amount of research has focused on whether and how learners can benefit from computer-animated pedagogical agents. Indeed, as educational technology continues to advance, learners have increasing numbers of opportunities to engage with novel, multimedia environments. Avatar instructors have been incorporated into a variety of electronic learning environments, and research suggests that these avatars can have positive effects on learning and motivation (e.g., Baylor, 2003; Holmes, 2007; Lester, Converse, Kahler et al., 1997; Lusk & Atkinson, 2007; Moreno & Mayer, 2007). Moreover, a recent meta-analysis of 20 experiments confirmed that gesturing pedagogical avatars are more effective than non-gesturing pedagogical avatars on measures of student learning (Davis, 2018). The current results provide insight into the students' experience learning from a gesturing avatar and suggest that opportunities for gesture mimicry may play a role.

Among our convenience sample of 24 middle school students, 80% of them produced a gesture that was similar in form to the avatar's gestures during their target explanation. One possibility is that these students were responding naturally to the task at hand. That is, regardless of whether the avatar had gestured or not, these students may have produced these task-relevant gestures in reference to these specific visual representations. This possibility would be consistent with previous research finding variations in learners' spontaneous production of relevant gestures on problem-solving tasks (e.g., Alibali et al., 2011; Church & Goldin-Meadow, 1986; Cook & Goldin-Meadow, 2006). Another possibility is that these students were mimicking the avatar's gestures and adopting them into their own gesture repertoires. This possibility would be consistent with previous research reporting spontaneous gesture mimicry in students learning from a human instructor. For example, Cook and Goldin-Meadow (2006) reported that nearly 50% of elementary school students mimicked a human instructor's gestures that portrayed a correct problem-solving strategy for a mathematics problem.

An experimental comparison of students' responses to a gesturing versus non-gesturing avatar is needed to tease apart these possibilities in a definitive way. However, evidence from the current study tentatively suggests that some students mimicked the avatar's gestures. First, students were relatively unlikely to produce these gestures at baseline. At pretest, only half of the students ever produced a gesture that was similar in form to the gesture's avatars, even though the items were conceptually and visually similar to the item the avatar presented. Second, of the task-relevant gestures that students produced at baseline, very few were similar in form to the gestures that students produced during their target explanation. Thus, there was a lack of alignment between the gestures that students spontaneously produced at baseline and those that they produced after watching the avatar. Indeed, the clear majority of gestures that the students produced after watching the avatar referred to the area-based representation (e.g., pointing to the value of the height of one of the rectangles and dragging the point to the value of the width of that rectangle to signal multiplication). These area-based gestures were remarkably rare at baseline, suggesting that students' gestures during their explanations reflected gestures that they observed from the avatar rather than gestures that they would have produced spontaneously. If true, then the current findings imply that students are likely to view avatars as social agents and subsequently engage in typical interactive gesture behavior, even in this novel technology-mediated instructional space.

In addition to demonstrating that students produced gestures that aligned with those of the avatar, the results also indicated that variations in students' gestures related to their performance on the post-lesson assessments. Students who mimicked at least one of the avatar's gestures had higher total scores than students who did not mimic any of the avatar's gestures. Further, this association was specific to our measure of gesture mimicry; our measure of gesture frequency during the target explanation was not significantly associated with post-lesson assessment scores. One possibility is that the mimicked gestures reflected higher prior knowledge of the target strategy. Although possible, we do not think this the most likely explanation, given that we controlled for pretest accuracy in the analysis.

A second possibility is that these gestures reflected students' attention to the lesson or knowledge obtained from the lesson. For example, perhaps the students who attended to the lesson were both more likely to notice the avatar's gestures and reproduce them and more likely to encode the instructed strategy in a way that enhanced their performance. Similarly, students who *gained* more knowledge from the lesson may have been more likely to display that knowledge, both in gesture and in their problem solving, relative to students who did not gain knowledge from the lesson. If either of these explanations are true, then gesture mimicry in the current study is functioning as an *indicator* of learning, much like students' scores on the assessments. In this case, one would expect an association between these two indices of learning. This is consistent with previous research that has used the contents of students' gestures to assess what students know about a topic (e.g., Alibali, Flevares, & Goldin-Meadow, 1997; Church & Goldin-Meadow, 1986, see also Kimbara, 2006).

