Title

Intermodality in Multimodal Learning Analytics for Cognitive Theory Development: A Case from Embodied Design for Mathematics Learning

Permalink

ISBN

9783031080753

Authors

Tancredi, Sofia
Abdu, Rotem
Balasubramaniam, Ramesh
et al.

Publication Date

2022

DOI

10.1007/978-3-031-08076-0_6

Intermodality in Multimodal Learning Analytics for Cognitive Theory Development: A Case from Embodied Design for Mathematics Learning

Sofia Tancredi[1][2], Rotem Abdu[3], Ramesh Balasubramaniam[4], and Dor Abrahamson[1]

[1] University of California Berkeley, Berkeley CA 94720, USA
[2] San Francisco State University, San Francisco, CA, 94132
[3] University of California Merced, Merced, CA, 95343
[4] University of Haifa, Haifa, Israel, 3498838
`sofiatancredi@berkeley.edu`

Abstract. Multimodal Learning Analytics (MMLA) grant us insight into learners' physiological, cognitive, and behavioral activity as it unfolds. In this chapter, we query the relations among modalities, *intermodality,* in the context of a design-based research program studying the relations between learning to move in new ways and learning to think in new ways. In part I, we reflect on how different methods have afforded purchase on the investigation, development, and elaboration of theoretical claims about the *multimodal* enactment of cognitive events, culminating in the use of Recurrence Quantification Analysis (RQA) to quantify the microgenesis of stable new patterns in hand movement and gaze. In part II, we analyze an RQA case study spanning across hand and gaze modalities to examine the emergence of *intermodal coordination* at a critical moment in the mathematical task. We conclude with implications and open questions around intermodality in embodied learning.

Keywords: multimodal learning analytics, perception, embodied cognition, mathematics, intermodality

The embodied turn in the cognitive sciences brought forth reconceptualizations of what it means to think and learn. Embodied cognition philosophers proposed that bodily activity is central to the activity of the mind; indeed, enactivists posited that cognitive structures emerge through repeated patterns in the perceptual guidance of motor activity (Varela et al., 1991). To take this stance seriously is to view learners' sensorimotor activity as playing a defining role in their cognition. Multimodal learning analytics here offer themselves as indispensable. Multimodal data are positioned to grant unprecedent access into cognitive activity.

To take an enactivist stance seriously is likewise to reimagine pedagogy to explicitly cultivate new ways of perceiving and moving as new ways of thinking (Abrahamson & Sánchez-García, 2016). Such an approach radicalizes the historical view in cognitive-developmental psychology research that mathematical reasoning has roots in sensorimotor activity (Piaget, 1968; Steffe & Kieren, 1994; von Glasersfeld, 1987). Design-based research program *embodied design* (Abrahamson, 2009, 2014, 2015a) imagines an enactivism-based pedagogy where teaching and learning are understood as the cultivation of new perceptual structures. More specifically, embodied design seeks to ground concepts in students' existing perceptuomotor capacities. Embodied designs create the conditions for learners to discover new ways of moving that ground mathematical concepts. Drawing on cultural–historical psychology (Vygotsky, 1926/1997), embodied design activities introduce disciplinary and symbolic artifacts as resources for enhancing students' pragmatic, epistemic, and discursive activity. Through this process, students come to perceive their own activity in ways that ground semiotic expression.

The Mathematics Imagery Trainer for Proportion (MIT-P) is one instantiation of embodied design. This chapter will trace the MMLA history of the MIT-P project, illustrating how different analyses afforded purchase on multimodal learning phenomena, iteratively informing the theorization of the relationship between movement and mathematical thought. We will focus on a current frontier: modeling system dynamics with Recurrence Quantification Analysis, discussing two studies applying this method: Tancredi et al. (2021), which tracks the dynamics of hand coordination in MIT-P problem-solving, and Abdu et al. (under review), which tracks the dynamics of eye movements in MIT-P problem-solving. Setting forth from these analyses, we query the multi in multimodality, beyond the phenomena arising in each modality towards centering the interactions between them. We present an in-depth case study looking closely across data from each of these analyses to examine the microprocesses of intermodal perceptuomotor learning.

1.1 Overview of the MIT-P Project

The Mathematics Imagery Trainer for Proportion (MIT-P) is a tablet-based math instructional design for learning proportional reasoning (Abrahamson & Trninic, 2011). As an embodied design, the MIT-P is built to foster a new way of moving through goal-directed action under designed constraints. The design also takes up sociocultural theory by introducing cultural-symbolic mathematical artifacts as resources to enhance action and discourse. In the activity, users manipulate two parallel bars on a touchscreen with their fingers. The bars start off red, and learners are tasked with figuring out how to turn them green and keep them green while moving their fingers. The bars turn green when the ratio of the left bar to the right bar is 1:2, that is, when the left bar is at half the height of the right bar (Fig. 1). The MIT-P is designed to productively disrupt the common "additive" assumption, often displayed by children first learning ratio, that the difference between two quantities should remain constant as the quantities increase, i.e., that a 1:2 ratio is equivalent to a 2:3 ratio and a 3:4 ratio. To fulfill the task directive of moving-in-green, users discover that instead of keeping the distance between their hands constant (instantiating an "additive" relation), they must instead increase this distance as the bars grow longer (instantiating a "multiplicative" relation).

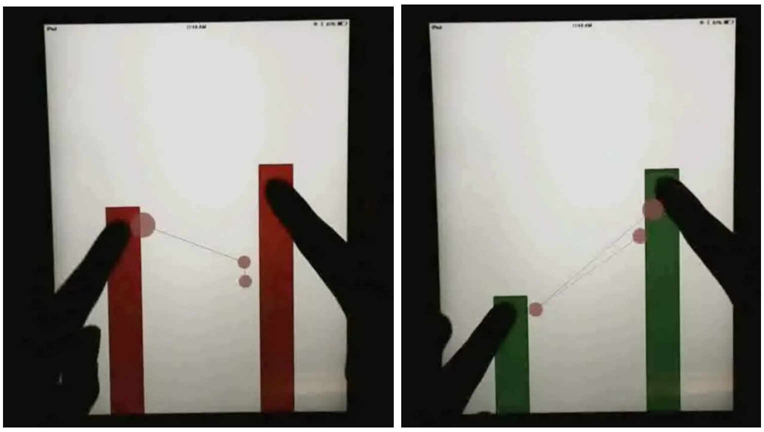


Fig. 1. Photos of user engaging with the MIT-P tablet application. In the first image, the bars are red because the encoded ratio, here 1:2, is not met. The bars turned green in the second example because the height of the left bar is in a 1:2 ratio with the height of the right bar.

1.2 Theoretical Framework: Intermodal Perception

Drawing on phenomenologist philosophers Husserl and Merleau Ponty, James Gibson (1966) defined perception as an active process that draws upon the relations among different perceptual organs to accomplish ecologically viable dynamic behavior. He wrote, "certain higher-order variables—stimulus energy, ratios, and proportions, for example—do not change" (p. 3). Eleanor Gibson (1969) elaborateD upon this view by distinguishing several types of intermodality[1]. One is *intermodal transfer*, whereby perceptual discrimination abilities in one modality carry over in another modality. She gave the example of how a man who lost his sight at 10 months and recovered it over the age of 50 could recognize letters visually, which he had only previously experienced through the tactile modality. Notably, intermodal transfer can be inhibited by the presence of distinctive modality-specific properties such as color (E. Gibson, 1969). Another type of intermodality is that of *amodal perceptual properties*. In contrast to modality-specific properties, such as the color blue, amodal properties are higher-order relational structures such as jerkiness, or a transition from rough to smooth. These properties are not modality-specific, although they reside in properties of light, sound, movement, or vibration. Perceptual development in Eleanor Gibson's model is a process of discovery of intramodal and intermodal invariant stimulus relations and transformations that become more specific and attuned to higher–order relations (E. Gibson, 1969).

