

# Multisensory Sensemaking: Children's Exploratory Behavior has Organizing Structure at Micro- and Macro-scales

Nancy Owens Renner (nrenner@ucsd.edu)

Department of Cognitive Science, University of California San Diego  
9500 Gilman Drive, La Jolla, CA USA 92093-0515

## Abstract

When children encounter museum exhibits, they find rich opportunities for action, perception, learning, and other forms of cognition. Can we see systematic organization in the children's behavior, and by extension, their cognition? What would count as evidence of this organization? Based on an account of cognition as embodied, situated, and culturally mediated, this research illustrates how some cognition can be directly observed, manifested through interactions among modalities, people and objects in a distributed cognitive system. This field study uses micro- and macro-analyses of behavioral data recorded on video to discover organizing structure in children's behavior and cognition, evidenced in allocation of visual and haptic attention, manipulation of objects, and use of written, spoken, and gestured language. At micro- and macro-scales, perceptual engagement with concrete objects precedes engagement with abstract concepts, as evidenced in motor behavior and content of speech and gesture. In this context, we describe learning as a process of adaptive coordination, rather than a product to be measured. Future studies will test observation-driven hypotheses related to development of perceptual skills and patterns of scientific thinking, promising relevance to educational practices.

**Keywords:** multimodal attention; distributed cognition; situated cognition; perceptual; conceptual; embodied learning; education

## Introduction

To describe the complexity of learning involves an ecological model built of interactions among human brains and bodies, cultural practices and artifacts, all subject to evolutionary and developmental change over time. "Learning ecosystems" represents the prevailing perspective among educational researchers and cognitive scientists today (NRC, 2009). This ecological view stands in stark contrast to first generation cognitive science and folk models of learning. These focus on symbolic information processing, often with a disregard for the role of context, culture, and history, and ignoring cognition as a biological phenomenon that includes emotion, sensory-motor experience, and probabilistic models of learning through experience (Gardner, 1985). This study uses observational methods and defines cognition as:

- embodied, i.e. a body is required for cognition, and the nature of bodily experience profoundly shapes cognition;

- situated, i.e. physical, social, cultural, and historical contexts all play a role in cognition;
- distributed, i.e. cognition at every scale manifests in nested systems with interacting parts. People live and act in cognitive ecosystems, in which brains, bodies, cultural and material worlds all exert mutual cognitive influence in any given moment and as they change over time (Hutchins, 2010). Distributed cognition provides a unifying principle.

## Research Methods

To understand how the designed environment structures cognitive activity, we collected and examined data with several questions in mind: Can we see systematic organization in the children's behavior, and by extension, their cognition? What would count as evidence of this organization? How do children move their bodies, use their senses, and utilize language to explore objects and engage with ideas? The research team collected video data over several days, including first-person perspective video, recorded with head-mounted cameras worn by elementary students, and third-person perspective video, recorded with hand-held cameras operated by researchers. These two perspectives provide a wealth of information about what children pay attention to, as well as their interactions in physical and social contexts. These two video data streams amount to approximately 16 hours of video, primarily focused on 10 participants (five girls, five boys, all bilingual), some for two visits to the exhibition, with supplemental video recorded at specific exhibits.

The research team indexed the video to create an outline of the event structure, making large-scale behavioral patterns and anomalies apparent for more detailed study. We transcribed and coded a subset of the video in a densely detailed manner (figure 1). The video coding system includes multiple dimensions that fall into broad categories related to modalities in coordination with physical objects, social interaction, speech, and reading behavior. The coded video permits a broad array of analyses, for example, duration and sequence of behaviors with individual exhibits, specific objects of attention (looking at and touching specimens, interactives, graphic panels), composition of social group, speech and gesture content, language (English or Spanish), reading, and behavioral configurations defined by multimodal coordination.

