Children and Parents Using Coordinated Multimodal Meaning Making During a Robot Coding Activity

Brian M. Andrus, Naomi Polinsky, Stephanie McCarty, Anjelique Bomar, Paige Smyth, David Uttal, Michael Horn, brianandrus2023@u.northwestern.edu, naomipolinsky2022@u.northwestern.edu, stephaniemccarty2021@u.northwestern.edu, anjeliquebomar2021@u.northwestern.edu, paigesmyth2021@u.northwestern.edu, duttal@northwestern.edu, michael-horn@northwestern.edu
Northwestern University

Abstract: Codable robots are now popular toys for families to play with at home. However, during play with codable robots, children may confront novel mathematical and spatial concepts that are central to achieving their coding goals. This work focuses on how parents use multiple modalities in representations with their children to elucidate and make meaning about these concepts. We present three episodes from a corpus of video recordings of parents and children completing a playful coding activity together in their homes. Our analysis reveals how parents support children during meaning making by strategically coordinating multiple modalities of communication, which become more concrete throughout the episodes. These concrete modalities, including physical artifacts and embodiment, support children’s understanding of spatial and mathematical concepts, which facilitates their continued play with the robot.

Introduction
Recently, the popularity of codable robot toys in homes has grown exponentially. Codable robots are small robotic toys that children control with programming apps (Bers, 2010). These toys are designed to engage computational thinking skills, such as problem scoping, iterating and debugging, and computational thinking concepts, such as sequencing and logic (Brennan & Resnick, 2012). To date, research on children’s play and learning with codable robots has primarily been conducted within formal learning environments (Ohland et al., 2019). However, there are many attributes of the home that are not available in schools which may contribute positively to children’s learning with these toys. In particular, the availability of parents may shape children’s thinking and learning (Rogoff, 2003). Thus, we examine parent–child interactions during play with a codable robot at home and ask how pairs communicate and make meaning about computational concepts that are central to their coding goals.

 Codable robots may be ideal for children’s learning at home because they afford constructionist principles like learning-through-doing (Papert, 1980). During play with codable robots, children practice generating procedures by setting their own goals for their robot’s behaviors. The child must engage in problem solving to translate these goals into logical computational sequences. Because the coding programs associated with the robots use block-based codes, children are able to prioritize their attention on arranging commands selected from a list, rather than on syntactic elements like punctuation. As a result, children can more readily test and iterate their code. During testing, children receive visual feedback that connects their block-based codes to tangible objects and observable events, revealing errors and inspiring reflection and iteration (Bers, 2010; Papert, 1980). Essentially, the clarity of the block-based codes paired with the tangibility of the robot creates an accessible way for children to practice computational processes.

However, knowing how to program their robot to may require engagement with spatial and mathematical concepts. For example, children often need to insert numerical values into their codes to indicate the distance, speed, and direction at which their robot should travel. Despite the tangibility of the robot, children may be challenged to connect these spatial and mathematical concepts to how their robot moves, resulting in impasses in children’s coding process. Yet due to the possibilities for fun and play involved in making a robot move, children may feel particularly motivated to overcome these impasses (Papert, 1980). Nonetheless, children may need additional support to apply spatial and mathematical concepts to continue making progress toward their goals. At home, parents may work with their children to overcome these impasses together, even though parents may be new to coding or may not have established pedagogical tools for supporting their children during coding activities. Thus, we ask how parents and children make meaning about spatial and mathematical concepts when working to overcome impasses during a coding activity. To answer this question, we virtually observed parents and children completing a playful coding activity with a robot together in their homes. This activity presented parent–child pairs with spatial and mathematical concepts that fell within children’s zone of proximal development (Vygotsky, 1978), such that children could make meaning about these concepts with the help of a parent.
Framework

We draw from work that considers a sociocultural approach to learning (e.g., Rogoff, 2003; Vygotsky, 1978), which emphasizes that knowledge is co-constructed through social and conversational exchanges. Each person in a social context contributes to the learning through their unique knowledge, skills, and repertoires of practice. For children, the social context of the home environment includes the availability and contributions of parents, who provide a critical social context for children’s meaning making and knowledge construction. Parents contribute to meaning making by taking on a guiding role, involving asking questions and directing children’s attention (Callanan & Oaks, 1992; Eberbach & Crowley, 2017). During learning interactions, parents closely monitor and assist with their children’s understanding and progress with a task by interpreting their verbal and nonverbal cues while attending to subtleties of the interaction (Goodwin, 2018). During play with codable robots, this monitoring behavior may allow parents to provide support to children in moments of impasse.