A third possibility is that mimicking gestures *facilitated* students' learning, rather than reflected it. That is, engaging in the explanation episode and producing gestures like those of the avatar may have altered (and improved) students' knowledge. This possibility is consistent with experimental studies showing that asking students to mimic the gestures of a human tutor can influence learning, both in mathematics (e.g., Cook et al., 2008; Goldin-Meadow et al., 2009) and in word learning (e.g., de Noojier et al., 2013; Tellier, 2008). Thus, gesture may play a causal role in learning, perhaps by giving learners an alternate, embodied way of representing ideas.

Although it is possible that a similar mechanism is driving the association between gesture mimicry and post-lesson assessment scores in the current study, the current data are correlational and cannot be used to tease apart these explanations. Future research that experimentally manipulates whether learners mimic an avatar instructor's gestures is needed. If mimicking the avatar's gesture does actually facilitate learning, then students who are encouraged to imitate an avatar's gestures should show greater learning than students who are not encouraged to do so or who are prevented from doing so. Promising work by Nathan and Walkington (2017) suggests that avatars in a game-based setting can direct the type of gestures that students make during mathematical reasoning, and that the types of gestures students mimic may relate to their learning. This research is laying the groundwork for future intervention studies that examine the possible causal effect of mimicking gestures from an avatar on a variety of learning outcomes.

Several additional limitations of the current study suggest other fruitful avenues for future research. First, our sample was a convenience sample of 24 middle school students, which limits the inferences we can make to the broader population. The sample also limits our ability to investigate variability in the rate and types of gestures that students produce in this novel learning environment. Future studies should incorporate larger, more diverse samples to better examine variability in gesture mimicry and to enhance the generalizability of the results. Second, our research was limited to a single avatar instructor. Avatars are sometimes perceived to lack natural tendencies and can enter the "uncanny valley," making it particularly important to examine variations in avatars' appearances and movements (e.g., Mori et al., 2012). The avatar in the current study was developed to test accounts of learning with gesture, for example by controlling eye gaze, head tilts, and so forth. Future research should continue to take advantage of the benefits of using avatar instructors in gesture research (see Cook et al., 2017), including testing variations in the type and appearance of the avatar being used. Third, as mentioned above, the current study was correlational in nature. Future research should conduct experimental comparisons between (a) gesturing avatars and non-gesturing avatars, and (b) gesturing avatars and gesturing humans, so as to better understand the mechanisms by which avatars and avatars' gestures influence student learning.

Despite these limitations, the current study provides novel insights into students' gestures as they learned from a mathematics lesson delivered by a novel teacher avatar. After viewing the lesson, students varied in their gesturing behavior, including whether or not they produced gestures that were similar in form to those modeled by the avatar. In large part, these gestures differed from those spontaneously produced at baseline. Further, students who produced such gestures learned more than students who did not. Our findings suggest that the *relations* between the teacher's gestures and the students' gestures may be particularly important to consider in learning contexts, even (or perhaps especially) when the teacher is a computer-animated pedagogical agent.

Funding

Support for this research was provided by the U.S. Department of Education, Institute of Education Sciences grant number R305A130016 to Alibali and training grant number R305B130007 as part of the Wisconsin Center on Education Research Postdoctoral Training Program in Mathematical Thinking, Learning, and Instruction at the University of Wisconsin – Madison.

Acknowledgements

The authors thank Voicu Popescu for his work developing the avatar instructor.

