It's Our Move: Mathematical Perception Emerges for Coordinating Joint Action

Dor Abrahamson[1,2], Sofia Tancredi[3], Andrew Xiao[1], Xiaoran Li[4], Michał Weiss[5],

Katarzyna Potega vel Żabik[6], and Justin K. Dimmel[7]

[1]Berkeley School of Education, University of California, Berkeley, USA
[2]Department of Mathematics Education, College of Education, Korea University, South Korea
[3]Wisconsin Center for Education Research, University of Wisconsin–Madison
[4]School of Linguistics and Language Resources, Beijing Language and Culture University, China
[5]Faculty of Psychology, University of Warsaw, Poland
[6]Copernicus Science Centre, Warsaw, Poland
[7]College of Education and Human Development & Maine Center for Research in STEM Education, University of Maine, USA.

Author Note

This study was not sponsored. It has not been presented or published elsewhere. We have no conflicts of interest to disclose. The authors wish to thank Hermann Badowski for his engineering expertise in building OЯTHO at the Copernicus Science Centre and working with us to reconfigure game conditions for this study. We are grateful to Przemyslaw Biecek and his students for their insightful analyses and visualizations of the data corpus, which informed our methodologies and theorizations.

Correspondence concerning this article should be addressed to Dor Abrahamson, Berkeley School of Education, 2121 Berkeley Way West, Berkeley, CA 94720-1670.
Email: dor@berkeley.edu

It's Our Move: Mathematical Perception Emerges for Coordinating Joint Action

ABSTRACT. Can children learn basic mathematical concepts through playing a game? OЯTHO is a two-player game implemented on a large multitouch digital tabletop. Players must cooperate to navigate a digital ball through a maze. One player, positioned along the table's length, controls the horizontal position of the ball, as the other player, positioned along the table's width, controls its vertical position. The game was designed to foster the development of *coordinate perception*: an acquired capacity to attend to objects as located in space at points specifiable by linearly independent axes. Theoretically, the game is conceptualized as steering players' gradual transition from situated sensorimotor entrainment to grounded mathematical practice. To evaluate the game's rationale, we randomized two of its aspects—(1) the sequencing of game conditions (n=254), and (2) track features players had to overcome (n=1,218)—testing their impact on young players' performance on an in-game quantitative coordinate task. We found that a sensorimotor-first sequence and encountering decision points most improved players' coordinate fluency. We offer a set of heuristic principles for designing collaborative games fostering mathematical perception, and we speculate on the prospects of integrating such games into mainstream education as conceptual entries to curricular modules.

Keywords: digital tabletop, Enactivism, joint action, mathematics education, perception

Public Significance. Educational institutions, such as schools, libraries, and community centers, are enjoying increased access to interactive technological platforms (tablets, tabletops, smartboards, etc.). Can children learn mathematics through playing games together on these platforms? Educational researchers found that, yes, carefully designed activities, where children manipulate digital objects together, may give rise to new intuitions that become mathematical understandings.

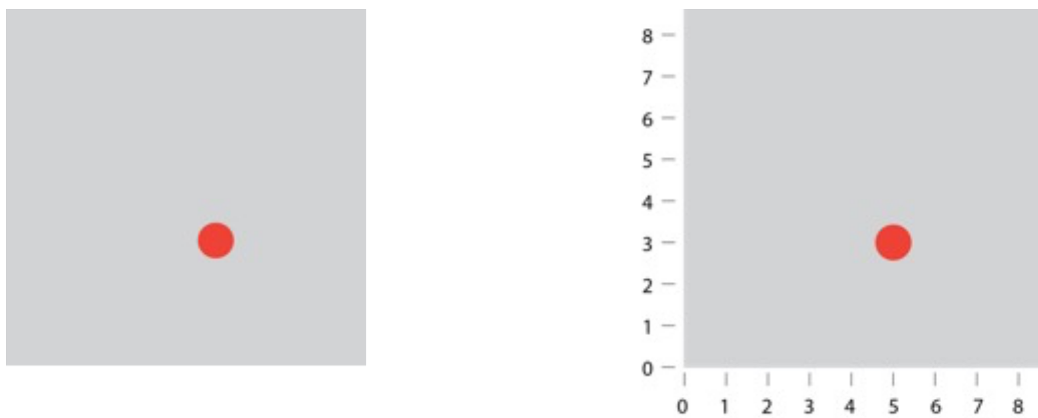
It's Our Move: Mathematical Perception Emerges for Coordinating Joint Action

Introduction

Consider the circle in Figure 1a. We attend to it as a standalone entity. We *could* note its location with respect to the height and width of the field or perhaps with respect to the middle of the field. But that might not be our first inclination. Now consider the circle in Figure 1b. Something in our attention shifts. We apprehend this object as bearing spatial affinity to two scalar values lying on orthogonal axes, respectively 5 and 3. In doing so, our eyes saccade vertically and horizontally along invisible auxiliary trajectories back and forth to those indices, evaluating the precision of our visual assessment. Moreover, the object now bears semiotic portent—it's supposed to mean something: it's not just a circle, it's a point. The two numerical values, and the entire diagrammatic system, serve as a frame of reference indexing the point's location and, in so doing, signifying its meaning as implicating information respecting an apparent relation among those numerical values. We are now apprehending the point as embedded in a Cartesian system. One might say we are exercising a particular form of spatial attention—*coordinate perception*—that configures an element in the field as cardinally triangulated, a bi-axial composite.