To provide the support necessary to overcome coding impasses, parents may call on the material artifacts and tools in their homes. In problem solving situations, such as coding a robot, cognition can be distributed among available material artifacts in the environment. People interact with these artifacts in social ways through embodied movements and actions (Hall & Stevens, 2015; Hutchins, 1995). Moreover, material artifacts support collective thinking by facilitating the construction of a shared visual representation, which promotes intersubjectivity and a common understanding of the relevant concepts (Stevens & Hall, 1998). A potential barrier to achieving intersubjectivity with material artifacts is that these objects and representations may be cognitively opaque cultural forms, such as complex tools (Goodwin, 2018). When these barriers arise, more experienced learners can support less experienced learners in using and understanding these forms to achieve intersubjectivity.

In addition to material artifacts, parents use a variety of strategies, many of which can be very subtle, to attune to their child’s understanding and find alternate means of explaining concepts. These strategies are often embedded in various forms of communication, such as talk, gestures, the use of physical artifacts, and subtle bodily movements to direct children’s attention (Tulbert & Goodwin, 2011), and could easily be overlooked. Studies of multimodality (Kress, 2010) inform our work because they help us tend to the affordances and functional specialism of these various modalities of communication and provide additional clues to why parents and children rely on different modalities at different times and for different purposes to convey meaning. Often, modalities are layered and used simultaneously or in contrast to nuance meaning (Kress, 2010). Given that children’s meaning making about mathematical and spatial concepts can be difficult to elicit verbally in the modalities are layered and used simultaneously or in contrast to nuance meaning (Kress, 2010). Given that children’s meaning making about mathematical and spatial concepts can be difficult to elicit verbally, more fine-grained and systematic approaches to understanding how learning happens interactionally.

Methods

We delivered a materials kit, containing a Sphero Bolt robot, two tablets (one for communicating with researchers and recording the process, and another for controlling the robot), and a felt rug with a map of an imaginary town, to families’ homes. During the study session, we guided families through the setup of materials, led a short tutorial on how to use the Sphero, and facilitated participation in a thirty-five-minute coding task. This task required pairs to “take their robot on a trip” to different locations on the floor map. Researchers turned off their cameras and microphones and took field notes throughout the activity. The tablet used to control the robot captured the pairs’ on-screen coding activities via a screen recording app. The second tablet was mounted on a tablet stand to record families’ interactions and movement through Zoom’s recording feature. For analysis purposes, the videos from both recordings were synced, enabling simultaneous viewing of the families’ activities, interactions, and code.

Basing our analysis on traditions of interaction analysis work (e.g., Erikson, 2006; Jordan & Henderson, 1995) we began by taking field notes during observations of the pair’s interactions and continued as we re-watched the video recordings alongside the pair’s on-screen code inputs. We grounded our observations and notes in a variety of modalities of interaction, which meant tending to details like changes in the pace of the activity, shifts in bodies, and evolving expressions of affect as indicators of important moments. In this way, we located key moments where children relied on assistance from their parent to overcome a challenge in their computational tasks. Using these aforementioned cues, we chose episodes that we saw as beginning with initial moments of impasse in achieving the activity objectives and concluding with moments of resolution. Then we generated multimodal transcripts to document content and characteristics of talk, gesture, gaze, and manipulation of material artifacts and the on-screen code. We relied on varied transcription styles allowing us to tend to and report on the information that we observed was relevant to the participants during their interaction. Through careful attention to these transcripts during interaction analysis work, our research team documented and shaped our analysis by considering the sequence and context of each utterance and action, and how the interaction progressed towards a
solution. This method of collaboration allowed us to draw on interdisciplinary expertise from the areas of learning sciences, developmental psychology, and computer science as we engaged in the process of discovering meaning making in the interaction. We see this spiraling process of data analysis (Creswell & Poth, 2016) as part of our foundation for identifying and refining the foci of our ongoing work, and also as our effort to see and notice meaning making happening from the perspectives of both the parents and the children (Hall & Stevens, 2015).