To answer the research questions, this study sought to describe a learning ecosystem and its organizational

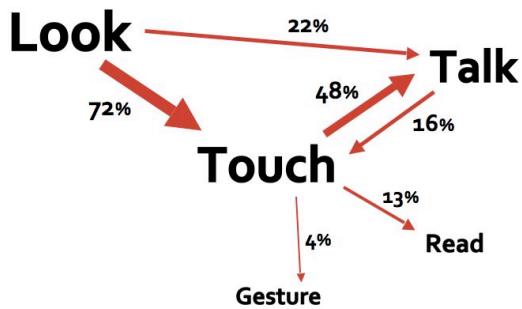
structure. Toward this goal, the research documents the abundance and diversity of behaviors, their distribution in time and space, and interaction effects. Sustained observation led to hypotheses about the organization of multimodal behaviors and content of speech. The coded video permitted testing of these hypotheses in a quantitative manner. Allocation of attention, sensory-motor engagement, and the production and timing of speech and gesture provide observable behavioral evidence of cognition in action, and illustrate learning as an adaptive process that takes place through interaction and over time.



**Figure 1: Behavioral configurations become visible in the coded video.** The research team used ELAN software and a customized behavioral coding scheme to represent the occurrence of behaviors in the horizontal rows; the red vertical line represents a behavioral configuration in one moment of time. Student's first-person view on left; researcher's third-person view on right.

## Results – Micro-analysis

Not surprising among sighted participants, all observed interactions with exhibits begin with looking. Focus of visual attention on an exhibit defines the beginning of an event. The majority of events proceed from looking only to looking and touching (figure 2). Most exhibit-related talk is preceded by touch. When focusing on exhibit-related behaviors, the data show a normative sequential pattern of Look-Touch-Talk in 48% of all events. Social interactions and physical constraints generate variations on this pattern.



**Figure 2: Behavioral sequences unfold with each exhibit interaction.** This diagram shows the frequency of occurrence in sequence for the first onset of behaviors. Look-Touch-Talk is the most common sequence of behaviors.

By re-representing the coded video data, we can see how multimodal behaviors combine in sequence revealing a distinctive behavioral profile that includes look, talk, manipulate, touch, gesture, and read (figure 3). This *ChronoVis* representation of the coded video shows a Look-Talk-Touch pattern. Here, the first onset of speech performs a social function, i.e. a proposal to access the exhibit (spoken in Spanish, demonstrating sensitivity to the recipient's language abilities). Once the child has access, his hands engage the object. He moves a knob on the front of the display (in his peripheral vision) while he looks at top of the display. He coordinates his hand motion with visual feedback to align single words in a cut-away window, which he reads aloud for himself one-by-one during pauses in his motion, "past—present—future." Then his hand sweeps over the textured surface of this representation of the San Andreas Fault, he looks at the graphic panel then back to the surface, and he moves the knob back and forth.

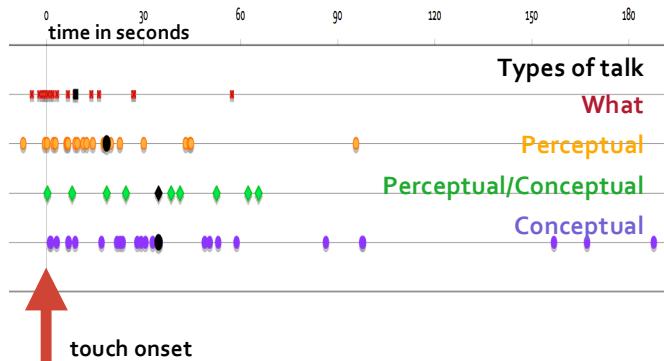
Another child arrives and asks, "What is this?" The first child returns the knob to reset position and reads "past," as he points to the word then traces an arc over the sculpted landscape. He moves the knob again, reads "present," underlines the word with his finger and briefly points at the landscape. He moves the knob a final time, reads "future," takes a brief look at the graphic, then cedes possession of the interactive to his classmate. In 24 seconds of activity, this child negotiates two social interactions in two languages, coordinates visual and haptic attention to guide sensory-motor exploration; he reads text to establish context and make some sense of the object, then he uses interaction with the object combined with gesture and speech to demonstrate the object's function, and perhaps its meaning, in response to a peer's question.



**Figure 3: Re-representation of video data in ChronoVis to illustrate how a behavioral profile unfolds in time.** Two fourth-grade boys interact with a plate tectonics display. Event duration: 24 seconds. X-axis = time.

When interacting with objects and exhibits, as mentioned above, children tend to look first, then they touch, and then they speak. The Look-Touch-Talk pattern varies along several dimensions, including the content of the speech.