The MIT-P is a fertile context for examining intermodality in that the activity of multiple systems participates and intertwines. Learners move their hands on a

---

[1] We use the term intermodality as coordination perceptual information across modalities. This usage is distinct from intermodality as multisensory integration (e.g., Ernst, 2008).

touchscreen, monitor their action visually, and dialogue with a tutor. Bodily participation of the hands and eyes is sensory and motor: the hands perform motor action *and* provide kinesthetic-proprioceptive feedback to the learner. Similarly, eye gaze detects color-feedback and supports visual proprioception (J. Gibson, 1966) of the hands' positions *and* moves in strategic, goal-directed ways to participate in the control of action. We begin by tracing strands of research investigating each of these modalities in the MIT-P project, showing how these threads converge towards an intermodal research agenda.

2 Multimodal MIT-P Analyses: A Brief History

As a design-based research project, the MIT-P has gone through numerous iterations and forms and has elicited evolving analyses. Data collected from users of the MIT-P includes eye tracking, tablet data, transcripts, and video. Early work on the MIT-P project used detailed qualitative analysis of video and eye-tracking data to study the microgenesis of the target movement pattern. Such analyses have been pivotal in identifying how learning unfolds in the MIT-P, including the strategies learners use (Abrahamson, Lee, et al. 2014) and the role and activity of the tutor (Abrahamson et al. 2011; Flood et al., 2020; Shvarts & Abrahamson, 2019). Later work took up a range of quantitative methods including statistical methods and machine learning to detect participants' strategies. Below, we describe the methods and findings of studies pertaining to the kinesthetic and visual modalities' role in the MIT-P.

2.1 Hand Movements

Touchscreen data have been effectively used to identify learner strategies. Machine learning was used to differentiate between sequential "A per B" strategies that focused on the relative displacement of each hand and "speed" strategies that focused on how one hand moved faster than the other (Pardos et al., 2018). These hand-movement strategies offer different entry-points into mathematical discourse about the activity. Other researchers applied machine learning to develop an intelligent tutoring agent that responded dynamically to MIT-P learners' needs (Abdullah et al., 2017).

Statistical analyses of MIT-P hand data have highlighted regime-switching dynamics within and between participants (Ou, Andrade, Alberto, Bakker, & Bechger, 2020; Ou, Andrade, Alberto, van Helden, & Bakker, 2020). These analyses used a mixture Regime-Switching Hidden Logistic Transition Process to identify characteristic hand movement-language transcript regimes and assign students to clusters according to their regime transitions. Participants were found to cluster as quick or slow discoverers. Within subjects, participants moved through three phases: an initial phase characterized by hands moving at the same heights, an intermediate phase of exploring different hand relationships, and a final phase of moving-in-green where the hands maintained the target ratio, with occasional relapses into Regime 2. These findings suggest that learners transition between stages with different characteristics in this embodied design environment, reminiscent of phase transitions between different stable regimes (Ou, Andrade, Albert, Bakker, & Bechger, 2020).

2.2 Eye Movements

In parallel, eye tracking studies of MIT-P data have revealed critical insights about student strategy not available to the tutor in real-time. Qualitative eye-tracking studies found that over the course of the task, participants began to shift their gaze towards new areas of interest beyond their fingers such as the projection of the left bar on the middle of the right bar (Abrahamson, Shayan, et al., 2016; Shayan et al., 2015). These gaze patterns in combination with analysis of participant's language about what they were doing suggested that participants were developing *attentional anchors* supporting the transition from additive to multiplicative movement and reasoning. Attentional anchors (Hutto, & Sánchez-García, 2015) are perceptual objects that come forth to facilitate motor action. The anchors are real or imagined relational features that can be manipulated to control activity. For example, dancers might imagine being pulled by a string emerging from the top of their head to control multiple aspects of their posture. Work in cognitive psychology (Mechsner, 2003, 2004) corroborates that the enactment of complex motor actions is often enabled by the perception of Gestalt organizing structures. In the context of the MIT-P, learners were found to exhibit idiosyncratic attentional anchors supporting moving-in-green such as the interval between their fingers, or imaginary lines connecting features on the screen (Abrahamson, Shayan, et al., 2016). Duijzer et al. (2017) corroborated and further described the emergence of attentional anchors quantitatively, analyzing the frequency, duration, visit count, and fixation count of area-of-interest gaze patterns alongside lag-sequential analysis of language transcripts. Eye-tracking study findings mark the importance of multimodality in the MIT-P context: hand positions alone do not capture the critical changes that allow learners to achieve fluency with the task.

2.3 RQA Analysis

Analyses of hand movements suggested the presence of phase transitions between different regimes as learners engaged with the MIT-P task. To further investigate this phenomenon, a next set of analyses turned to dynamical systems theory (Kostrubiec et al., 2012; Thelen & Smith, 1994) and the related nonlinear analysis method of Recurrence Quantification Analysis (RQA) to describe the dynamics of the different phases present in MIT-P problem-solving. We pause here to briefly introduce RQA as context for our RQA-based case study in the coming section before describing MIT-P work with this method to date. RQA is a nonlinear method for studying the structure of a dynamical system by detecting repetition patterns (Marwan, 2007). RQA is particularly apt to the study of dynamical systems in that 1) RQA does not assume linearity, allowing for the study of complex systems with interaction-dominant dynamics, and 2) RQA treats variability as a quality of the system rather than as noise, making it powerful for understanding noisy, nonstationary data (Webber & Zbilut, 1994). RQA begins with the construction of a recurrence plot, a plot that compares every state in a time series to every other. These plots can be constructed for an individual time series, quantifying its dynamics, or a pair of time series, quantifying the degree and stability of these systems' coupling. Features of the resulting recurrence plots are then quantified, such as the proportion of matching states (recurrence rate), percent of points falling on diagonal lines indicating a repeated sequence (determinism), the average length of a repeated sequence (meanline), distribution of repeated sequence lengths (entropy), and duration of continuous repetitions (trapping time). These metrics reflect the repetition, predictability, stability, disorder, and duration of connected states in the time series. Originating in physics and consequently applied in the study of physiology, cognition, joint action, economics, and communication, RQA remains rare in math and science education research, with a handful of notable exceptions (e.g., Fleuchaus et al., 2020; Stephen et al., 2009).

The MIT-P has been the first context for the application of RQA to embodied design data. A first study (Tancredi et al. 2021) modeled the dynamics of the bimanual system using RQA as learners discovered and began to move fluently in green. A second (Abdu et al., under review) modeled the dynamics of the gaze patterns in relation to phases of hand movement. Both RQA studies segment participants' time series according to phases appearing in the majority of interviews: the initial phase corresponds to Exploration, searching for an initial green, following by Discovery, where the participant spends substantive time in green positions, and finally, Fluency, wherein the participant moves both hands simultaneously while keeping the bars green.

