Chapter 18

An embodied cognition perspective on symbols, gesture, and grounding instruction

Mitchell J Nathan

18.1 Introduction

The need to understand and predict behaviour in complex settings such as the classroom and the workplace elevates the importance of the role of context and communication in building models of cognition. Embodied cognition is an emerging framework for understanding intellectual behaviour in relation to the physical and social environment and to the perception- and action-based systems of the body. By reconsidering cognition with regard to interactions with the world, rather than in terms of the sequestered computational nature of the mind, embodied cognition recasts many of the central issues of the study of thought and behaviour. One of the ways that cognition is seen as embodied is through the close relation of hand gestures with thinking and communication. In this chapter, I investigate how gestures enact symbols and thereby ground the meaning of abstract representations used in instructional settings.

Central to this inquiry are two principles that follow from the theory of embodied cognition (Wilson 2002): that cognition is situated and that cognitive work is off-loaded onto the environment. A third principle – that off-line cognition is body based – is also considered, but reframed to include the influences of social interactions, along with sensorimotor processes, as mediating cognitive behaviour even when others are not present. This last issue is developed more fully in the closing section of this paper.

Instruction is a communicative act. As such, what research reveals about instruction, and the learning that ensues, informs our understanding of the production and comprehension of language. When instruction is examined as discourse, it becomes clear that gesture is ubiquitous (Alibali and Nathan, in press). A review of the literature (e.g., Church et al. 2004; Perry et al. 1995; Roth 2003; Singer and Goldin-Meadow 2005; Valenzeno et al. 2003) also shows that gesture use is influential. Teachers’ uses of gestures have been shown to facilitate learner comprehension of material and performance on later tasks. For example, students are more likely to reiterate ideas from tutors’ speech when that speech is accompanied by matching gestures (Goldin-Meadow et al. 1999). Preschool students’ learning and explanations of bilateral symmetry is enhanced when instruction makes use of appropriate gestures (Valenzeno et al. 2003). Gestures also help
to elaborate teachers’ explanations of complex material (Roth and Lawless 2002) and to
direct students’ attention to important features of formal mathematical representations
(Stevens and Hall 1998). Furthermore, there is evidence that gestures are more often used
in instruction when the curricular material is unfamiliar or abstract (Alibali and Nathan,
in press). Taken together, these studies suggest that teachers’ gestures can have a substantial,
positive impact on students’ learning of complex material.

An embodied view of cognition would seem to have a great deal to contribute to our
understanding of the nature and influence of gestures. Yet, the literatures on gesture
research and on embodied cognition are largely independent. The aim here is to show
that embodied cognition can inform research on gesture, and that findings from the
gesture literature can contribute to embodied theories of cognition. To foreshadow my
position, I argue first that gestures are physically realized forms of communication that
have the capability to serve as signs at all of Peirce’s (1909) levels of signification,
performing iconic and symbolic functions, as well as an indexical role. For example,
I show examples of how gesture use in instructional settings physically enacts symbols,
and how the reoccurrence of distinct features of instructional gestures provide cohesion
over the length of a discourse. This serves to evoke a consistent meaning even when the
referent varies or is absent. In this way, gestures exhibit *symbolic off-loading* that helps to
convey complex and abstract ideas (Iverson and Goldin-Meadow 1998; Krauss 1998),
and to offset the large cognitive demands of high-bandwidth exchanges typical of
instructional interactions (Alibali and Nathan 2005).

Secondly, gestures can provide grounding for abstract ideas and representations.
Grounding, as an aspect of neural processing, provides an account of how sensorimotor
experience gives rise to concepts (Havas, personal communication; also see Pulvermüller,
Chapter 6, this volume). Yet, grounding is also used to describe a mapping between an idea
and a more concrete referent, such as an object in the world (Glenberg et al., Chapter 1,
this volume; Gunderson and Gunderson 2006), in order to facilitate meaning making
(e.g., Glenberg and Robertson 1999; Harnard 1990; Lakoff and Johnson 1999; Lakoff and
Nunez 2000). One definition of grounding that follows from Peirce (see Glenberg et al.,
Chapter 1) is to view signs as a means of grounding, in that they connect the interpretant
or idea to something tangible, such as an object or event. Similarly, Roy (2005, as presented
by Glenberg et al., Chapter 1) suggests that grounding denotes the processes by which an
agent (human or machine) relates mental structures to external objects.

In instructional settings, this latter form of grounding is often marked by *linking gestures* that provide conceptual correspondences between familiar and unfamiliar
entities. This is particularly evident when the focus of instruction is on learning to
understand and produce formal (i.e., conventionalized) representations that have an
arbitrary mapping between a representational form and its meaning, as with symbols in
Peirce’s sense. For example, a teacher may provide a gesture as a link between a mathematical symbol, such as $L$, and the length of a geometric shape by pointing to $L$ and then
tracing the longest side of a rectangle (Alibali and Nathan 2005).

Instructional settings are interesting places to study cognition. They are complex,
social environments that involve multiple participants with varied capabilities and goals;
they are relatively uncontrolled environments; subject to real-time constraints on thinking and action; and are focused on intellectually demanding activities such as learning and transfer, teaching, metacognition, and reflective thought that take place over extended periods of time. The evidence presented here is primarily in the form of discourse, using both illustrative excerpts and quantitative content analyses, rather than results from experimental design. While the methodological considerations of studies of classroom instruction and cognition are significant, the considerations of these issues are intended to invigorate the debate and foster progress toward a comprehensive theoretical framework for discussing symbolic and embodied perspectives on cognition as they occur naturally in instructional settings.