References

Adamo-Villani, Nicoletta, Wilbur, Ronnie, Eccarius, Petra, & Abe-Harris, Laverne. (2009). Effects of character geometric model on the perception of sign language animation. *Proceedings of the 2009 Second International Conference in Visualization* (pp. 72–75). Washington, DC: IEEE Computer Society. https://doi.org/10.1109/VIZ.2009.23

Anasingaraju, Saikiran, Wu, Meng-Lin, Adamo-Villani, Nicoletta, Popescu, Voicu, Cook, Susan W., Nathan, Mitchell, & Alibali, Martha W. (2016). Digital learning activities delivered by eloquent instructor avatars: Scaling with problem instance. In *Proceedings of SIGGRAPH Asia 2016 Symposium on Education*, Article 5, 1–7. New York, NY: ACM. https://doi.org/10.1145/2993352.2993355

Alibali, Martha W., Flevares, Lucia M., & Goldin-Meadow, Susan. (1997). Assessing knowledge conveyed in gesture: Do teachers have the upper hand? *Journal of Educational Psychology*, *89*, 183–193. https://doi.org/10.1037/0022-0663.89.1.183

Alibali, Martha W., Nathan, Mitchell, Church, Ruth Breckinridge, Wolfgram, Matthew S., Kim, Suyeon, & Knuth, Eric J. (2013). Teachers' gestures and speech in mathematics lessons: Forging common ground by resolving trouble spots. *ZDM*, *45* (3), 425–440. https://doi.org/10.1007/s11858-012-0476-0

Alibali, Martha W., Spencer, Robert C., Knox, Lucy, & Kita, Sotaro. (2011). Spontaneous gestures influence strategy choices in problem solving. *Psychological Science*, *22* (9), 1138–1144. https://doi.org/10.1177/0956797611417722

Bauer, Patricia J. & Mandler, Jean M. (1992). Putting the horse before the cart: The use of temporal order in recall of events by one-year-old children. *Developmental Psychology*, *28* (3), 441. https://doi.org/10.1037/0012-1649.28.3.441

Baylor, Amy L. (2003, July). The impact of three pedagogical agent roles. In *Proceedings of the second international joint conference on Autonomous agents and multiagent systems* (pp. 928–929). ACM. https://doi.org/10.1145/860575.860729

Berch, Denise, Singleton, Jenny L., & Perry, Michelle. (1995). Many hands make learning math easy. Paper presented at the biennial meeting of the Society for Research in Child Development, Indianapolis, IN.

Black, Matthew, Chang, Jeannette, Chang, Jonathan, & Narayanan, Shrikanth. (2009, November). Comparison of child-human and child-computer interactions based on manual annotations. In *Proceedings of the 2nd Workshop on Child, Computer and Interaction* (p. 2). ACM. https://doi.org/10.1145/1640377.1640379

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

Buisine, Stéphanie & Martin, Jean-Claude. (2007). The effects of speech–gesture cooperation in animated agents' behavior in multimedia presentations. *Interacting with Computers*, *19* (4), 484–493. https://doi.org/10.1016/j.intcom.2007.04.002

Cassell, Justine, Sullivan, Joseph, Prevost, Scott, & Churchill, Elizabeth. (2000). *Embodied conversational agents*. Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/2697.001.0001

Chartrand, Tanya L. & Bargh, John A. (1999). The chameleon effect: the perception–behavior link and social interaction. *Journal of personality and social psychology*, *76* (6), 893–910. https://doi.org/10.1037/0022-3514.76.6.893

Church, Ruth Breckinridge, Ayman-Nolley, Saba, & Mahootian, Shahrzad. (2004). The role of gesture in bilingual education: Does gesture enhance learning? *International Journal of Bilingual Education and Bilingualism*, *7* (4), 303–319. https://doi.org/10.1080/13670050408667815

Church, Ruth Breckinridge & Goldin-Meadow, Susan. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition*, *23* (1), 43–71. https://doi.org/10.1016/0010-0277(86)90053-3

Cook, Susan W., Duffy, Ryan G., & Fenn, Kimberly M. (2013). Consolidation and transfer of learning after observing hand gesture. *Child Development*, *84* (6), 1863–1871. https://doi.org/10.1111/cdev.12097