Hand Cross-RQA[2]**.** The first relationship of interest in the MIT-P context was that between the hands. The ratio of the left hand to the right is the conceptual core of the MIT-P activity. How do the nonlinear dynamics of this core bimanual coordination pattern evolve? Whereas previous analyses of hand data identified strategies or phases, RQA enabled comparison of hand coordination within each of these phases in terms of their stability, predictability, order, and duration of connected states. In a cross-RQA analysis of the left- and right-hand data (Tancredi et al., 2021), we found that when participants reached Discovery, they exhibited an increase in the RQA determinism metric, reflecting an increase in the predictability and coupling of the hands. When participants reached Fluency, this yielded a leap in RQA metrics of recurrence rate and meanline, reflecting an increase in coupling, stability, and predictability. These findings corroborate that a qualitative change arises between

---

[2] Cross-recurrence quantifies dynamical aspects of the coordination of two time series (here, the left- and right-hand position time series). Auto-recurrence quantifies recurrent patterns within a single time series. The gaze data are categorically coded according to areas of interest relative to the hands such as the top of the left bar or the middle of the right bar (see Abdu et al., under review). Thus, cross-RQA analysis compares the continuous hand position time series, whereas the categorical auto-RQA compares the categorical gaze time series to itself.

finding-piecemeal-greens and moving-in-green through which a new, more stable, more predictable coordination emerges.

Gaze Auto-RQA. RQA analysis of bimanual data showed abrupt reconfiguration into “moving-in-green.” Do we see similar changes in gaze patterns? How might gaze be bound up in the gradual constitution of bimanual coordination? Is pattern emergence an independent, unimodal phenomenon? Another study (Abdu et al., under review), examined the sequence of gaze fixations in areas of interest corresponding with the dynamic MIT-P solution space over the course of the activity with categorical RQA. Whereas prior gaze analysis unveiled types of patterns in the data overall, RQA was able to shed light on the dynamics of these patterns within each phase of hand coordination. This permitted looking across participants’ idiosyncratic solutions to identify a common quality of fluent performance: gaze-pattern dynamic stabilization. Abdu et al., (under review) found a decrease in entropy during the Discovery phase indicating an increase in the level of order in gaze behavior after finding green, as well as an increase in recurrence rate, determinism, meanline, and trapping time in the Fluency phase. These findings showed that gaze patterns were more ordered in Discovery than in Exploration, and more stable, repetitive, and predictable during Fluency than in prior phases.

Hand and gaze RQA analyses validated the distinct dynamics of Exploration, Discovery, and Fluency phases of the MIT-P task for both hand coordination and gaze patterns. The convergence of both hand and gaze dynamics towards greater stability in Fluency suggests that the dynamics of these two systems may be related. We will investigate this hypothesis in the next section by juxtaposing the moment-to-moment progression of hand and gaze dynamics for a focal participant.

3 From Multimodal Gaze and Hand Movement to the Intermodal Emergence and Stabilization of Attentional Anchors—an RQA Case Study

The diverse multimodal analyses of MIT-P data point to the interactions between hand and gaze as critical to the process by which learners increase their grip on the problem space, culminating in mathematical insights. We will delve into the case of a single participant appearing in both the Tancredi et al. (2020) and Abdu et al. (under review) analyses.

3.1 Research question

How do the relations between the visual and kinesthetic modalities change as a prototypical participant gains fluency with moving multiplicatively?

3.2 Methods

This case study is drawn from data collected through semi-structured clinical interviews with 39 grade 5 and 6 participants in the Netherlands trying the MIT-P application for the first time (Duijzer et al., 2017). We focus on the first phase of the interview, in which participants first learn the multiplicative movement pattern. Multiple data streams were recorded from these interviews: eye–tracking data using Tobii x2-30, video recordings taken from the participant's perspective, and audio recording of participants' dialogue with the interviewer. We selected this participant, pseudonymed Finn, because their overall trends in RQA mirrored cross-participant trends for both hand cross-RQA and gaze auto-RQA.

To contextualize results within the student's specific learning trajectory, we began with a qualitative analysis of the audio-video recording. We also calculated the correlation between each hand's y-location and y-axis gaze position in each task phase as an initial reflection of the relationship between these two modalities. For a more in-depth analysis of this relationship, we generated windowed RQA plots for this participant's categorical gaze time series and for their two-finger touchscreen locations over time. We focused on the RQA metrics of recurrence rate, determinism, and meanline for this analysis. The windowed plots (Fig. 5-7) consist of the three RQA metrics taken for a window of the coming 50 seconds, sampled for every second of the time series (lag width = 5). Gaze data were coded according to 14 areas of interest: the top, middle, and bottom of the left and right bar, the space between the bars segmented according to the same divisions as the right bar, and the spaces above each bar (coding detailed in Abdu et al., under review). We added a graph of hand heights over time to figures and overlaid the transitions from Exploration to Discovery (0:39) and Discovery to Fluency (3:42) identified in Abdu et al. (under review) onto the windowed plots for reference in connecting the case study to findings from these prior studies.

3.3 Results

Overview: The Case of Finn. We describe Finn's learning process to contextualize our analysis within the broader task-interviewer-environment system. When the task began, Finn found green almost immediately (Fig. 2a) while gazing at his right finger (0:00-0:09). He then raised both hands with the right slightly higher than the left, keeping a roughly constant distance between the bars and maintaining his gaze on the right bar (Fig. 2b) (0:09-0:31). He then reset his fingers to the bottom of the screen and began to raise them together, looking at the tops of the bars (Fig. 2c) (0:31-0:41).

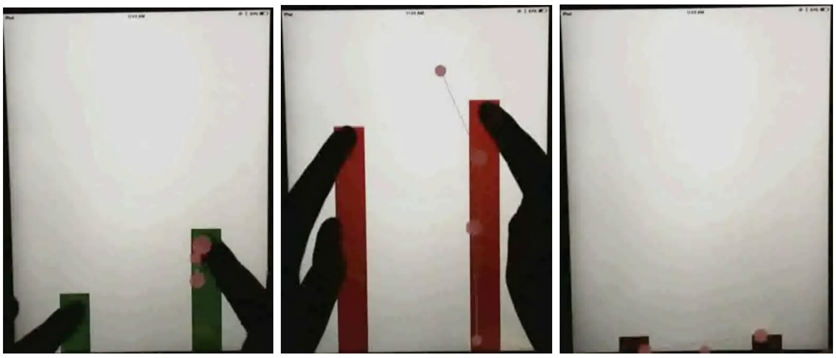
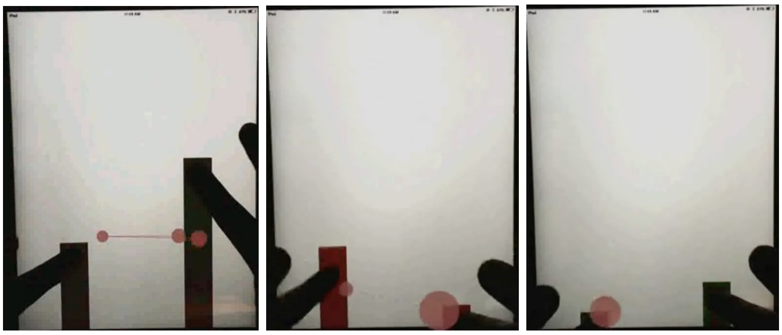


Fig. 1. Finn's activity during the Exploration phase. (a) Finding a first green location. (b) Raising fingers with a set distance between them. (c) Resetting to bottom of the screen.