Findings
The following episodes show parents and children co-constructing visual representations and coordinating their interactions around those representations to understand mathematical and spatial concepts related to their task. Our findings provide detailed descriptions of how and why this process helps the pair to overcome moments of impasse. Similarities among these episodes reveal patterns in the interactions and demonstrate ways that parents and children use everyday meaning making to solve problems. The differences demonstrate the adaptability of this process in response to the nature of the concepts the pair encounter and the resources available to them.

Episode 1: Co-creating a computational plan using a durable representation
This segment of interaction follows Landon and his mother as they plan the trip their robot will take on the map (Figure 1a). The pair agree their robot will visit six non-adjacent destinations. These criteria play a role in how they choose their destinations while planning. Landon begins by communicating that his plan involves moving the robot from a house to the swimming pool, and then to school, but he soon demonstrates difficulty by starting over several times. As he recites this plan, his mother gazes at each location on the map and responds to him verbally. This creates a back-and-forth flow of talk consisting of his supplied information, her response, and their changing tones (see Table 1). At the moments when his mother is following along, she provides verbal affirmation (e.g., lines 2 and 6). When she is not following along, she pauses their flow (lines 4 and 18) with questions meant to clarify the locations to which Landon is referring. At these moments, Landon responds by recruiting gesture (Figure 1a), which he eventually adds to his communication of every destination. In line 10 the mother’s pause serves to draw Landon’s attention towards the adjacency of his two sequential stops, which violates one of their established criteria. On Landon’s third attempt, his mother halts their sequence of talk and shifts her orientation away from their activity on the map by sitting up and looking around the room while looking around for tape and a marker to label their routes (line 24). She recruits material artifacts for co-constructing a visual representation to help keep track of their plan, allowing the pair to proceed with more certainty (Figure 1b).

Figure 1. Landon gestures to refer to stops on his route (a). Recreation of Landon’s final route (b).

Table 1. Transcript of Landon and his mother discussing the procedure for their code.

<table>
<thead>
<tr>
<th>Line</th>
<th>Talk</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Landon: Right. So the ↑swimming pool.</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Parent: Yeaa↑ah</td>
<td>moves gaze to swimming pool</td>
</tr>
<tr>
<td>3</td>
<td>Landon: To school.</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Parent: (?) Where's school?</td>
<td>searches the map with gaze</td>
</tr>
<tr>
<td>5</td>
<td>Landon: Hhh. Um, so actually- so- to our ↑house. to ↑school</td>
<td>points to each spot as he names them</td>
</tr>
<tr>
<td>6</td>
<td>Parent: Okay.</td>
<td></td>
</tr>
</tbody>
</table>
Landon: To the swimming pool
Parent: Right.
Landon: And then-
Parent: Oh wait. But they can't be right next to one another.
Landon: °Oh° ((grunted))
Parent: ↑It's okay.
Landon: So, let's go to our ↑house.
Parent: Yea
Landon: To the. ↓School
Parent: Right
Landon: To Craig and ↑Chris' house.
Parent: Which is?
Landon: Right here
Parent: Okaaaay
Landon: To. The Circus? (. ) Which is right here
Parent: Mhmm.
Landon: And toooo ↓the swimming pool.
Parent: Okay, so you know what we should do?

Note. In this first episode, we use a modified version of Jefferson’s (2004) transcription notation to emphasize features of talk. Arrows represent rising or falling pitch, dashes mark cut-offs, degree signs enclose quieter talk, and a period in parentheses marks a short pause.

At a phase in their activity when the pair had to establish and agree on specific plans, the meaning that was communicated between them to achieve intersubjectivity happened interactionally and was built step-by-step with each utterance and action. Over time, pairs successively layered modalities of communication to offer clarity and stabilize understanding of their shared plan (Stevens & Hall 1998). Accordingly, the pair incorporated more physical and durable representations as time progressed, which leveraged their respective affordances towards more effective descriptions. They started with talk, making use of specific organizational features of their exchange like sequenced turn-taking and patterns of shifting intonation (Goodwin & Heritage, 1990). These subtleties allowed them to specify points where their understanding needed alignment for continued cooperation. To further align this understanding, they next relied on gestures (pointing) and then material objects (tape). This material numbering system had several affordances allowing them to carry out their task effectively. Their placement helped visualize the sequence of stops with a durable representation, yet their placement was malleable, which reserved room for changes to the plan. The labels also highlighted the spatial relationships necessary to confirm that they were aligning to the non-adjacent stop criteria.