After presenting evidence that gestures enact symbols and provide grounding of novel and abstract ideas and representations, some broader themes of embodied cognition are considered. First, I consider these findings with respect to the distinctions made between ‘on-line’ and ‘off-line’ aspects of cognition, and their relation to principles of situated cognition, and symbolic off-loading of cognitive work. I then examine these points in relation to computer-based metaphors of cognition. Finally, I revisit the debate within the literature that the primary role of gesture is either for lexical access (chiefly serving the speaker) or communication (chiefly serving the audience; Alibali et al. 2000; Krauss 1998). I attempt to use principles of embodied cognition to integrate these two views under a new framework that acknowledges ways that internalized body-based behaviours from social interaction can mediate individual thought processes.

18.2 Use of signs in instructional gesture

Let \( X \) equal \( X \).

– Laurie Anderson (1982), from Big Science

In Chapter 1, Glenberg et al. orient us to the use of symbol as one of the types of signs in Peirce’s theory of semiotics.\(^1\) A symbol, in this view, is a sign that takes on meaning through its connection to other entities (other signs, or actual objects or events) via arbitrary rules or conventions. While not identical to its use in Newell’s (1990) physical symbol system hypothesis,\(^2\) symbols in this sense share many of the properties as they are commonly used in cognitive models. For example, Glenberg and Robertson (2000) note that within cognitive psychology, symbol meaning historically ‘arises from the syntactic

---

\(^1\) ‘... I had observed that the most frequently useful division of signs is by trichotomy into firstly likenesses, or, as I prefer to say, icons, which serve to represent their objects only in so far as they resemble them in themselves; secondly, indices, which represent their objects independently of any resemblance to them, only by virtue of real connections with them, and thirdly symbols, which represent their objects, independently alike of any resemblance or any real connection, because dispositions or factitious habits of their interpreters insure their being so understood.’ (Peirce 1909, p. 460–1).

\(^2\) This is in contrast to the embodied view of a perceptual symbol system, where perceptual symbols represent concepts that are contextualized, situated tokens (Barsalou, Chapter 13, this volume; but also see Barsalou 1999; Barsalou et al. 1993).
combination of abstract, amodal symbols that are arbitrarily related to what they signify’ (p. 379). Icons are distinguished from symbols in that they make this connection through resemblance, while an index connects to a sign through causal or spatiotemporal relations. Embodied cognition theorists, in contrast, argue that to take on meaning, symbols need to be ‘embodied and grounded in sensorimotor experience’ (Glenberg et al., Chapter 1, p. 4).

18.2.1 Language d’action

In this regard, consider the account of the creation of a sign presented by LeBaron and Streeck (2000). An instructor of a home-repair class teaches a lesson on the use of a scraper tool for adding and removing mortar compound to drywall. He picks up one scraper tool from among several on the counter before him, grasps it as if using it to scrape away excess compound, and performs a vertical up–down motion in midair to demonstrate its use. The actual drywall is not present, and neither is the compound. The instructor then repeats this action with a larger scraper, and then a much smaller one (a putty knife). A bit later, the instructor raises an empty hand shaped as though holding a scraper to show how it can be used to apply mortar compound to (an invisible) drywall surface.

In this example, a manual gesture serves a partly indexical, partly iconic, partly symbolic relation to a tangible event or object. Its indexical role is to the prior episode. Iconically, the instructor is using hand shape and arm motion to evoke the use of a tool. It is also symbolic, not in the sense that it is arbitrary, but because the gestural sign use and meaning transcends the specific event–object that led to its origin. The hand motion comes to be a type, referencing a class of tools that all conform to this hand shape and perform a similar function using rhythmic arm motions. Furthermore, for some of the novices in the class, the specifics of drywall and compound are unfamiliar, and the actions do not evoke their own, primary experiences of applying or removing mortar, but rather simply evoke the (ungrounded) conventions of the earlier demonstration. Indeed, the instructor’s gesture act is a sign in this discourse.

As Lebaron and Streeck (2000) state, ‘The formation of a symbol is a defining moment in the fabrication of shared knowledge because it allows the participants to focus upon and evoke previously shared experiences and to plan and conduct shared activities in their wake’ (p. 118). It is a case of language d’action, an abstraction of a specific act.

18.2.2 Catchment

One of the important attributes of the gesture used to signify the scraping tool is its recurrent shape and motion. McNeill and Duncan (2000) use the term catchment to describe those communicative events where distinct features of a gesture recur over the length of a discourse. The recurrence can be signalled by the form of the hand shape, its location, orientation, motion, rhythm, and so on. One interpretation of catchment draws directly on the idea that speech and gesture are mediated by a common set of language production processes (McNeill 1985). Following from this model is the claim that the recurrence of a common idea or image for the speaker is exhibited by the re-enactment of a particular gesture. From this, it follows that the re-enactment marks the recurrence
of the original image or thought. Thus, catchments convey a semantic association for the speaker, even when they are directed at different objects or in different contexts. It also follows that catchments can contribute to establishment of coherence for the listener.

Catchment behaviour is evident during instruction and, along with speech, can be regarded as ways that instructors attempt to provide continuity of meaning across representations (Alibali and Nathan 2005). As an example, consider the basketball problem as an entry-level problem in a sixth-grade early algebra unit.

The basketball problem
Consider the following: ‘Mr. Robinson and his four daughters want to buy a special, autographed basketball. Mr. Robinson’s daughters will each pay the same amount. Mr. Robinson will contribute $18 himself. If the ball costs $42, how much will each daughter pay?’