Figure 1

Coordinate Perception as an Acquired Cross-Product Spatial Orientation to Points on the Plane



a. *b.*

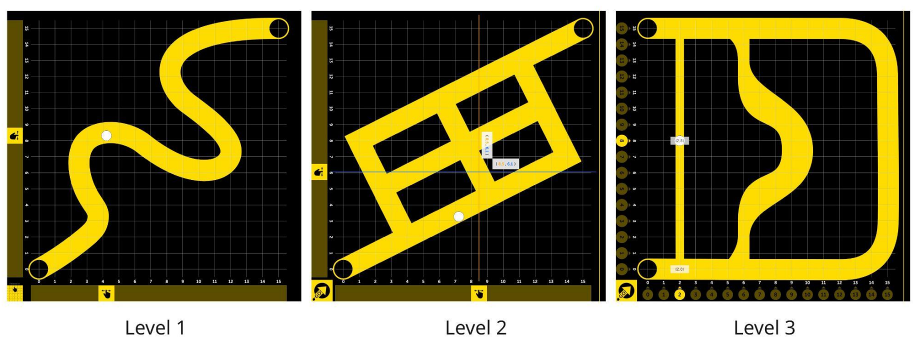
Note. a. A circle lying on the plane; *b.* The same circle, now perceived as a data point expressing the intersection of two scalar numerical values (5, 3) lying respectively on the *x*-axis and *y*-axis.

Seeing the point in Figure 1b as a bi-axial composite of orthogonal indices is not a genetic inclination, but rather a skilled professional practice pertaining to a historical artifact, the Cartesian system. As such, coordinate perception needs to be learned. Typically, sensorimotor routines for making sense of informational artifacts, such as the Cartesian system, are mediated through cultural epigenesis, whereby more-knowledgeable others recruit multimodal semiotic resources, such as speech and gesture, to regulate novices' perceptual behavior (Shvarts, 2018). As such, coordinate perception is a form of educated perception (Goldstone et al., 2010).

Can children develop such coordinate perception without explicit instruction? If not through direct instruction, then what might steer a young child to perceive a point on the plane as located at the intersection of two imaginary crosshairs emanating from orthogonal axes? From an ecological-psychology perspective, what “field of promoted action” (Reed & Bril, 1996) could catalyze the emergence of coordinate perception as the child’s spontaneous cognitive solution to a problem of enacting some goal movement (Abrahamson, 2014)? Could this learning process, moreover, be an exciting experience, perhaps playful?

Figure 2

OЯTHO: A Joint-Action Game for Developing Coordinate Perception of the Plane



Note. OЯTHO’s three maze-like levels progress from controlling the ball continuously, by sliding the *x*- and *y*-controls, to controlling it with a series of coordinates.

By creating *OЯTHO* (Potega vel Żabik et al., 2024), a two-person “digital-minigolf” game, we sought to intrinsically motivate players’ development of coordinate perception as their practical solution to the emergent problem of co-enacting axially distributed movement, where one player manipulates only the ball’s *x*-trajectory and the other player manipulates only its *y*-trajectory (see Figure 2; reminiscent of Etch-a-Sketch controls). We synthesize literatures of enactivist cognition (Varela et al., 1991) and joint social action (Sebanz et al., 2006) to theorize that: (1) mathematical knowledge is rooted in perceptuomotor skill (Hutto & Abrahamson, 2022); (2) new perceptual orientations guiding action emerge as solutions to motor-control problems (Mechsner et al., 2001); (3) educational settings can be created whereby motor-control problems give rise to new perceptuomotor skill undergirding the cognition of curricular mathematical concepts, such as proportion, trigonometric relations, or linear functions (Alberto et al., 2022; Wei et al., 2025); (4) mathematical concepts can be socially distributed across designed joint-action spaces (Abrahamson & Wilensky, 2007); (5) and, therefore, joint-action motor-control settings, too, may likewise give rise to proto-mathematical skill. More broadly, we draw on enactivist epistemology to frame learning as an emergent state of dynamic equilibrium among intentional agents applying cultural–historical tools to perform adaptive practices situated within specific environments (Shvarts & Abrahamson, 2023).

Joint action is broadly defined as two or more actors working together to achieve a shared goal (Sebanz & Knoblich, 2021). Prior work has established that joint actors are highly able to inter-regulate, mutually coordinate, and fine-tune their respective movements to complete tasks in shared spaces. For example, Török and colleagues (2019) found that people working together to steer a ball through an obstacle course would choose path segments that minimized the total length of route through the course, even if that meant one player had to navigate a longer segment (see also Schmitz et al., 2017, on *simultaneous* navigation). The design of OЯTHO elaborates on this joint-action capacity. Our hypothesis was that a mathematical concept—Cartesian coordinates—could be instantiated in the form of a collaborative task, where individual-action affordances and joint-action constraints would induce the players' emergent understanding of the concept as they struggled to coordinate their actions.