Finn continued into the Discovery phase by holding his hands in an identified green position, gazing at the top of the left bar and just below the top of the right, drawing his eyes from one to the other almost horizontally (Fig. 3a) (0:41-1:19). He moved his fingers slightly up and down, exiting and re-entering a green position and repeating the visual pattern. At this stage, Finn articulated that to make green, "you have to put both of them on a certain thing that turns green." Then, to illustrate this point (1:19-1:32), Finn inverted which bar was taller (Fig. 3b), commenting: "If you put one here and one goes down, then it doesn't turn green," then restoring the right-bar higher position, "but if I keep this one here, it will be" (Fig. 3c). Prompted to find other greens, Finn tried raising his hands with a set distance between them, eventually adjusting his left hand downwards to find another green (Fig. 3d) (1:32-1:41). During this time, Finn displayed a range of different gaze patterns, prominently featuring finger to finger and left finger to below the right finger. Finn then alternately raised the left and right bars higher (Fig. 3e and f), frequently gazing not only at the fingers but also at the halfway projection of the shorter bar on the taller one. He found a series of green positions this way (1:41-3:42). The tutor then prompted Finn (3:42) to try moving such that the bars stayed green continuously.



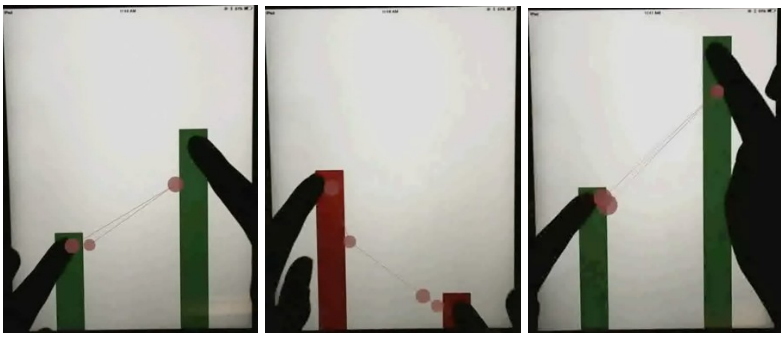


Fig. 1. Finn’s activity during the Discovery phase. (a) Holding green and drawing gaze from top of left bar to midway up right bar. (b-c) Finn shows the interviewer how a left-higher position does not turn green whereas a right-higher position does. (d) Finn works to find other greens, moving his hands at a set distance and adjusting. (e-f) Finn continues to try left-higher and right-higher positions, often gazing at the fingers and a point below the higher finger.

Finn then initiated his most fluent performance of movement-in-green (3:42-5:07). In the Fluency phase, Finn found an initial green and looked from finger to finger (Fig. 4a). He then began to move slowly up the screen, moving his right and left fingers in sequence. As he did so, in addition to the tops of each bar, Finn’s eyes frequently jumped to the middle of the taller right bar (Fig. 4b and c). When asked to explain what he was doing, Finn responded, “I think there must be a certain distance.”

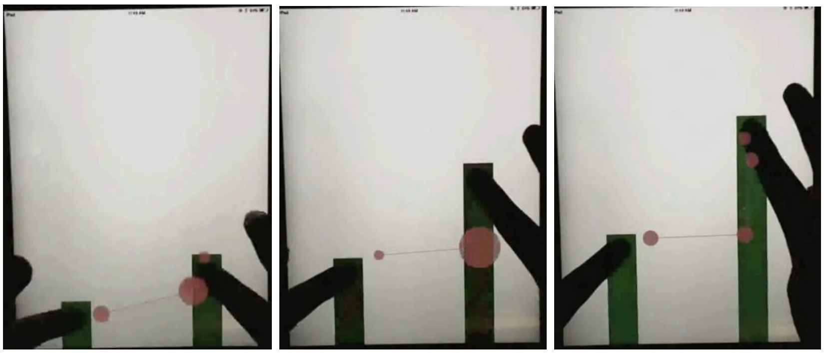
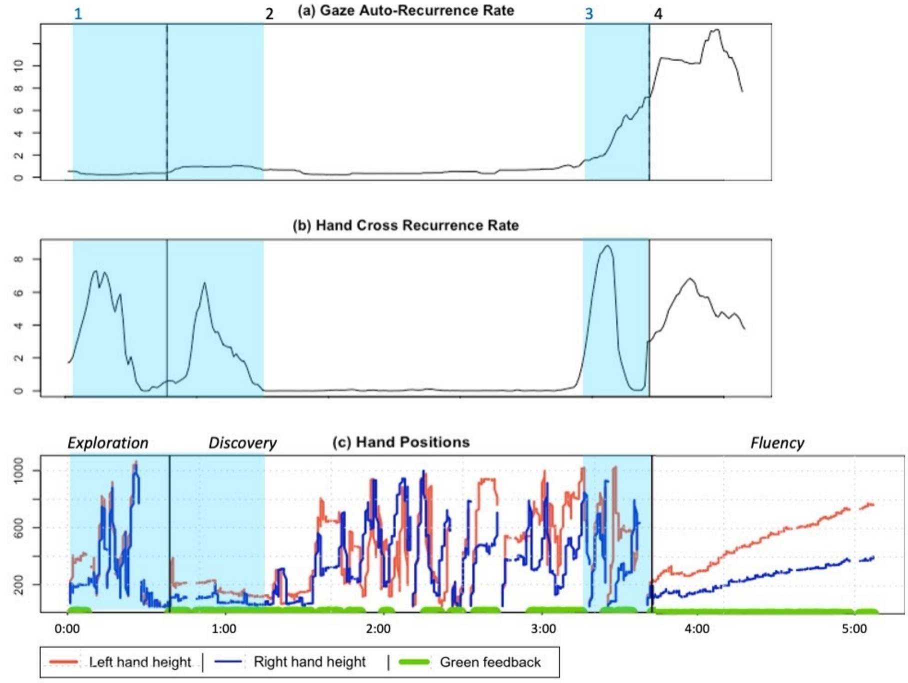


Fig. 1. Finn’s activity during the Fluency phase. (a) Finding an initial green position. (b-c) Gradually moving up the screen, adjusting to keep the bars green. Gaze patterns include the midpoint of the right bar.

Hand-Gaze Correlation. From Exploration to Discovery to Fluency, hand heights and gaze height became increasingly correlated. Initially during the Exploration phase, there was no correlation between hand heights and gaze height (left hand: r(225) = -0.002, p = 0.979; right hand: r(233) = -0.017, p = 0.792). During the Discovery phase, there was a moderate correlation between hand heights and gaze height (left hand: r(1427) = 0.360, p<0.001, right hand: r(1436) = -0.017, p <0.001).

During the Fluency phase, there was a high correlation between hand heights and gaze height (left hand: $r(802) = 0.697$, $p<0.001$; right hand: $r(803) = 0.069$, $p<0.001$). Variance in hand and eye movements became more related over the course of the task.

Recurrence Rate. Gaze auto-recurrence rate indicates the frequency with which states (here, specific areas of interest such as the top of a bar) are repeated. Hand cross-recurrence rate indicates the degree of coupling between the right- and left-hand heights. Finn's trajectory showed a high-low-high pattern of hand coupling (Fig. 5b). An initially stable coordination pattern (5b-1) destabilized for a period of low coupling (Fig. 5b-2) before culminating in a new stable coordination pattern (Fig. 5b3-4). In contrast to the hands, the gaze started out with low recurrence and maintained this throughout the Discovery phase (Fig. 5a1-2). At the end of the Discovery phase (section 3), both gaze recurrence rate and hand coupling[3] increased sharply, yielding high recurrence in both modalities in the Fluency phase (Fig. 5a-4, 5b-4).