Over the course of a couple of days, students generated two mathematical accounts of the basketball problem that are recreated in Figure 18.1. The equation on the right side, the solution equation, was generated first. It was constructed by the students as a way to summarize their steps of executing an arithmetic solution strategy (Nathan and Koedinger 2000) called unwinding. Unwinding is highly effective across a range of tasks and grade levels, from middle school through college (Koedinger et al. in press; Koedinger and Nathan 2004; Nathan et al. 2002). During unwinding, the student runs the relations stated in the problem situation backwards and ‘unwinds’ them by performing the inverse arithmetic operations. But unwinding as a solution method has its drawbacks and can easily be thwarted. It also emphasizes a procedural approach of arithmetic reasoning at the expense of an algebraic approach that highlights the structural relations of the problem. Consequently, mathematics instructors want students to understand a more general and powerful method, direct modelling, where the situation of the problem stated above is represented directly by algebraic expressions that explicitly identify the quantitative relations between known and unknown values.

In a later lesson, as a way to introduce the situation equation that directly models the relations in the problem statement, the teacher had students act out a skit of the basketball problem (transcript 1):

Transcript 1: The basketball problem skit

Teacher: Okay? Let’s see it.

T: You got to talk really loudly; you’re on stage now. I need everybody, I need the audience to be really quiet.

<table>
<thead>
<tr>
<th>Situation Equation</th>
<th>Solution Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>amount each</td>
<td>amount each</td>
</tr>
<tr>
<td>$4 \times \text{daughter paid} + $18 = $42</td>
<td>\text{daughter paid}</td>
</tr>
</tbody>
</table>

**Fig. 18.1** Classroom whiteboard depicting the parallel structure of the situation (left-hand side) and solution equations.
T: This is a performance, please!

Daughters: Wow look at this basketball!

D:Cool! It’s $42 altogether, wow that’s a lot! Can you help pay for it?

T: (To actors:) Nope, now nice and loud now.

Father: Um, I’ll give eighteen dollars and you girls give the rest.

D: We need more dollars!

T: (To actors:) Okay, freeze!

T: (To class:) Have they described the situation so far if you have the numbers on the board?

T: Yep, they have, right? They mentioned $42; they asked Dad to help; Dad said he’d kick in $18; and they figured out how much left, left they’d have to pay.

T: (To actors:) OK. I’m going to let you finish off the problem even though we haven’t done it on the board yet. Go ahead and just act out the situation; now what are you going to do?

D: (Unintelligible).

T: Okay you gotta gotta be really loud.

D: How much are we each going to have to pay?

D: (The four daughters in unison:) Well 24 divided by 4 is 6.

D: (To other actors:) So we … (unintelligible). So all of us have to… (giggling)

T: Talk it over with Dad, maybe finish it off here …

D: (All four in unison:) We’re each gonna pay six dollars.

F: (Unintelligible).

T: Okay can you say that really loud again Matt, cos that was really good. (Unintelligible) listen over there.

F: Altogether plus the $18 dollars that’d be $42, so we can buy the… basketball

T: Give ‘em a hand.

Class: (Applause)

The teacher relied on students’ understanding of this relatively simple situation – made apparent through roleplay – to show that the situation equation is really a mathematical summary of the known and unknown quantities and their relations described in the story problem. It was not merely the retelling of the story, but the acting itself that helped establish the proper situation model and gave meaning to the situation equation (lines 4–7). For example, known and unknown information was differentiated through the use of numbers or letters, respectively (lines 8–10). Even operations were enacted. The students playing the four daughters grounded the multiplicative relation (4x) by announcing in unison how much each would pay (line 20). The physical separation of the daughters and the father is also consistent with the separation of the quantities that they each contribute to additive portion of the algebraic model. This enactment of the basketball problem
shows the value of body-based structures and spatial relations for thinking about analytical expressions that model the situation.

There is empirical evidence that this approach is effective for improving learning.3 This work on simulation relates to recent work by Glenberg and colleagues (Glenberg and Mehta, Chapter 2, this volume; Glenberg et al., in press) using real and imagined manipulation to enhance arithmetic story problem solving with elementary school children. That work frames the challenge for story-problem solving as performing proper indexing, and scaffolding the indexing so it actually happens. It shares the focus on establishing proper mapping. As Glenberg et al. hypothesize, ‘when the structure of the mental model can be easily mapped onto the mathematical operations, performance is dramatically improved.’ Learning gains with the ANIMATE system for solving algebra story-problems provides further support for this instructional approach. In that study (Nathan et al. 1992), those who learned to generate and manipulate algebraic expressions with reference to a running animation of the referent situation described by the story-problem showed greater test performance, more story inference making, and more self-correction than users in comparison groups with and without the simulations.

After establishing the two equations shown in Figure 18.1, the teacher invited the students to consider the similarities between the two symbolic representations in Figure 18.2. She set out to foster links between these two different symbolic representations, and to convey a mapping between the solution steps (solution equation) and the structural relations (situation equation) that she believes will help students to understand the meaning of each mathematical representation in terms of the other. To establish the links, the teacher uses gesture and speech within instructional communication. One thing to note is how the teacher tries to avoid the problem of symbol-to-symbol mapping that leaves the meaning of new symbols ungrounded (Glenberg and Robertson 2000; Harnad 1990).

Initially, students focused on the common elements: as originally written, they each had $42, $18, and a 4. The teacher used these observations to highlight how elements of the two representations map from one side of the board to the other even though their mathematical relations were different (see transcript 2).

Transcript 2: Relating the situation and solution equations (with indices to gestures in [i]).

Teacher: How are they different? I’d like to keep going on what G3 started to say; how they’re written. Tell me something more about how they’re written that’s different. Boy-1?

B1: They’re backwards.

T: What do you mean they’re backwards?