Cook, Susan W., Friedman, Howard S., Duggan, Katherine A., Cui, Jian, & Popescu, Voicu. (2017). Hand gesture and mathematics learning: lessons from an Avatar. *Cognitive Science*, *41* (2), 518–535. https://doi.org/10.1111/cogs.12344

Cook, Susan W. & Goldin-Meadow, Susan. (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. https://doi.org/10.1207/s15327647jcd0702_4

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

Davis, Robert O. (2018). The impact of pedagogical agent gesturing in multimedia learning environments: A meta-analysis. *Educational Research Review*, *24*, 193–209. https://doi.org/10.1016/j.edurev.2018.05.002

ELAN (Version 5.0.0-beta) [Computer software]. (2017, April 18). Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved from https://tla.mpi.nl/tools/tla-tools/elan/
Ericsson, K. Anders & Simon, Herbert A. (1993). *Protocol analysis*. Cambridge, MA: MIT press. https://doi.org/10.7551/mitpress/5657.001.0001
Fyfe, Emily R., Alibali, Martha W., & Nathan, Mitchell. (2017). The promise and pitfalls of making connections in mathematics. In E. Galindo & J. Newton (Eds.), *Proceedings of the 39th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 717–724). Indianapolis, IN: Hoosier Association of Mathematics Teacher Educators.
Goldin-Meadow, Susan. (2005). *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.
Goldin-Meadow, Susan & Alibali, Martha W. (2013). Gesture's role in speaking, learning, and creating language. *Annual Review of Psychology*, *64*, 8.1–8.27. https://doi.org/10.1146/annurev-psych-113011-143802
Goldin-Meadow, Susan & Beilock, Sian L. (2010). Action's influence on thought: The case of gesture. *Perspectives on Psychological Science*, *5* (6), 664–674. https://doi.org/10.1177/1745691610388764
Goldin-Meadow, Susan, Cook, Susan W., & Mitchell, Zachary A. (2009). Gesturing gives children new ideas about math. *Psychological Science*, *20* (3), 267–272. https://doi.org/10.1111/j.1467-9280.2009.02297.x
Holmes, Jeffrey. (2007). Designing agents to support learning by explaining. *Computers & Education*, *48* (4), 523–547. https://doi.org/10.1016/j.compedu.2005.02.007
Hostetter, Autumn B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin*, *137* (2), 297. https://doi.org/10.1037/a0022128
Kimbara, Irene. (2006). On gestural mimicry. *Gesture*, *6*(1), 39–61. https://doi.org/10.1075/gest.6.1.03kim
Kelly, Spencer D., Church, Ruth Breckinridge, & Alibali, Martha W. (2017). Understanding gesture: Description, mechanism, function. In Ruth Breckinridge Church, Martha W. Alibali, & Spencer D. Kelly (Eds.), *Why gesture? How the hands function in speaking, thinking, and communicating* (pp. 3–10). Amsterdam: John Benjamins. https://doi.org/10.1075/gs.7.01kel
Kontra, Carly, Goldin-Meadow, Susan, & Beilock, Sian L. (2012). Embodied learning across the life span. *Topics in Cognitive Science*, *4* (4), 731–739. https://doi.org/10.1111/j.1756-8765.2012.01221.x
Lester, James C., Converse, Sharolyn A., Kahler, Susan E., Barlow, S. Todd, Stone, Brian A., & Bhogal, Ravinder S. (1997, March). The persona effect: affective impact of animated pedagogical agents. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems* (pp. 359–366). ACM. https://doi.org/10.1145/258549.258797
Libertus, Klaus & Needham, Amy. (2010). Teach to reach: The effects of active vs. passive reaching experiences on action and perception. *Vision Research*, *50* (24), 2750–2757. https://doi.org/10.1016/j.visres.2010.09.001
Louwerse, M. Max, Graesser, Arthur C., McNamara, Danielle S., & Lu, Shulan. (2009). Embodied conversational agents as conversational partners. *Applied Cognitive Psychology*, *23*, 1244–1255. https://doi.org/10.1002/acp.1527