Background information for the next two episodes
In the following episodes, parents and children are working on their code and have set a critical variable for navigating their robot called a reference heading. Thereafter, heading commands that incorporate numerical heading values tells the robot which direction to move. Initially setting a reference heading allows the child to arbitrarily decide a direction at which the robot is facing forward, or at zero degrees (0°). All subsequent heading choices within these individual codes are in relation to this global reference direction. For example, the code Spin asks children to input a heading. If the child inputs 90°, the robot will turn ninety degrees with respect to the forward-facing direction chosen for the reference angle. Setting headings is central to the following episodes.
Episode 2: Employing embodiment to make sense of headings

Alan and his mother have set their reference heading to the direction facing away from the wall and are troubleshooting the headings needed to move their robot from the police station to the train station (Figure 2a). After a few missed attempts, Alan’s mother attempts to clarify by tracing the desired route by pointing from “the police station” (Figure 2b) to “the train station” (Figure 2c). However, Alan’s gaze is instead fixated on the tablet as he inputs a new heading for another iteration, which does not go as planned. His mother now suggests that he sits against the wall while physically moving him to align his body’s forward-facing position to match the direction they set as zero before she tries to explain again (Figure 2d). After another attempt from his position against the wall, Alan’s mother leans over, again demonstrating with her hand the direction the robot needs to go. However, at this moment, her speech and physical orientation are not aligned with Alan’s perspective (Figure 2e), and his next result is still off. Then, Alan’s mother demonstrates a reflexive awareness of how she can respond differently to her son’s understanding by recruiting a new modality of communication (Goodwin, 2000), insisting that they stand up together to embody the robot in the same position on the map. Facing their reference heading while looking together at the tablet (Figure 2f), she shows Alan with her arm that the angle is behind and to the right. With this embodied enactment, he finally makes sense of the values and inputs 116°, a heading that works.

The tangible representations of a walkable map, along with the observable and physical outcome of the code enacted on it, allowed the pair to share knowledge and grapple with spatial and mathematical relationships of angles and headings (Horn, 2018; Ramey & Uttal, 2017). At key moments, Alan’s mother led their meaning making through her multimodal communication, facilitating Alan’s exploration of spatial relationships. By attuning to Alan’s understanding and responding reflexively, she situated elements of their code in the material environment around them, providing rapid just-in-time feedback through both material and social mechanisms, which has been shown to support learning (Berland, 2016). Notably, she does so with an embodied relatedness that parents often possess (e.g., sitting closely to share visual perspectives, moving her son with care, and standing directly behind him while looking over his shoulder), which she employs to help choreograph his attention to the spatial relationships relevant for completing his task (Tulbert & Goodwin, 2011).
Episode 3: Sharing a material artifact to understand headings

Winter and her mother are using a compass tool placed around the robot to decide on the appropriate heading value to make it move to their next destination, “a sushi restaurant” (Figure 3a). After Winter assigns the reference heading, her mother asks what heading she chose. We see this question as an invitation for deliberation. Winter examines the compass and responds hesitantly, “Uh zero?” (Figure 3b). In response, her mother provides verbal, tonal, and gestural feedback which implicitly indicates that Winter should reconsider (Figure 3c). To process her mother’s feedback, Winter briefly moves her finger counterclockwise along the compass. She then gestures with her hand to the left and asks if the direction she is pointing in is “more like 10” (Figure 3d). Her mother interprets from the response that Winter is still having difficulty estimating the angle adjustments based on a point farther away from the compass. Demonstrating reflexive awareness that her previous explanation using language and gesture was insufficient, Winter’s mother pauses to suggest that they “get grandpa’s ruler”, which she races from room-to-room excitedly to find. This visual representation they construct with material artifacts serves to make the ephemeral radius that they both are referring to a durable representation. With the ruler in place (Figure 3e), Winter’s mother sees an opportunity to reset the robot’s reference heading to match the heading of their first destination. They coordinate their actions as Winter’s mother holds the ruler in place while Winter positions a finger of her right hand in line with it on the compass tool. With her finger held in place marking the spot to which she needs to reset their reference heading, she sets the value (Figure 3f depicts a recreation for clarity). Next, Winter’s mother checks for understanding by asking Winter the new value they need to input to reach the destination. Winter quickly responds with a definitive and punctuated “zero”, an input that achieves the goal.

Figure 3. Winter and her mother use a ruler to help with selecting headings.