B1: Um, like 18 and 42 are in the end of that [0] equation.

3 Grounding abstract representations may be most helpful during the beginning stages of learning to work with formalisms because they also tend to model relatively simple mathematical relations. However, as the complexity of the relations increases, there may be greater advantages incurred when working with amodal representations (Koedinger et al., in press).
Having established that the two equations were made up of some of the same elements, the teacher asked students to consider what was different between the two equations (see transcript 2, line 1). This prompt elicited a far deeper awareness from students of how the representations were related. The teacher’s use of linking gestures (discussed in more detail below) surfaced when a student (line 2) described the relationship between the two equations as ‘backwards’. In response, the teacher probed further (line 3). The first boy (B1) referred to ‘that equation’ (line 4) and used a pointing gesture to provide a reference (i.e., the situation equation) for his deictic statement. However, the teacher provided a more specific reference (lines 6–7) as she walked over to the situation equation and slid her whole hand (see Figure 18.2a) to underscore the specific elements (the plus-18 and the equals-42) that B1 referred to. While B1 was still talking, the teacher then walked over to the other side of the board and performed the same hand and arm motion – a catchment (see Figure 18.2b) – under the difference expression ($42 – $18) of the solution equation (line 9). The act served more than an indexical role for the speaker’s words, it also established through the recurrent hand shape that a correspondence existed for the two expressions across the different representations. This conveyed a structural mapping across the two equations, much like that discussed in the analogical reasoning literature (e.g., Gentner 1983).

18.3 Grounding and linking

It seems implicit, but worth stating overtly, that in order to serve a grounding role worldly objects and events need to already have some familiarity or meaning assigned to them. In line with this, Lakoff and Nunez (2001, p. 49) argue that for a symbol to be understood it must be associated with ‘something meaningful in human cognition that is ultimately grounded in experience and created via neural mechanisms.’ With this in

---

4 These students have not previously discussed mathematical terms like ‘inverse relations’, and it is understandable that they would need to use colloquial terms like ‘backwards’ to try to describe what they see.
mind, it is worth surveying some instructional interactions to see how educators introduce symbols and help students assign meaning to them.

18.3.1 Grounding instruction

It is difficult to imagine the study of classroom learning and instruction – especially in secondary and college-level courses on mathematics and science – without the presence of external symbols (systems of notation) and representations. External representations – particularly formal representations such as variables, equations, graphs, and schematics – figure prominently in the classroom and in curricular materials. Much of the focus of mathematics and science education centres on manipulating, interpreting, and producing external representations (Laurillard 2001).

External representations also figure prominently in Peirce’s triadic theory of signs. ‘A sign, or representamen, is something which stands to somebody for something in some respect or capacity.’ (Peirce, c. 1897). This description of external representation clarifies its distinction from internal, mental representations, while providing its connection. As Peirce states, ‘It [the sign] addresses somebody, that is, creates in the mind of that person an equivalent sign, or perhaps a more developed sign. That sign which it creates I call the interpretant of the first sign.’ (Peirce, c. 1897).

Grounding as a connection or mapping is commonly inferred when studying communicative situations. Grounding is generally attributed through elaborated speech, the display and manipulations of objects and external representations, and through various forms of gesture, including pointing and representational gesticulation. Instructional settings are particularly rich with attempts to provide meaning for signs, since these are sensorially demanding contexts where new concepts and formal representations are often being introduced and applied.

---

**Fig. 18.2** The teacher used an arbitrary gesture to index elements in the situation equation in panel A with a student’s speech. She exhibited a catchment by repeating this same gesture for corresponding elements of the solution equation in panel B to establish their relationship across the representations.
My colleague, Martha Alibali, and I have used the term *grounding instruction* to highlight ways in which teachers exhibit grounding in the classroom (Alibali and Nathan 2005). One such instance of grounding instruction draws again from the basketball problem episode.

Transcript 3: Teacher grounds a student’s speech (with indexes to gestures in [i]).

**Teacher:** [10] Adding, subtracting, multiplying, dividing. How are some of those the same and some of those different?

**[Gesture-10: counts each operation off with her fingers]**

**T:** B3? You had your hand up. No? B1?

**B1:** Um, the timesing and dividing are… timesing was there and dividing's there… [11]

**[11: Teacher walks over from the right (solution equation) side of the board to the situation equation side and points to the multiplication symbol.]**

A student notes (without explicitly saying it) that the inverse operations of multiplication and division appear in corresponding locations across the two representations. Here, the teacher links B1’s reference to 'timesing' (transcript 3, lines 55–56) by walking over to the situation equation and pointing beneath the multiplication operator.

From a conversational standpoint, gesturing to elaborate another speaker’s speech is unusual, though my colleagues and I have seen this in other datasets of classroom discourse (Alibali and Nathan, in press; Nathan et al., in press). One reason the teacher might do this is to clarify what B1’s utterance actually meant – ‘timesing’ is not a standard mathematical term – while avoiding interrupting him. But the move is pedagogical and communicative, as well as indexical. It linked the word with the sign for the entire class, and was not merely a clarification request between interlocutors as is typically found within conversation analysis (e.g., Sacks et al. 1974). This move suggests the teacher’s desire to elicit students’ ideas to convey mathematical content, thereby engaging other students while still maintaining control over the mathematical meaning of that speech. In this way, linking can bring together the best that student and teacher have to offer.