When analyzed across multiple exhibit interactions, talk related to what an object is and the object's concrete perceivable features tends to occur early in the interaction; talk with abstract conceptual content occurs later (figure 4). A particular form of perceptual talk—"what" questions and comments ("What's that?" "It's a fossil")—happen at the beginning of the interaction most often. Other perceptual talk includes describing observable features ("They're so tiny") or perceptual processes ("Look, touch it"). Conceptual talk refers to abstract ideas, unseen processes, or representational significance, and perceptual/conceptual talk contains both concrete and abstract content. These tend to come later in sequence. This work builds on a study of types and frequencies of learning talk in a museum, which informed our categorization (Allen, 2002). Of all speech, these types of exhibit-related talk totaled 46% and talk with social content totaled 38% (e.g. "Get in line!"). The remaining 16% had ambiguous content or was inaudible. To categorize the content and function of speech is certainly problematic, because any utterance may carry multiple meanings and perform several functions. Nevertheless, we sought the defining characteristic for each utterance and reduced the high-dimensional space of speech to look for sequential patterns.



**Figure 4: The timing of when talk follows touch depends on the type of talk.** The origin represents the first touch onset. Black shapes indicate the average latency.

With this field study, we set out to describe cognition evident in freely behaving humans, and posit structures that organize the children's activity at multiple scales. The quantitative data here are meant to describe the magnitude of these phenomena. Future quantitative analyses will explore in finer-grained detail how these patterns vary with and without adult mediation, and with different kinds of displays and various objects to touch.

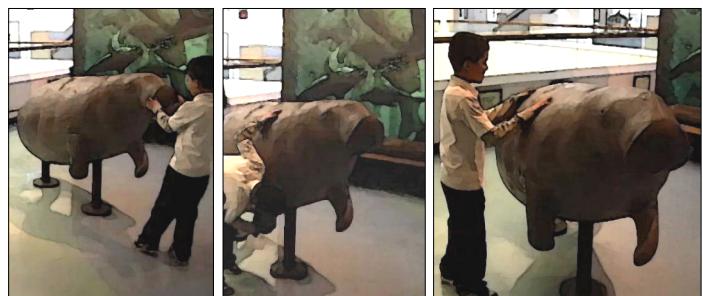
## Results – Macro-analysis

Based on this micro-analysis of behavior at a time-scale of milliseconds, we can also describe macro-scale behavioral patterns organized into larger activities. By engaging multiple modalities (sensory, motor, and communicative), the children use their bodies as if to explore tacit questions.

Sensory-motor behavior and language content characterize activities described in these ways: children engage in active sensing; they discover exhibit uses (or affordances); they look for cause and effect; they use their bodies to express ideas; they demonstrate, interpret, and explain. Specific examples below (with still shots from the video data) serve as abbreviated illustrations of a broader class of examples.

### Active sensing • *What is it?*

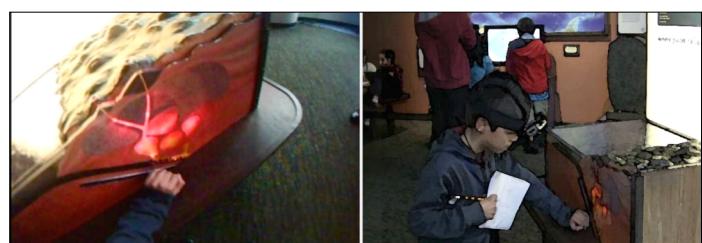
Sensory-motor exploration dominates the children's behavior. The children learn about the properties of objects and environments by experiencing relationships between movement and the senses (Nöe, 2004). They move bodies, hands, heads, and eyes, driving sensory input through multiple modalities simultaneously (figure 5). Using their bodies, often involving pointing gestures and sometimes expressed in words, the children ask the question "what is it?" Often when children read, they trace the text with their fingers, they voice the words, and sometimes read in unison. For them, reading is a physical, social, exploratory activity. Multisensory displays allow for complimentary and reinforcing experience through many channels.



**Figure 5: Children engage in active sensing** using visual, tactile, motor, proprioceptive, and auditory systems.

### Discovery of exhibit uses (or affordances) •

**What can I do with it?** Children's perceptual skills help them to find how their bodies fit with objects in the environment and where they can take action (Gibson, 1979.) Through imitation learning, observation of others plays a role in finding affordances. With a plate tectonics interactive display, a child pushes down on a knob (figure 6). With this movement, he initiates a series of visual representations: one plate dives beneath another and begins to melt (subduction), magma rises to Earth's surface (volcanism), and the overlying continental plate lifts up (uplift). Interaction also provides multisensory experiences of cause and effect.