Lusk, Mary M. & Atkinson, Robert K. (2007). Animated pedagogical agents: Does their degree of embodiment impact learning from static or animated worked examples? *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *21* (6), 747–764. https://doi.org/10.1002/acp.1347

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

Meltzoff, Andrew N. & Moore, M. Keith. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198* (4312), 75–78. https://doi.org/10.1126/science.198.4312.75

Moreno, Roxana & Mayer, Richard. (2007). Interactive multimodal learning environments. *Educational Psychology Review*, *19* (3), 309–326. https://doi.org/10.1007/s10648-007-9047-2

Mori, Masahiro, MacDorman, Karl F., & Kageki, Norri. (2012). The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine*, *19* (2), 98–100. https://doi.org/10.1109/MRA.2012.2192811

Moses, Robert P. & Cobb, Charles E. (2001). *Radical equations: Math literacy and civil rights*. Boston: Beacon Press.

Nathan, Mitchell J. & Walkington, Candace. (2017). Grounded and embodied mathematical cognition: Promoting mathematical insight and proof using action and language. *Cognitive research: principles and implications*, *2* (1), 9. https://doi.org/10.1186/s41235-016-0040-5

NGACBP (National Governors Association Center for Best Practices) & Council of Chief State School Officers. (2010). *Common Core state standards for mathematics*. Washington, DC: Authors.

de Nooijer, Jacqueline A., Van Gog, Tamara, Paas, Fred, & Zwaan, Rolf A. (2013). Effects of imitating gestures during encoding or during retrieval of novel verbs on children's test performance. *Acta Psychologica*, *144* (1), 173–179. https://doi.org/10.1016/j.actpsy.2013.05.013

Noma, Tsukasa, Zhao, Liwei, & Badler, Norman I. (2000). Design of a virtual human presenter. *IEEE Computer Graphics Applications*, *20* (4), 49–85. https://doi.org/10.1109/38.851755

Parrill, Fey & Kimbara, Irene. (2006). Seeing and hearing double: The influence of mimicry in speech and gesture on observers. *Journal of Nonverbal Behavior*, *30* (4), 157–166. https://doi.org/10.1007/s10919-006-0014-2

Ping, Raedy & Goldin-Meadow, Susan. (2008). Hands in the air: Using ungrounded iconic gestures to teach children conservation of quantity. *Developmental Psychology*, *44* (5), 1277–1287. https://doi.org/10.1037/0012-1649.44.5.1277

Popescu, Voicu, Adamo-Villani, Nicoletta, Wu, Meng-Lin, Rajesakaran, Suren D., Alibali, Martha W., Nathan, Mitchell J., & Cook, Susan W. (2014, April). Animation killed the video star. Paper presented at the CHI (Conference on Human Factors in Computing Systems), Satellite Workshop on *Gesture-Based Interaction Design: Communication and Cognition*. Toronto, Canada.

Singer, Melissa A. & Goldin-Meadow, Susan. (2005). Children learn when their teacher's gestures and speech differ. *Psychological Science*, *16* (2), 85–89. https://doi.org/10.1111/j.0956-7976.2005.00786.x

Tellier, Marion. (2008). The effect of gestures on second language memorisation by young children. *Gesture*, *8* (2), 219–235. https://doi.org/10.1075/gest.8.2.06tel

Valenzeno, Laura, Alibali, Martha W., & Klatzky, Roberta. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology*, *28* (2), 187–204. https://doi.org/10.1016/S0361-476X(02)00007-3

Yeo, Amelia, Cook, Susan W., Nathan, Mitchell J., Popescu, Voicu, & Alibali, Martha W. (2018). Instructor gesture improves encoding of mathematical representations. In T. T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (Eds.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society* (pp. 2723–2728).