The linking gesture has the potential to be a powerful pedagogical technique for grounding the meaning of unfamiliar symbols and representations. Linking gestures were evident in a detailed study of a middle school algebra classroom (Alibali and Nathan, in press). The focal lesson involved illustrations of pan balance scales, like that seen in Figure 18.3, with various combinations of objects of unknown weight. As the teacher described it, ‘We’re going to translate some of these… pans into equations.’ Three pan balances were presented in order: two that balanced (and led to equations; see Figure 18.3), and one that did not (and led, therefore, to an inequality).

After a class discussion of balance and weight, the teacher set out to show the parallel structure between the pan balance and algebraic sentences. In particular, she wished to highlight how balance related to equality (and to the equal sign), how lack of balance related to inequality (and to the greater-than sign), how collections of the same object related to scalar multiplication, how collections of different objects related to addition (and to the ‘+’ operator), and how the unknown weight of any object related
to the use of variables. Second, the teacher supported students' constructions of algebraic models of the pan balances. Third, she demonstrated how equations could be simplified through meaningful operations on the pan balance, such as cancelling out identical variables (objects) from opposite sides of the algebraic sentence (removing identical objects from both sides of the pan balance), and how variables (objects of unknown weight) could be substituted for one another.

In that study, the unit of analysis was the verbal utterance. Each utterance was coded as speech alone, or speech with writing, pointing, or representational gestures (a combination of iconic and metaphoric gestures as they are described by McNeill 1992). An utterance could be accompanied by any number of gestures. In addition, each gesture referent was coded as referencing either the pans (considered the most concrete referent), the algebraic symbols and equations written during the lesson, or the links the teacher made between the pans and symbols (the most abstract referent).

During the 23-minute lesson, 158 utterances were collected that addressed the instructional topic and occurred when the teacher was on camera. Some form of gesture accompaniment, mostly pointing, occurred in 74% of the teacher utterances. The teacher also exhibited linking behaviour. Through careful placement and timing of her gestures, the teacher established a connection between the physical point of balance and the equal sign. Normally, we would not expect the equal sign to be a novel symbol for middle school students, since this is generally introduced in early elementary grades during arithmetic lessons. However, the role of the equal sign in algebra is broader then when used in arithmetic. Specifically, algebraic equations use the equal sign to convey an equality relation between two expressions (i.e., the two sides of the equation), and this relation must hold even when the value of these expressions may be unknown. In contrast, the arithmetic meaning of the equal sign indicates that one side of an equation produces the value given in the other side when the rules of computation are properly followed.

How did the teacher set out to convey this alternative interpretation of the equal sign symbol? First, she used pointing, indexing the fulcrums as ‘wedges’. In speech she made the link that the wedges will operate ‘as the equals sign.’ She then wrote an equal sign directly under, and in close proximity to a wedge, setting up the relation by location and establishing the line of symmetry defined by the pan balance as also serving as the line of symmetry for the forthcoming equation.
I want to provisionally claim that the teacher intended these accompanying actions to serve as a form of grounding for her students. That is, I assume that the link she established by drawing the equal sign in that way was meant to be interpreted by students as connecting the mathematical idea of the equality relation to the physical concept of balance. The scale was chosen because it was believed to be familiar and meaningful to the students (an assumption that was checked earlier in the lesson), and it physically embodied the notion of an equality relation (as opposed to producing an answer) through physical balance of equal weight (actually torque) (cf. Lakoff and Nunez 2001).

Grounding instruction of this sort varied with the instructional referent during the lesson. Verbal utterances by the teacher that made links between the equations and the pans was more likely to be accompanied by some form of grounding, through gesture and writing, than were utterances that referred exclusively to algebraic equations or pans: $\chi^2(1, 158) = 4.13, p < 0.05$. The number of grounding acts per utterance also differed significantly across referents, $F(2, 155) = 3.60, p < 0.03$. The teacher used significantly more grounding acts per utterance when talking about pan–equation links than when talking about pans or algebraic equations by themselves.

The teacher’s use of grounding instruction also changed over the course of the lesson. The mean number of grounding acts per utterance was high for the initial pan balance equation (mean = 0.78, standard error = 0.08). It then decreased during the second problem of the same type (0.40, 0.07), but increased again for the third, novel problem, which extended the principle to an inequality (0.67, 0.08): $F(2, 260) = 5.35, p < 0.01$. This pattern suggests that the teacher used gestures most often when she introduced materials that she expected to be new and unfamiliar (i.e., on problems 1 and 3) and this practice faded when she believed that the material had become familiar to students. Thus, there is support for the claim that the teacher’s use of gestures was to serve a grounding role during instruction.

Lastly, Alibali and Nathan (in press) found that the frequency of grounding instruction was influenced by student demands. The teacher was more likely to gesture after a student utterance, such as a question, than before (56% vs. 31%). Further, the mean number of gestures per utterance for the teacher was significantly greater immediately following student utterances than immediately preceding them (0.69 vs. 0.33, $t(35) = 3.17, p < 0.01$).

It appears that student needs, in the form of questions or exposure to new and abstract material, trigger more gesturing on the part of the teacher. This is noteworthy, since there is empirical evidence that gesture use by teachers can improve student learning and subsequent gesture production in children (Cook and Goldin-Meadow 2006). We now consider the findings reported above in terms of the theory of embodied cognition, and explore their implications for understanding cognitive behaviour in complex settings.

18.4 Discussion

My aim for this chapter was to show that embodied cognition can inform research on gesture, and that findings from gesture research can contribute to embodied theories of cognition. The findings on instructional gesture presented here included the investigation
of cognitive behaviour *in situ* and considered the influences of social and environmental interactions on that behaviour. We saw that instructors can use the physical and dynamic forms that gestures take to enact symbols and to reinstate them (using catchments) at different points in the discourse. Furthermore, those enactments can be transported across contexts and they can denote a class of entities (e.g., scraper tools). In addition to observing examples, classroom data were used to show when gesture served a grounding function. For example, we saw that the referents of a lesson that were more abstract and less familiar were more likely to be the subject of gestures, particularly gestures we expect to serve a grounding role that, presumably, help students understand the lesson and foster meaning making. We also saw that the teacher’s use of gesture was responsive to student needs and occurred more frequently after students made demands on the teacher. Now we re-examine these findings of the nature and role of gesture in terms of contemporary theory of embodied cognition.