**Figure 6: Children discover exhibit uses.** Student's first-person view on left; researcher's third-person view on right.

### Look for cause and effect • *What will happen if...*

The discovery of cause and effect relations is a means to develop perceptual skills (Piaget & Inhelder, 1969). Active sensing (or finding sensory-motor dependencies) and discovery of object uses (or affordances) embody children's exploration of causality. Children also test the limits of their actions relative to responses in physical and social realms. They seem to explore patterns of dependence, which they can use to infer causal structure (Schultz & Gopnik, 2004).



Figure 7: Children look for cause and effect.

In a room filled with exhibits that move when pushed and pulled, this student finds a sandstone block with a ridge down the middle (figure 7). He places his fingers on the ridge, pulls toward his body, and tests to see if the rock will move. In contrast to several other objects that he set in motion, he discovers that this rock remains immobile.

### Use the body to express ideas • *How do I make sense of this?*

Abundant research suggests that gesture and speech derive from the same source (McNeill, 2005). Both give the researcher information about the children's perceptual experience and conceptualization. Children use gesture to express ideas that they can't easily express in words, due to language development (especially among bilingual children) (Goldin-Meadow, 2003) or because gesture can convey spatial and temporal qualities in a manner different from speech (Parrill & Sweetser, 2004). Children use gesture as they speak to model two forms of volcanic eruption in the subduction display shown below and in figure 6. They use their hands to perform an oozing spread of lava or a violent explosion. Gesture complements speech: "when it builds too much pressure (hands move up rapidly) it goes up!" (fig 8).



Figure 8: Children use multiple modalities to express ideas.

In our sample taken when the children were freely exploring the exhibits, representational gestures were uncommon relative to pointing gestures, comprising 20% of all manual gestures. Half of all representational gestures occurred at the subduction display; most gestures originated strategically where the volcanic vent meets Earth's surface. We conjecture that the children used their hands to fill a gap in the display's representational content. The Earth's cross-section shows glowing magma rise to the surface, but then what? The children seize the opportunity—with fluid virtuosity—to enact the story's eruption climax using their hands, sometimes with explosive vocal sound effects.

### Demonstrate, interpret, and explain • *Can I show you / can I tell you?*

In this social setting, the children comment on the exhibits, they share what they perceive and know, often in short sentences accompanied by gesture anchored to the objects. Their demonstrations, interpretations, and explanations may be spontaneous or, as in the video stills (figure 9), elicited by a teacher.

In this example, their guided lesson focuses on plate tectonics, and the children must identify different types of plate boundaries. The teacher asks what the children see in the interactive exhibit; they convey in words coupled with body movements the upward thrust of the uplifted plate resulting from subduction. The teacher's orientation lesson with an activity sheet make explicit a conceptual framework for plate tectonic theory with three types of plate boundaries. Highlighting this structure enables the children *to see different features of the display than they saw before*.



Figure 9: Children demonstrate, interpret, and explain.

When enacting tectonic uplift, the morphology of the gestures express a shape-for-shape mapping (Taub, 2001) in which the palms represent the horizontality of the Earth's crust moving in an upward direction (figure 9). When the class activity specifically focused the children's attention on the structure of the plate boundary and its consequences, the children produced uplift gestures nearly equal in number to volcano-type gestures. When children independently explored the subduction display, the majority of their gestures expressed aspects of volcanism, possibly an indicator of its salience. Different forms of mediation provided by the teacher, i.e. guided observation and an explicit conceptual framework represented in diagrammatic form, facilitated new forms of active perception for scientific thinking. The children used their bodies coupled

with objects within a framework for social participation to enact a personal and collective understanding of a geologic phenomenon.

## Discussion

Each one of these macro-scale activities are comprised of micro-scale behaviors that involve interaction with objects and other people through looking, touching, talking, moving, and reading. The interaction described in the micro-analysis results (figure 3) illustrates how these individual multimodal behaviors integrate into macro-scale cognitive activities. In that example, the child begins with active sensing to discover the object's affordances and cause-and-effect relations. He seems to make sense of the object as evidenced in the sequence and coordination of sensory-motor behaviors, the way that he demonstrates the object's function, and using very economical means, explains an aspect of the object's meaning to a peer.