---

[3] The brief drop in hand coupling at the end of this stage is likely a by-product of the participant attempting to generalize the ratio to an inverted position where the left hand was above the right.

Fig. 1. Recurrence rate over time for gaze and hands. (a) Gaze auto-recurrence rate starts low and increases at the end of the Discovery phase. (b) Hand cross-recurrence starts high, drops during Discovery, and increases at the end of the Discovery phase.

Determinism. Determinism captures the predictability of a system. Finn's hands began highly deterministic (Fig. 6b1-2), became less deterministic through the Discovery phase (6b2-4), and then once again reached high levels of determinism (Fig. 6b-5). During most of these first two phases, Finn's learning trajectory showed distinct determinism patterns in hand and gaze. During phases when hand determinism was high (Fig.6-2), gaze determinism was low, and vice versa (Fig. 6-3). These divergent dynamics resolved in section 4 as both the hands and gaze exhibited a decrease in determinism (Fig. 6-4), followed by high levels of determinism through to the end of the task (Fig. 6-5).

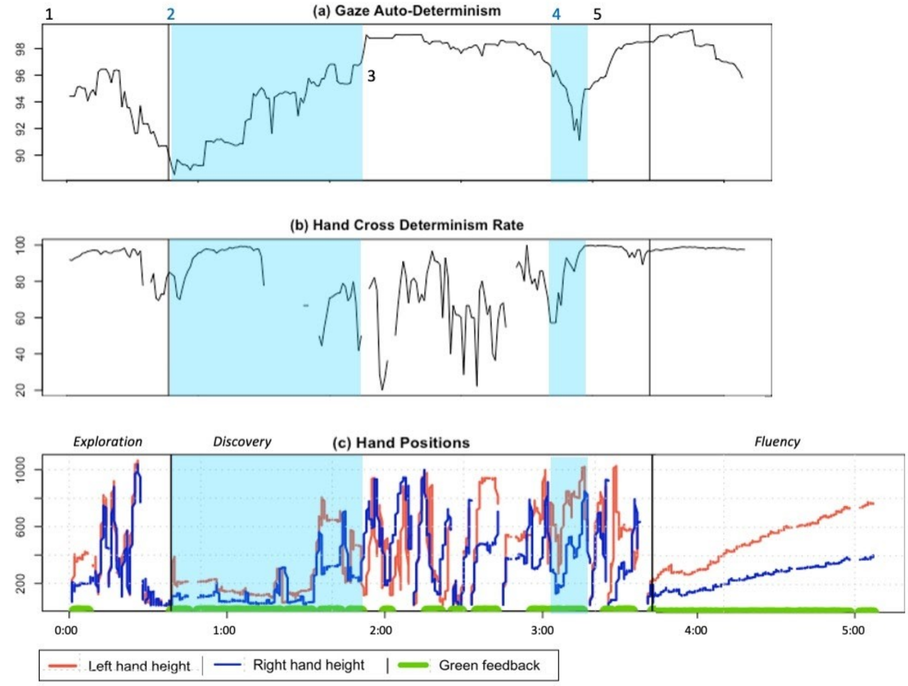


Fig. 1. Determinism Over time for gaze and hands. (a) Gaze determinism initially drops towards the end of Exploration, alternately drops and increases throughout the Discovery phase, and increases at the end of the Discovery phase. (b) Hand determinism begins high in Exploration and early Discovery, becomes lower and more variable through mid-Discovery, and increases to a consistently high rate at the end of Discovery.

Meanline. The meanline RQA metric captures the stability of a system. For the gaze auto-RQA, this metric reflects the mean duration of repeated sequences. For the hand cross-RQA, this metric reflects the mean duration of coordinated bimanual actions.

Hands and gaze started off with high stability (Fig. 7-1). As the participant discovered greens and lingered in them (Fig. 7-2), the gaze destabilized. As the participant attempted to move-in-green, the gaze became more stable once again (Fig. 7-3) with hand coupling becoming less stable. Just prior to the onset of Fluency, hand-hand meanline spiked, marking the beginning of a new stability (Fig. 7-4) lasting throughout the Fluency phase (Fig. 7-5). Gaze meanline also increased at this time, much more moderately (Fig. 7a-4).

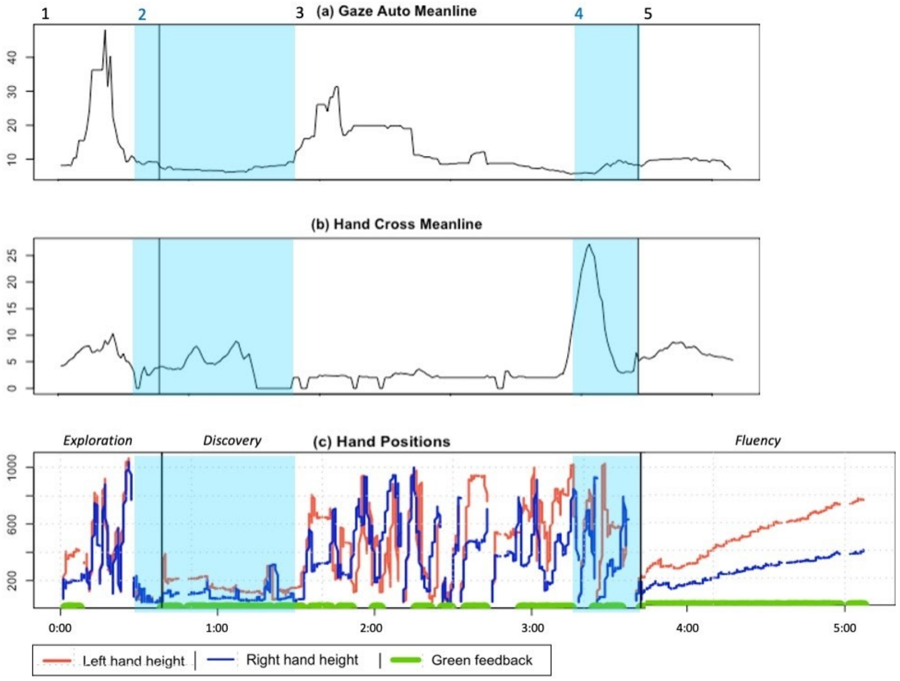


Fig. 1. Meanline over time for gaze and hands. (a) Gaze meanline starts high, drops at the onset of the Discovery phase, increases and decreases through mid-Discovery, and increases slightly at the end of Discovery leading into Fluency. (b) Hand meanline starts relatively high, drops mid-Discovery phase, spikes sharply at the end of Discovery, and remains relatively high in Fluency.

4 Discussion

This case study participant exhibited high levels of cross-hand recurrence rate, determinism, and meanline at the start and end of the time series, with lower levels in between (Figure 5b, Figure 6b, Figure 7b). The evolution from a predictably-coupled, stable state through a less-stably-coupled and less-deterministic phase, through to another predictably-coupled, stable state suggests a process of phase transition. Rather than a gradual increase, the participant's bimanual coordination changed suddenly in what Kostrubiec et al. (2012) might term a bifurcation. In contrast to shift, where a system maintains the same overall configuration, but with a slight pattern change, bifurcation is an abrupt reconfiguration. The right-hand–left-hand system shifts from one dynamically-stable coordination-pattern attractor to another, not unlike a horse breaking from a canter into a gallop. However, this transition included a lengthy transitional period of relative instability.