Appendix

The table below presents the full results from the primary ANCOVA model reported in the results section. The ANCOVA includes gesture mimicry (mimickers vs. non-mimickers) as a between-subjects variable and total post-lesson assessment scores (out of 9) as the dependent variable. Four covariates were included in the model: accuracy on the pretest (out of 6), use of task-relevant gestures at pretest that were similar in form to the avatar's gestures (yes vs. no), gesture rate during the explanation (per minute), and linking condition (exposed to verbal link vs. not exposed).

Predictor	*Df*	*Mean square*	*F*	*p*	η_p^2
Intercept	1	37.866	8.450	.009	.319
Gesture mimicry	1	30.344	6.771	.018	.273
Pretest accuracy	1	12.490	2.787	.112	.134
Task-relevant gestures at pretest	1	0.084	0.019	.893	.001
Gesture rate	1	2.443	0.545	.470	.029
Linking condition	1	13.474	3.007	.100	.143

Address for correspondence

Nicholas A. Vest
Department of Psychology
University of Wisconsin – Madison
1202 West Johnson Street
Madison, WI 53706
USA
navest@wisc.edu

Biographical notes

Nicholas A. Vest graduated from Indiana University in 2016 with a B.S. in Psychology and a certificate in Neuroscience. Following his undergraduate studies, Nicholas pursued his research interests as the lab manager of the Learning, Education, and Development Lab at Indiana

University (PI: Emily Fyfe). He is now a second-year graduate student in the Department of Psychology at the University of Wisconsin – Madison (Advisor: Martha Alibali). His research investigates numerical cognition with a focus on the development of abstract mathematical concepts. He seeks to understand how these concepts form through examining gesture and action.

Emily R. Fyfe is an Assistant Professor in the Department of Psychological and Brain Sciences at Indiana University. She is the director of the Learning, Education, and Development Lab where she conducts her research on the development of mathematics knowledge and problem solving. Emily's goal is to understand how children think and learn about math, both independently and with instructional guidance. Emily received the Rising Star Award in 2017 by the Association for Psychological Science and was a recipient of a 2018 Understanding Human Cognition Scholar Award by the James S. McDonnell Foundation to pursue her research on mathematical thinking.

Mitchell J. Nathan is a Vilas Distinguished Professor of Learning Sciences in the Educational Psychology Department at the University of Wisconsin – Madison. He holds faculty appointments in the Department of Curriculum and Instruction, the Psychology Department, and the Wisconsin Center for Education Research. Mitchell's research is largely rooted in cognitive, embodied, and social perspectives on learning and instruction in STEM fields. He examines the nature of representations in both their internal and external forms, instruction, problem solving, and learning. He is a Fellow of the International Society of the Learning Sciences.

Martha W. Alibali is Vilas Distinguished Achievement Professor of Psychology and Educational Psychology at the University of Wisconsin – Madison, and she is a principal investigator at the Wisconsin Center for Education Research. She conducts research at the interface of developmental psychology, cognitive psychology, and mathematics education. A central focus of her work is on role of gestures in learning and instructional communication. Martha is a past recipient of the Friedrich Wilhelm Bessel Research Prize from the Alexander von Humboldt Foundation, and she is co-author (with Robert Siegler) of the cognitive development textbook, *Children's thinking*.

Author Queries

- Please provide a complete reference for the citation '(de Noojier et al., 2013)' in this article.
- Please provide a citation for the reference id "CIT0004 (Alibali, Martha W., Nathan, Mitchell, Church, Ruth Breckinridge, Wolfgram, Matthew S., Kim, Suyeon, & Knuth, Eric J. (2013)), CIT0006 (Bauer, Patricia J. & Mandler, Jean M. (1992)), CIT0009 (Black, Matthew, Chang, Jeannette, Chang, Jonathan, & Narayanan, Shrikanth. (2009, November)), CIT0013 (Chartrand, Tanya L. & Bargh, John A. (1999)), CIT0038 (Meltzoff, Andrew N. & Moore, M. Keith. (1977))" since citation is missing in the article.