Sensory engagement is for making sense of a situation. Children focus attention of multiple modalities simultaneously, tuning effort and perception to gain what is useful in a given context. For children in a science museum, analysis of their multimodal interaction brings insight about the organization of their behavior and how their behavior creates a blend of perceptual and conceptual experiences.

The form and dynamics of the children's multimodal exploration, the content of their speech, and their use of gesture, suggests that they ask and answer questions with their bodies and words. Based on observation of children's behavior, I propose that educators can work with, not against, children's predispositions toward multimodal exploration and curiosity. Children seem to engage exhibits with an inquiry approach, beginning with questions that blend: "What is it? / What can I do with it? / What will happen if...?" Under some conditions, the inquiry proceeds to "How do I make sense of this? / Can I show and tell?". This description may not surprise the reader. The surprise is the degree to which formal and informal education does not exploit children's robust social and sensory-motor skills to make sense of the world.

Conceptual understanding builds on perceptual experience, as interacting with objects and phenomena generates the raw material of experience from which to create memories, extract features, form categories, make causal inferences, and develop abstract concepts. Social interaction can facilitate this process, as peers and teachers point out important things to notice and explain their significance. In this study of children's behavior, we see the multi-faceted intertwined nature of cognition, action, and perception. Bodies sense and move and interact with objects, other bodies, and ideas. Talk can mediate experience by identifying, describing, explaining, and making meaning. This coordination of resources—of the brain, body, social and material worlds—constitutes learning in action.

## Conclusion

Children's efforts to make sense of their perceptual experiences exert organizing structure on behavior and cognition. The body is the vehicle of cognition, doing its work through physical and social interaction. The children engage in a continuous project of bringing perceptual, motor, and other cognitive resources of the body into coordination with structure in the world. Making sense of a situation may require, even demand, multisensory engagement. Future work will explore: How can we deliberately create structure in the learning environment that offers opportunities to *practice learning as a set of skills*? These skills include learning how to see, learning how to coordinate multimodal perceptual experience, and to use perceptual experience to solve problems, make inferences, and produce generalizable explanations. Future work will focus on advancing perceptual learning theory, developing analytical methods, and creating an educational design framework that promotes practice of perceptual skills for conceptual learning. For educators, designers, and researchers of learning and cognition, rich territory for exploration lies at the intersection of body, senses, and world.

## References

Allen, S. 2002. Looking for learning in visitor talk: A methodological exploration. In Leinhardt, Crowley, Knutson (Eds.) *Learning conversations in museums*.

Gardner, H. 1985. *The mind's new science*. New York: Basic Books.

Gibson, J.J. 1979. *The ecological approach to visual perception*. Boston: Houghton Mifflin.

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

Hutchins, E. 2010. Cognitive ecology. *Topics in Cognitive Science*, Vol. 2, Issue 3: 1-11.

McNeill, D. 2005. *Gesture & thought*. Chicago: The University of Chicago Press.

National Research Council. 2009. *Learning Science in Informal Environments: People, Places, and Pursuits*. Committee on Learning Science in Informal Environments. Bell, P., Lewenstein, B., Shouse, A., Editors. Washington, DC: The National Academies Press.

Nöe, A. 2004. *Action in perception*. Cambridge: MIT Press.

Parrill, F., Sweetser, E. 2004. What we mean by meaning: Conceptual integration in gesture analysis and transcription. *Gesture*, 4:2, 197-219.

Piaget J. & Inhelder, B. 1969. *The psychology of the child*. New York: Basic Books.

Schulz, L., Gopnik, A. 2004. Causal learning across domains. *Developmental Psychology*, 40, 2, 162-176.

Taub, S. 2001. *Language from the body: Iconicity and metaphor in American Sign Language*. Cambridge: Cambridge University Press.

This study was conducted by Nancy Owens Renner, in cooperation with the San Diego Natural History Museum and the San Diego Unified School District. Research oversight was provided by University of California San Diego Cognitive Science faculty Edwin Hutchins, Jim Hollan, and Rafael Núñez. Research assistance was provided by Alba Basurto, Gina Bello, Richard Caballero, and Jordan Davison. Parents granted permission for their children to participate in this research by signing release forms made available in English and Spanish.