In contrast to the hands, gaze data does not initially exhibit high recurrence or determinism (Figure 5a, Figure 6a). During the Exploration stage, the case-study participant showed low gaze recurrence, suggesting that his eyes visited varied locations. This finding is consistent with the Gibsonian view that the function of exploratory movement is to expose new possibilities for action (J. Gibson, 1966), a process of information-detecting behavior that then supports perception of affordances for action attuned to local conditions (Adolph, 2019). Here, the gaze would be offering new possibilities for action through exploration of a variety of locations relative to the hands.

During the Discovery phase, the case study shows a destabilization of the initial system as the participant begins to find greens. Finding green introduced a new constraint on movement. As a result, the participant exhibited alternating meanline and determinism dynamics in the two focal modalities; gaze determinism was high when bimanual determinism was low, and vice versa (Figure 6a and b) and gaze meanline was high when bimanual meanline was low (Figure 7 and 7b). Towards the end of the Discovery phase, a coordination of coordinations (Piaget, 1970) develops wherein the coordination between the left- and right hands in turn enters into coordination with newly developed gaze structures spanning different screen locations. This recalls Piaget's theory of reflective abstraction, whereby higher-order knowledge arises from observations on lower-order coordinated actions and reorganizing them into new coordinations (Piaget, 1970; see also Abrahamson et al., 2016). Here, higher-order perceptual structures emerge through coordination between modalities that in turn informs coordination of each modality.

Although gaze does not share hand coordination's initial stability, both hand and gaze synchronously stabilize at the onset of the Fluency phase. The convergent increases in determinism across hands and gaze show an emergent meta-coordination that exhibited high predictability. The co-occurrence of increased gaze recurrence and increased hand coupling in this case study suggests a mutual influence of gaze pattern and hand-–hand coupling. Only when the gaze pattern became more consistent did the hands recouple into a multiplicative pattern. The eyes became more deterministic and consistent only as the whole system solved the movement problem. The participant's solution was not simply multimodal, attained in each modality separately; it was intermodal, with hands informing gaze, and gaze informing hands[4]. Only through higher-order coordination did the participant fully solve the movement problem posed by the task to move their hands smoothly in green. These results show that the phase transition arising in hand coordination fits within a larger dynamic. Whereas the initial hand–hand coupling was intramodal, the ultimate hand–hand–gaze coupling

---

[4] Although not visual or kinesthetic, it is worth noting that the auditory modality also played a role here as the tutor offered prompts and encouragements that solicited changes in dynamics, most notably by encouraging the participant to try moving-in-green.

that emerged was intermodal. This suggests that learning to perceive the central ratio variable involves intermodal coordination of vision and kinesthesia. Consistent with dynamic system theory, the impact of hand and gaze changes is not additive; instead, we observe hands-informing-eyes-informing-hands iteratively, reciprocally, intermodally. The emergent movement is greater than the sum of its individual gaze and hand components, where each modality's pattern is impacted by the other, giving rise to a qualitatively new movement form of smooth multiplicative movement. The process of specifying information detection that supports enactment of the target bimanual coordination pattern spans the kinesthetic and visual systems in these participants.

One interpretation of the emergence of this new coordination pattern would be that coupling between independent modalities gives rise to an emergent perceptual composite that guides motor action. From this view, multiplicative movement is the result of self-organization whereby sensorimotor couplings that promote task performance are maintained. Here, students achieve the enactment of a pre-specified movement form that dynamically elicits positive feedback (green) from an interactive technological device by maintaining an invariant relation between manually instantiated spatial magnitudes.

Each motor skill has its own problem space characterized by particular affordance relations (J. Gibson, 1966). The MIT-P design has at its heart a Gibsonian higher-order variable (J. Gibson, 1966) of ratio, which participants learn to perceive through action. Although the MIT-P activity instantiates ratio in a situated context, ratio conceptually is scale-free, transcendent beyond specific spatial magnitudes. Following Gibson, we conjecture that the attentional anchor perceptual structure is in fact not a visual-specific structure, but a spatial–relational "amodal" one, perceivable through different modalities. Thus, whereas ratio's detection here is specified by information in the visual–kinesthetic array, it could also be instantiated using other modalities in the context of other instructional designs.

Some more recent ecological psychology work has challenged Gibson's view of multiple, overlapping perceptual systems, arguing instead for a single irreducible supermodal perceptual system (Stoffregen & Bardy, 2001; Stoffregen et al., 2017). In this view, direct perception is a characteristic of a global array made up of the irreducible superordinate structures across information available through the different sense organs. Here, perceptual systems function relationally and cannot be understood independently. Behavior simultaneously affects multiple ambient energy arrays, and relations between different ambient energy forms provide critical information about the animal–environment interaction. Supermodal structures exhibit emergent properties, more than the sum of their sensory system parts. In this view, learning involves perceptual–motor differentiation of relevant structures in a global array spanning multiple forms of ambient energy and determining which referents are relevant. For example, learning to somersault involves discovery and control of the relations between vestibular, biomechanical, and optical patterns of energy. From this view, embodied designs like the MIT-P create the conditions in the global array for the detection of supermodal structures. From this perspective, learning to move multiplicatively in the MIT-P would be construed as a process of perceptual–motor differentiation of emergent structures already in the global array through discovery and control of emergent relational properties across the visual and kinesthetic sensory organs. The onset of fluency would be the result of supermodal direct perception of irreducibly superordinate invariant structures.

In general, skill-learning processes furnish attentional anchors available to deploy in novel contexts. Per Bernstein, learning develops new capacities that can be deployed in situations that differ along multiple parameters (Bernstein, 1996). According to E. Gibson (1969), distinctive modality-specific properties like color can inhibit intermodal transfer. An open question is how qualities specific to the visual modality such as color might affect learners' capacity to generalize or extend the ratio concept. In this activity, the color green begins as a task objective, then becomes task feedback, eventually taking on further meaning as a conceptual placeholder in reference to the new type of relation discovered through the activity (Abrahamson et al., 2011). Later in the activity, as the multiplicative mathematical expression is introduced, green fades from its discursive role as redundant to broader mathematical inquiry. Research on perceptual learning suggests that differences in learning pathway impact the adaptability of a coordination pattern attractor to novel situations (Yamamoto et al., 2020). Generalization of learning-through-the-MIT-P is a theory-generative direction for further empirical work.

A meaningful context for continued exploration of these questions are versions of the MIT-P activity designed for blind learners that use haptic and/or auditory feedback (Abrahamson et al., 2019; Tancredi et al., in press). Congenitally blind individuals have been found to have differently organized occipital lobe processing, processing spatial information using brain areas that respond to visual input in sighted individuals (e.g. Burton, 2003). In addition to kinesthetic, proprioceptive, and tactile information, auditory information is also used spatially. J. Gibson wrote,

> Visual proprioception, more than any other kind, is important for the guidance of purposive actions in the surrounding world. In an habitual motor skill, one can fall back on the feeling of movement for guidance but in any new task one needs visual control (J. Gibson, 1966, pp. 37–38).

Work with blind learners can investigate alternate pathways to achieve motor skill together with what E. Gibson described as intermodal equivalence. How might the dynamics of the activity in these versions compare and contrast with the color-feedback visual–kinesthetic version, for sighted versus congenitally blind learner, acknowledging the different cognitive architectures arising from different lifelong modal experiences?