The ruler placed over the compass is simultaneously a physical representation of their prediction, a tool for greater precision, and an object around which to interact and build further meaning. Accordingly, the mother recruits this everyday object, a tool that is familiar to them both, to generate a locally productive (Hall & Stevens, 2015) flexible learning system (Hike, 1989) in support of the pair’s co-construction of knowledge. In this way, they are able to use material artifacts to represent numerical headings (Hutchins, 2005). Winter’s mother monitors the physical engagement with this artifact to evaluate Winter’s knowledge, the focus of her attention, and the use...
of this strategy that is meant to help make sense of angles and headings. Staying close to the data only allows us to speculate about the cultural significance of the label “grandpa’s ruler.” However, the fact that the pair label it as their grandfather’s possession, and that Winter exhibits such excitement as she races to various rooms in her house to locate it, indicates how the familiarity and intimacy of objects available in the home may be central to families’ practices of constructing visual representations for problem solving.

Discussion and conclusion
This paper described how parents and children make meaning about mathematical and spatial concepts encountered during a playful coding activity at home. In all three cases, children reached moments of impasse during the activity when they grappled with mathematical and spatial concepts. For Landon, this moment surfaced from a need to simultaneously remember and consider information about routes, sequences, and spatial relations. For Alan and Winter, these moments were specific to understanding and manipulating numerical headings. Upon these moments of impasse, all three parents provided critical support that allowed their children to accomplish the coding task at hand. In turn, the pairs were able to continue engaging with the computational aspects of the activity.

Our analysis reveals a theme that parents adapt their communication to shift from verbal to more physical and material representations over the course of meaning making about mathematical and spatial concepts involved in computation with their children. This evolution in modalities emerged as parents closely monitored their children’s use of multiple sign systems (Goodwin, 2000), including verbal communication, gestures, and computational input to the programming software. These sign systems served as representations that revealed how children were understanding the content. Parents reflexively responded to this understanding by improvising their communication strategies and recruiting various modalities, such as gestures, embodiment, and the use of material artifacts. In recruiting these resources, parents choreographed their children’s bodily movements and directed their children’s attention. These behaviors occurred in an intimate and seamless way that is unique and second nature to the parent–child relationship.

Ultimately, this evolution in modalities of communication was critical for the pair’s meaning making about the mathematical and spatial concepts that emerged during the coding activity. By recruiting gestures and material artifacts and by coordinating embodiment, parents could reframe and reorient their children’s perspectives, while also creating more stable representations of concepts (Hutchins, 2005). This reorientation led the pair to converge their attention on the same system of visual practices and to share the same reference frame. Consequently, the pair achieved intersubjectivity (Goodwin, 2000), allowing them together to establish the use of cognitively opaque cultural tools, such as the compass, and the parent to provide more complete explanations of concepts, phenomena, and practices. In this way, we see parents controlling and constraining parts of the activity that may be too challenging for the children while working alone (Wood et al., 1976). Although not inherently computational themselves, the mathematical and spatial concepts and computation involved in the coding activity are mutually reinforcing, such that a child’s understanding of these concepts supports their abilities to engage in the computation. Thus, this process by which parents reflexively adapt their modalities of communication and direct their child’s attention at critical moments demonstrates the invaluable contribution of the meaning making that happens during computational thinking activities between parents and children.

In light of what we see as a dearth of information about how children learn computational thinking skills at home, this research depicts a micro-level analysis documenting the learning process that occurs when children collaborate with a parent on computational tasks. The method we employ offers a granular lens through which we could see the process of collaborative and distributed meaning making during computational tasks and identifies patterns in how this process unfolds. Moreover, our analysis works to account several attributes of the home learning environment, such as the availability of a parent, and acknowledges how this context facilitates parents’ and children’s abilities to call on familiar material objects, prior knowledge, and social resources that can support their meaning making. Accounts like these not only offer a more fine-grained account of computational meaning making but can incorporate and emphasize more culturally relevant and everyday forms of learning that happen across many contexts, and with many cultural forms (Horn, 2018), tools, and material resources. Therefore, this account represents an important perspective on children’s computational thinking and learning at home.

References


**Acknowledgements**

We want to express our gratitude to the families who welcomed us (virtually) into their homes for this research. Thanks to fellow students in our lab that offered their assistance, as well as members of our video data club for their thoughtful insights. Also, thanks to our reviewers for contributing helpful comments and suggestions. This work is funded in part by a graduate research grant from The Graduate School at Northwestern.