In reviewing the central tenets and claims of embodied cognition, Wilson (2002) makes a critical distinction between ‘on-line’ and ‘off-line’ aspects of cognition. On-line behaviour includes those forms of activity that are rooted in the environment within which it occurs, so that cognitive behaviour is affected by situational and contextual factors such as real-time demands, spatial constraints, and feature-based affordances (e.g., Goldstone et al., Chapter 16, this volume; Kirsch and Maglio 1994). During on-line cognitive behaviour, the mind helps navigate through the world, both metaphorically and literally. Because of this, Wilson (2002) notes, ‘the mind can be seen as operating to serve the needs of a body’ (p. 635). On-line aspects of cognition are very important for learning, as well as situated performance, and constitute one road to the development of artificially intelligent agents that learn to engage in purposeful interactions with their environment and other agents (Roy, Chapter 11; Steels, Chapter 12).

Still, Wilson suggests exercising caution so as not to overstate the situated nature of cognition. While much of what people do is on-line, she argues that a signature aspect of cognition is the ability to operate off-line, drawing on mental representations of the world that are removed from the events and things that are the object of attention, such as planning, recall, and running mental simulations. From an embodiment perspective, off-line processes may utilize purely mental processes and structures to reason about the world, but to do so they engage the same sensorimotor processes involved in situated cognition and action, when the objects of cognition are distal or even imagined. In this case, as Wilson (2002) notes, the body now serves the mind.

By Wilson’s account, off-line behaviour ‘is by definition not situated.’ (p. 626). However, Wilson applies a narrow view of situated behaviour, one that is not universally accepted by cognitive scientists (e.g., Greeno et al. 1996). Obviously, all cognitive processing occurs in some situational context. An alternative to Wilson’s view that off-line processing is not situated is to distinguish between the referent situation that is being modelled (but not indeed present) by the off-line processes, from the new setting within which off-line reasoning is taking place. The climate and available associations of the immediate context, even during so-called ‘off-line processing,’ can be shown to influence one’s reasoning. One unfortunate but apparently quite robust example of this is when
perceptions of stereotype threat impede the intellectual performance of under-represented minorities when they are in mixed-race or mixed-gender testing situations (Inzlicht and Ben-Zeev 2000; Steele 1997; Steele and Aronson 1995). In these cases, members of stereotyped groups may, when their minority status is primed by the current situation, invoke other, threatening settings that serve to hamper their performance. Furthermore, the referent (but distal) situation can also influence off-line cognition. Consider the case of planning: it is rare indeed for a plan to be completely specified. Contingency planning (Suchman 1987), which is highly responsive to the immediate context, is commonplace enough – often planned for, explicitly, in marking out the specified and unspecified portions of a plan – as to be considered a fundamental part of planning. One consideration is that there may, in fact, be no off-line cognition at all. Through enactment, we are really on-line, even when removed from the situation.

The select research on gesture reviewed here also showed support for the claim that some of the cognitive work involved in instruction is off-loaded onto the environment and makes direct use of the environment to support cognitive activity (see Steels, Chapter 12, this volume; Schwartz and Black 1999). Symbolic off-loading, in particular, can help mitigate the inherent limitations of memory and attention, and also allow agents in the environment to ground the meaning of novel and abstract ideas and representations. In this vein, we saw how teacher use of linking gestures made explicit connections between novel and familiar entities. For example, the novel relational interpretation of the equal sign was connected to the fulcrum of a balance scale. We also observed how the introduction of an algebraic use of letters was made by connecting the letters to the weights of the objects they were intended to represent. As a final example of off-loading, we saw how gestures provided explicit links between representations. One particularly elaborate way this was shown was in the analogical mapping constructed by a teacher to establish an element-by-element correspondence between two equations representing different aspects of an algebraic characterization of a problem situation. Computer animations and simulations also provide this kind support and appear to be a natural way for students to learn the meaning of formal representations (Goldstone et al., Chapter 16, this volume; also see Nathan et al. 1992). In the language of embodied cognition, one can say that these communicative forms are attempts to ground these new ideas and symbols, in that they establish connections that potentially service comprehension and inference-making processes.

The prevalence of gesture use during instruction, and communication more generally, raises an issue fundamental to the validity of computer-based metaphors of intellectual behaviour, as depicted in classical information processing theory (Newell 1990). In its canonical form, an information processor sits at the core of the cognitive architecture. Outside of this ‘central processing unit’ (CPU) are ‘peripherals’ – sensory inputs that provide information from the perceptual systems, and actuators that are output systems for acting upon the world (Eisenberg 2002; Wilson 2002). But what about the case when a hand gesture is the primary goal of a speech act, as when a teacher wants to make a linking gesture between a familiar object and a new representation? Data on teaching shows that the use of gesture is tightly woven to teachers’ language production. Its role in
this complex cognitive activity is not peripheral, but may in fact be central to the pedagogical and communicative goals of the instructor. By drawing on the metaphor of mind-as-CPU (and body-as-peripheral), the cognitive science community marginalizes the influences of body-based behaviours, and unnecessarily demarcates what is and is not legitimate phenomena within the study of cognition.