This chapter presents a nascent foray into intermodality grounded in a first case study. Limitations of this particular case include that the participant exhibited lower levels of fluent movement-in-green at the end of this part of the interview than other participants. It is likely that gaze stability and the coupling between hands and gaze would develop further in future stages of the interview. These dynamics warrant examination in a greater number of participants. Further work within and beyond the MIT-P can deepen understanding of these issues through inter-participant analyses quantifying the relationship between gaze and hand dynamics.

5 Conclusion

We observed a case study in which a student progressed through exploration, discovery, and fluency in an embodied design context. Reaching fluency in this case was equivalent to reaching intermodal coordination: the stabilization of new gaze patterns together with smooth bimanual movement. These analyses shed light on the process of perception-for-action by which perceptual structures emerge through and for coordinating bimanual actions. Exploratory movements functioned to generate possibilities for action, leading to discovery of new intermodal, if not supermodal, affordances for action not available to the independent modalities individually. In this environment, the learner assembled his body into a tool for solving the problem, coordinating hands and gaze into one dynamically stable assemblage. These results suggest that multimodal data streams are more than mere entry points into a cognitive phenomenon, and are actively interacting, mutually influencing components of a larger dynamical system. Beyond multimodality, these data reveal conceptual learning as intermodal coordination. Rather than a centrally controlling, modality-agnostic concept nurturing from each modality, a parsimonious explanation for our results is that what we call learning is constituted by the self-organized coordination of different modalities. Interacting with the embodied design environment effectively cultivated the perception of an intermodally-invariant, relational structure through the soft-assembly (Richardson & Chemero, 2014) of a temporary hand–gaze coalition into a coupled relation instantiating multiplicative reasoning.

The study bears multiple implications for educational practice and research. For educational practice, the present study supports the thematic conjecture of action-based embodied design (Abrahamson, 2014), namely that task- and environment design can elicit novel, conceptually-salient perceptuomotor dynamics. For educational research, the study creates an auspicious empirical platform for heady philosophical and theoretical deliberations over foundational metaphysical framings of the mathematical learning process. More broadly, we welcome nuanced conversations with researchers following the various flavors of 4E cognition, including enactivism and ecological psychology (qv. Di Paolo et al., 2021), over explanatory models that best account for our empirical findings and how these accounts might enhance our educational design for intersectionally diverse students (Tancredi et al., in press). A consistent focus of research remains the relation between enactment, perception, language, symbols, and concepts. How do the various 4E theories model the educational process by which students come to reason mathematically about their enactment? More specifically, what would be a 4E account for the semiotic micro-process by which symbolic artifacts take on enactive meanings? For partial accounts, readers are referred to our ongoing publications (e.g., Abrahamson, 2019; Abrahamson et al. 2011; Abrahamson et al., in press).

We propose that RQA is a useful method for ongoing work on intermodality. RQA's sensitivity to complex idiosyncratic dynamics at multiple, nested, sequentially dependent interacting timescales supports the research process of evaluating, deepening, and elaborating theory. RQA also provides a means to map the application of coordination-dynamics constructs onto messy processes such as tablet-based mathematics learning. Learning data, particularly in the context of embodied design, provides a rich context that engages both fundamental biological perception–action processes and cultural forms, from mathematical symbols to language. Our results suggest the traction of RQA on questions of intermodal coordination. Some future directions for this work include cross-participant analyses of the intermodal dynamics in multimodal systems, work with participants with a greater diversity of sensorimotor profiles, and work expanding the unit of intermodal analysis to incorporate the tutor-tutee dyad, spanning hand, gaze, language, and tutor movements. In turn, such analyses can feed back into the design of responsive embodied-interaction settings for mathematics education, and perhaps beyond.

References

1. Abdu, R., Tancredi, S., Abrahamson, D., & Balasubramaniam, R. (under review). A complex systems outlook on hand-eye coordination in mathematical learning. In M. Schindler, A. Shvarts, & A. Lilienthal. (Eds.), Eye-tracking research in mathematics education [Special issue]. *Educational Studies in Mathematics.*
2. Abdullah, A., Adil, M., Rosenbaum, L., Clemmons, M., Shah, M., Abrahamson, D., & Neff, M. (2017). Pedagogical agents to support embodied, discovery-based learning. In J. Beskow, C. Peters, G. Castellano, C. O'Sullivan, I. Leite, & S. Kopp (Eds.), *Proceedings of 17th International Conference on Intelligent Virtual Agents (IVA 2017)* (pp. 1–14). Springer International Publishing. https://doi.org/10.1007/978-3-319-67401-8_1
3. Abrahamson, D. (2014). Building educational activities for understanding: An elaboration on the embodied-design framework and its epistemic grounds. International Journal of Child-Computer Interaction, 2(1), 1–16. https://doi.org/10.1016/j.ijcci.2014.07.002
4. Abrahamson, D., Flood, V. J., Miele, J. A., & Siu, Y.-T. (2019). Enactivism and ethnomethodological conversation analysis as tools for expanding Universal Design for Learning: The case of visually impaired mathematics students. *ZDM Mathematics Education, 51*(2), 291–303. https://doi.org/10.1007/s11858-018-0998-1
5. Abrahamson, D., & Sánchez-García, R. (2016). Learning is moving in new ways: The ecological dynamics of mathematics education. *Journal of the Learning Sciences, 25*(2), 203–239. https://doi.org/10.1080/10508406.2016.1143370
6. Abrahamson, D., Lee, R. G., Negrete, A. G., & Gutiérrez, J. F. (2014). Coordinating visualizations of polysemous action: Values added for grounding proportion. *ZDM Mathematics Education, 46*(1), 79–93.
7. Abrahamson, D., Shayan, S., Bakker, A., & Van der Schaaf, M. F. (2016). Eye-tracking Piaget: Capturing the emergence of attentional anchors in the coordination of proportional motor action. *Human Development, 58*(4–5), 218–244.
8. Abrahamson, D., & Trninic, D. (2011). Toward an embodied-interaction design framework for mathematical concepts. In P. Blikstein & P. Marshall (Eds.), *Proceedings of the 10th Annual Interaction Design and Children Conference (IDC 2011)* (Vol. "Full papers", pp. 1–10). IDC.
9. Abrahamson, D., Trninic, D., Gutiérrez, J. F., Huth, J., & Lee, R. G. (2011). Hooks and shifts: A dialectical study of mediated discovery. *Technology, Knowledge, and Learning, 16*(1), 55–85.
10. Adolph K, E (2019). An ecological approach to learning in (not and) development. *Human Development*, *63,* 180–201. https://doi.org/10.1159/000503823
11. Bernstein, N. A. (1996). Dexterity and its development. In M. L. Latash & M. T. Turvey (Eds.), (pp. 3–235). Lawrence Erlbaum Associates.
12. Carvalho, E. M., & Rolla, G. (2020). An enactive-ecological approach to information and uncertainty. *Frontiers in Psychology 11*, 588. https://doi.org/10.3389/fpsyg.2020.00588.
13. Collins, A. (1992). Towards a design science of education. In E. Scanlon & T. O'shea (Eds.), New directions in educational technology (pp. 15–22). Springer.
14. Duijzer, A. C. G., Shayan, S., Bakker, A., van der Schaaf, M. F., & Abrahamson, D. (2017, 2017-February-08). Touchscreen tablets: Coordinating action and perception for mathematical cognition. In J. Tarasuik, G. Strouse, & J. Kaufman (Eds.), Touchscreen tablets touching children's lives [Special issue] [Original Research]. *Frontiers in Psychology, 8*(144). https://doi.org/10.3389/fpsyg.2017.00144

15. Ernst, M. O. (2008). Multisensory integration: a late bloomer. *Current Biology*, *18*(12), R519-R521.
16. Fleuchaus, E., Kloos, H., Kiefer, A. W., & Silva, P. L. (2020). Complexity in science learning: Measuring the underlying dynamics of persistent mistakes. *Journal of Experimental Education, 88*(3). https://doi.org/10.1080/00220973.2019.1660603
17. Flood, V. J., Shvarts, A., & Abrahamson, D. (2020). Teaching with embodied learning technologies for mathematics: Responsive teaching for embodied learning. *ZDM Mathematics Education, 52*(7), 1307–1331. https://doi.org/10.1007/s11858-020-01165-7
18. Gibson, E. J. (1969). *Principles of perceptual learning and development.* Meredith Corporation.
19. Gibson, J. J. (1966). *The senses considered as perceptual systems.* Houghton Mifflin.
20. Gowlland, G. (2019). The sociality of enskilment. *Ethnos, 84*(3), 508–524. https://doi.org/10.1080/00141844.2018.1455726
21. Hutto, D. D., & Sánchez-García, R. (2015). Choking RECtified: Embodied expertise beyond Dreyfus. *Phenomenology and the Cognitive Sciences, 14*(2), 309–331. https://doi.org/10.1007/s11097-014-9380-0
22. Kostrubiec, V., Zanone, P.-G., Fuchs, A., & Kelso, J. A. S. (2012). Beyond the blank slate: Routes to learning new coordination patterns depend on the intrinsic dynamics of the learner—experimental evidence and theoretical model. *Frontiers in Human Neuroscience, 6*. https://doi.org/10.3389/fnhum.2012.00222
23. Marwan N., Romano M.C., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports, 438*, 237–239.
24. Mechsner, F. (2003). Gestalt factors in human movement coordination. *Gestalt Theory, 25*(4), 225–245.
25. Mechsner, F. (2004). A psychological approach to human voluntary movements. *Journal of Motor Behavior, 36*(4), 355–370.
26. Ou, L., Andrade, A., Alberto, R. A., Bakker, A., & Bechger, T. (2020). Identifying qualitative between-subject and within-subject variability: A method for clustering regime-switching dynamics. *Frontiers in Psychology*, *11*, 1136. https://doi.org/10.3389/fpsyg.2020.01136
27. Ou, L., Andrade, A., Alberto, R., van Helden, G., & Bakker, A. (2020). Using a cluster-based regime-switching dynamic model to understand embodied mathematical learning. In K. Verbert, M. Scheffel, N. Pinkwart, & V. Kovanonic (Eds.), *Proceedings of the 10th International Conference on Learning Analytics & Knowledge* (pp. 496–501). ACM. https://doi.org/10.1145/3375462.3375513
28. Pardos, Z. A., Hu, C., Meng, P., Neff, M., and Abrahamson, D. (2018). "Classifying learner behavior from high frequency touchscreen data using recurrent neural networks," in *UMAP'18 Adjunct: 26th Conference on User Modeling, Adaptation and Personalization Adjunct* (pp. 317–322). ACM. https://doi.org/10.1145/3213586.3225244
29. Piaget, J. (1970). Genetic epistemology (E. Duckworth, Trans.). Columbia University Press.
30. Richardson, M. J., & Chemero, A. (2014). Complex dynamical systems and embodiment. In L. Shapiro (Ed.), *The Routledge handbook of embodied cognition* (pp. 39-50). Routledge.
31. Shayan, S., Abrahamson, D., Bakker, A., Duijzer, A. C. G., & Van der Schaaf, M. F. (2015). The emergence of proportional reasoning from embodied interaction with a tablet application: An eye-tracking study. In L. Gómez Chova, A. López Martínez, & I. Candel Torres (Eds.), Proceedings of the 9th International Technology, Education, and

Development Conference (INTED 2015) (pp. 5732–5741). International Academy of Technology, Education, and Development.

32. Shvarts, A., & Abrahamson, D. (2019). Dual-eye-tracking Vygotsky: A microgenetic account of a teaching/learning collaboration in an embodied-interaction technological tutorial for mathematics. *Learning, Culture and Social Interaction, 22*, 100316. https://doi.org/10.1016/j.lcsi.2019.05.003
33. Steffe, L. P., & Kieren, T. (1994). Radical constructivism and mathematics education. *Journal for Research in Mathematics Education, 25*(6), 711–733.
34. Stephen, D. G., Dixon, J. A., & Isenhower, R. W. (2009). Dynamics of representational change: Entropy, action, and cognition. *Journal of Experimental Psychology: Human Perception and Performance, 35*(6), 1811–1832.
35. Stoffregen, T. A., & Bardy, B. G. (2001). On specification and the senses. *The Behavioral and Brain Sciences, 24*(2), 195–261. https://doi.org/10.1017/s0140525x01003946
36. Stoffregen, T. A, Mantel, B., & Bardy, B. G. (2017). The senses considered as one perceptual system. *Ecological Psychology, 29*(3), 165–197. https://doi.org/10.1080/10407413.2017.1331116
37. Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. MIT Press.
38. Tancredi, S., Abdu, R., Abrahamson, D., & Balasubramaniam, R. (2021). Modeling nonlinear dynamics of fluency development in an embodied-design mathematics learning environment with Recurrence Quantification Analysis. *International Journal of Child-Computer Interaction*, 100297. https://doi.org/10.1016/j.ijcci.2021.100297
39. Tancredi, S., Chen, R. S. Y., Krause, C., & Siu, Y.–T. (in press). The need for SpEED: Reimagining accessibility through Special Education Embodied Design. In S. L. Macrine & J.M. Fugate (Eds.), *Movement matters: How embodied cognition informs teaching and learning.* M.I.T. Press.
40. Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. MIT Press.
41. von Glasersfeld, E. (1987). Learning as a constructive activity. In C. Janvier (Ed.), *Problems of representation in the teaching and learning of mathematics* (pp. 3–18). Lawrence Erlbaum.
42. Vygotsky, L. S. (1926/1997). Educational psychology (R. H. Silverman, Trans.). CRC Press LLC. (Original work published 1926)
43. Webber, C. L., & Zbilut, J. P. (1994). Dynamical assessment of physiological systems and states using recurrence plot strategies. Journal of Applied Physiology, 76(2), 965–973. *10*(1), 3058. https://doi.org/10.1038/s41598-020-60066-7
44. Worsley, M., Abrahamson, D., Blikstein, P., Bumbacher, E., Grover, S., Schneider, B., & Tissenbaum, M. (2016). Workshop: Situating multimodal learning analytics. In C.-K. Looi, J. L. Polman, U. Cress, & P. Reimann (Eds.), "Transforming learning, empowering learners," Proceedings of the International Conference of the Learning Sciences (ICLS 2016) (Vol. 2, pp. 1346–1349). International Society of the Learning Sciences.
45. Yamamoto, K., Shinya, M., & Kudo, K. (2020, 2020/02/20). The influence of attractor stability of intrinsic coordination patterns on the adaptation to new constraints. *Scientific Reports,* 3058. https://doi.org/10.1038/s41598-020-60066-7