The important role of gesture during instruction, and communicative settings more broadly (McNeill 1992), highlights its importance as a visual component of speech production in socially mediated settings. For example, speakers gesture often in communicative settings, and do so more when their listeners can see them (Alibali and Don 2001; Alibali et al. 2001). Yet gesture also occurs in settings where speakers are not in visual contact, such as when people are speaking on the phone or otherwise out of visual range, and when the speakers are congenitally blind and have never seen gesturing modelled (Iverson and Goldin-Meadow 1997). For these reasons, some have suggested that gesture is not primarily intended to serve the audience, but rather to aid the speaker with lexical access, particularly when the utterances deal with spatial phenomena (Krauss 1998). Both competing claims about the primary role of gesture can be seen as compatible with embodied cognition, particularly that cognition is off-loaded onto the environment, and that even off-line cognition co-opts sensorimotor processes and structures that operate during situated action. However, it is difficult to reconcile them with one another (Alibali et al. 2000). In these closing remarks, I consider how a broader treatment of embodiment that includes the social as well as the physical environment can point the way to an overarching framework for gesture production that addresses both communicative and non-communicative functions.

There is growing awareness of the social nature of human thought and learning (e.g., Greeno et al. 1996). One avenue recently explored is how gesture serves to create a common frame of reference or common ground among speakers (Nathan and Alibali 2006). It has long been acknowledged that the establishment of common ground – what many scholars refer to as intersubjectivity – is an essential aspect of effective communication, and that speakers devote a great deal of effort and time ensuring this at a discourse level, even though they may fail to agree on the specific matter at hand (for a review, see Nathan et al., in press).

I propose that communicative and noncommunicative functions of gesture can be brought under one theoretical frame that addresses intersubjectivity in both individual and socially situated settings. Contemporary cognitive neuroscience is exploring basic interpersonal processes such as imitation, empathy, and the ability to impute the intentions of others – all behaviours that hinge on intersubjectivity – in an effort to understand fundamental mechanisms of both individual and social behaviour. During discourse, speakers learn that establishment of intersubjectivity is an effective way to further their social aims during conversations. Researchers studying the behaviour of mirror neurons (e.g., Rizzolati et al., 2001) hypothesize that these neural circuits permit people to directly understand the actions of others because these actions evoke in us the bodily states that we would normally occupy if we initiated those same actions. Mirror neurons are specially evolved and selected areas of brain circuitry ‘that allow us
to appreciate, experience, and understand the actions we observe, the emotions and the sensations we take others to experience’ (Gallese 2003a, p. 525), by constituting them in intersubjective relation to our own actions and feelings. From this empathetic response, speakers constitute a shared manifold of intersubjectivity with those with whom they interact (Gallese 2003b). The use of gesture and other body-based behaviours add to the establishment of this shared manifold to the degree that they invoke similar sensations in the listener/observer. This allows interlocutors to provisionally enter into a shared social space, even though their ideas and interpretations may differ from one another.

Through social interaction, and by observing the social interactions of others, I hypothesize that interlocutors internalize the social and cognitive behaviours that lead to effective discourse. This includes, in no small measure, gestures and other actions that facilitate the appropriate selection of words so that the shared manifold is maintained. As proposed within sociocultural theory (e.g., Vygotsky 1978; Wertsch 1979) these interpersonal interactions help to determine the development of intrapersonal (i.e., individual) processes. The body-based behaviours that were so effective in socially mediated thinking are co-opted to fulfil the goals of individual thinking, such as planning and lexical access. Those gestures that foster intersubjectivity among speakers also assist us in articulating our own thoughts to ourselves (and others) in a clear manner. In this way, we are drawing on socially mediated processes such as gesture even when we are seemingly ‘alone with our thoughts.’

**Debate**

**Friedemann Pulvermüller:** Thank you very much for your interesting talk. It seems to be clear that gestures somehow help in understanding an explanation. However, it could be at different levels. There is no doubt that gestures help, for example, in structuring the discourse by making emphasis – this is the word I want to attend to. Or I could make an index using a hand gesture that says that something needs to be kept in mind, in the case of understanding a formula. So certain processes can be symbolized or indexed, but on the other hand explaining or symbolizing addition and multiplication might be extremely difficult or impossible. So can you maybe comment on that?

**Mitchell Nathan:** It’s true, I don’t think a teacher would only want to use gesture as their only source of instruction or gesture even in speech. I didn’t show a video of this skit but there’s a really nice part here where the four daughters all say in unison, ‘We need to contribute our money to buy this basketball.’ And there’s sort of this element of oh wow, that’s like multiplication, you know. Here’s these four girls at the same time, conveying this idea. As a teacher you wouldn’t by any means want to narrow down your repertoire of how to convey the range of ideas, even just in mathematics, and gesture isn’t going to be suitable for everything. I agree. What’s remarkable is how pervasive it is, and it’s also a little bit of a concern how kind of under the radar it’s gone for so long. For instance, we do very little in teacher education around what teachers should gesture. We don’t talk about it at all – I don’t really know of any
programs that do talk about it. However, it’s clear that it could be used to one’s advantage tremendously.

**Walter Kintsch:** Should we spank the students when they invent something as intelligent as the ‘unwinding’ strategy? Because you really don’t want them to do this. I mean, it works, but to scale up you want them to use the equation.

**Nathan:** A big debate that I think that happens in education, and certainly relevant to algebra instruction, is where you want people to end up versus how you want them to get there. So I think we have a lot of agreement about where we want everyone to end up. But I think it’s pretty clear that if we try to do it directly by teaching these formalisms, they don’t get there. So, for instance, think of the teacher who believes that symbolic reasoning precedes story-problem solving, and reserves story problems until students reach a certain level of equation manipulation. Well, some of those students are never gonna get across that threshold. They’re never going to get to exercise the reasoning they may have for solving a story problem, because the teacher feels like they’re sparing this student from this onerous task that they’ll never be able to accomplish because they have a developmental model that puts story-problem solving as farther out than equation solving. But in fact, we have some developmental data that compares the developmental models, symbol precedence (i.e., symbol first) versus verbal precedence. And it shows that the verbal precedence is a far better fit for our data across two datasets. So it would be a mistake for a teacher to withhold story-problem solving. They’re not doing the student a favour. In fact, what we show in the intervention studies is that you can leverage students’ reasoning on these more accessible verbal problems and get them to do the symbolic representation, and even get them to solve nonlinear problems in 7th and 8th grade, which often they don’t get to do anyway. So I think there’s a lot of promise for this method. I understand your intuition, and I think we should always be careful about it, but I think that as you look at the developmental trajectory issue, you start to realize that we need to be looking at how to move people. And that’s why this progressive formulization is such an appealing idea.

**Antoni Gomila:** My first comment is that it looks like what is more important for algebra teaching is probably reading comprehension. Language understanding through texts, maybe even more than gestures, I don’t know. The other thing is that your talk nicely illustrates the point that Wertheimer already made in his classical productive thinking theory, explaining that to understand is to know how to proceed, in a given context. The context can be very different, and not just to repeat the same situation again and again to apply the same equation again and again, but to be able to codify a representation of a problem in more abstract terms.

**Nathan:** Yeah, I think that’s a great connection, thank you. That’s nice.

**Francis Quek:** Ah, just a couple of things. Number one, you showed that when teachers do the equation the first time there are more gestures, and then the second time the number of gestures drop. And you started by talking about catchments that tend to hold certain features to show that there’s continuity of ideas. Did you see that?
Did you see that although the number of gestures declined, the form of the gesture was the same so the student could be cued in that it’s the same problem?

Nathan: I can’t say that I’ve systematically looked at it across the entire dataset. But I can tell you we see lots of examples of this catchment phenomenon. And what’s interesting to me about catchments – I don’t know if Duncan and McNeill talk about it quite this way – but there’s a way in which a repeated hand form that’s meant to convey the same idea across different settings is like a symbol, Right? And one of my favourite examples is Streeck and LeBaron’s work with a vocational education teacher showing students how to use a scraper for a drywall, like putting the compound on the wall. He picks up the tool and introduces this tool idea to a room that’s basically novices, but then later just the hand form, so there’s no tool in his hand anymore; the hand form is used to reinvoke the idea of the tool. And then it gets used to reinvoke other tools that have the same function but are not exactly the same tool. So, like, a spackling device, which is not so different than a drywall compound tool, also gets referenced by this hand form. I think it’s an incredibly powerful idea. So we do see it, but we haven’t had a cause to systematically pull it out of that dataset. Unless, you know, you have some hypotheses we could talk about.

Quek: One last question: I just wanted to introduce this so we might discuss this later. We have talked a lot about pointing and deixis, but pointing at a reference involves three things, right? There is the thing pointed to, the field of reference, and the point of view of reference, which is actually very odd for instruction, whether I’m looking from here or from there. And good teachers would actually shift their shoulders – Randi Engle from Stanford has been working on this – they would actually give a different viewpoint just by subtly pointing at the same thing but from a different position. So if you’re looking at something from a viewpoint of inequality as opposed to an equality, the teacher might do this change of position, and that cues a student in to know that something new is happening. And, of course, the third aspect of pointing action is how you point. So the question I have, given that there’s so much pointing in the work that you see, have you looked at not so much the field of reference but source of reference. That might actually be more telling.

Nathan: We’ve looked at how pointing is involved in linking. I’m not sure if that’s exactly the same thing you’re asking for. We have looked at how it’s used to link across two representations, say an equation and, like, an equation on one side and an inverse of that equation on the other side. Martha Alibali has some nice data on the formula for area and the parts of a shape of the object, essentially to dereference the variables in the area formula. Is that what you’re asking?

Quek: Randi Engle did this. She showed teachers explaining a lock mechanism to a student. And, of course, when explaining a lock mechanism you actually look at the lock from the top, from a viewpoint of the tumblers of the key. And she showed that teachers would consistently shift positions. The teacher would consistently use different viewpoints to cue the students in, and that’s one of the hardest things to do in
teaching, to explain the changes in abstraction and changes in viewpoint. And this seems to be much better cued in gesture than in speech.

Nathan: I’m not sure how much cause that would come up for inscriptions like equations where the viewpoint probably doesn’t change that much from teacher to student. So, because she’s looking at three-dimensional objects and kind of moveable systems, it seems that it’s more relevant.

Quck: Thank you very much.

Author note

I would like to thank my colleague Martha Wagner Alibali for her extensive contributions to this program of research. I also wish to thank members of the STAAR research collaborative. In addition, David Havas was very generous with his time and feedback during our frequent conversations on these issues and reviews of an earlier draft of this chapter. I thank Rob Goldstone for his thoughtful review of this work. Finally, my warmest regards to Amy French for her openness to the scrutiny that comes with participating in classroom research. Preparation of this paper was funded in part by funds from the University of Wisconsin School of Education. This research was funded through a grant entitled ‘Understanding and Cultivating the Transition from Arithmetic to Algebraic Reasoning,’ by the Interagency Educational Research Initiative, an alliance of the National Science Foundation, the US Department of Education’s Institute of Educational Sciences, and the National Institute of Child Health and Human Development within the National Institutes of Health.

References