This paper reports on our efforts to determine whether OЯTHO achieved its design objective, what game mechanics and any other activity elements of OЯTHO appear to have fostered coordinate perception, and how individuals' learning process may be shaped by their need to negotiate their joint action. In particular, the game has seven mazes (tracks) and three modes (levels) of play, and we were interested to understand whether and how the levels' game mechanics, level sequence, and track features contribute to developing coordinate perception. The first empirical report from our research project, this paper experimentally manipulated key aspects of gameplay to test the resulting impact on players' in-game performance with quantitative coordinates. To ground these results, we draw on qualitative analyses of audio–video and eye-tracking data from an illustrative player dyad in action to unpack the theorized learning mechanism at play.

Research Questions

We applied mixed methods to investigate empirically how learners can develop fluency with Cartesian coordinates through playing an action-based embodied design. Our research questions address whether the game impacts coordinate fluency (RQ1), what design features impact gains (RQ2), and how interactions with game features can give rise to fluency (RQ3):

1. **Does playing an axially distributed joint-action game change children's coordinate perception?**

Drawing on Enactivism's epistemology of mathematical cognition (Hutto & Abrahamson, 2022), we hypothesized that learning to move jointly (Level 1), then beginning to reference static positions (Level 2) together build capacity for children to use Cartesian coordinates effectively (Level 3, an in-game measure).

2. **What features of game encounters occasion opportunities for coordinate perception to emerge?**

Does greater exposure to track features that require either complex motor (sharp turns) or social (decision points) coordination enhance coordinate perception skills? We hypothesized that both motoric and social complexity will yield greater learning gains, because these coordination breakdowns will surface players' implicit sensorimotor entrainment to an explicit discursive plane, whereby cultural ontologies take form as cognitive structures (Maturana & Varela, 1992).

3. **What is the mediating function of spontaneous multimodal communication in players' epistemic shift from action to concept?**

Building on RQ2, we analyze qualitatively the cognitive function of interpersonal communication in bootstrapping sensorimotor entrainment into mathematical discourse. The multimodal semiotic resources players recruit in their efforts to regulate their fellow player's perceptuomotor orientation to the shared display, we hypothesized, give tangible form to covert sensations, establishing these utterances as consensual dyadic means of coordination (Kita et al., 2017).

Methods

Participants

After preprocessing, we obtained 254 dyads (1,143 trials) for RQ1 and 1,218 dyads (9,049 trials) for RQ2. Most dyads were familiar partners (RQ1: 93.7%, RQ2: 91.1%), with a small proportion of strangers (RQ1: 3.5%, RQ2: 4.9%) or missing reports (RQ1: 2.8%, RQ2: 4.0%). The mean age was 9.46 years (SD = 1.73) for RQ1 and 9.33 years (SD = 1.75) for RQ2, with average within-dyad age differences of 1.65 (SD = 1.66) and 1.67 years (SD = 1.61), respectively, a relatively homogeneous age composition.

For RQ3, we collected eye-tracking and video data on-site from 3 player dyads, using Tobii Glasses and fixed cameras, then selected one focal dyad for the qualitative case study.

Materials

OЯTHO is installed in a research-designated zone (LivingLab, see Figure 4) at the Copernicus Science Centre, in Warsaw, Poland. The installation runs on an 86'' multi-touch tabletop with a dedicated Windows 11 application and custom logging. Before play, visitors read an on-screen information notice, provide informed consent, and enter minimal demographics (age and relationship status). Participation proceeds without on-site supervision.

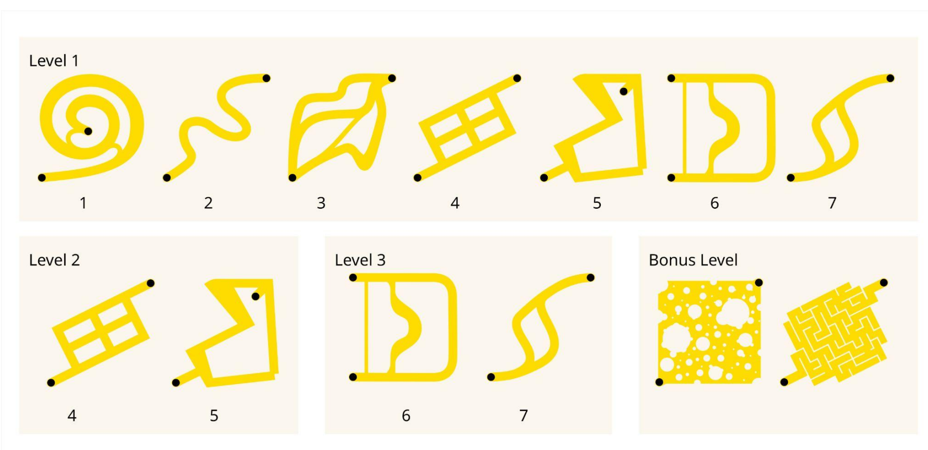
Figure 4

OЯTHO in the LivingLab Research Zone at the Copernicus Science Centre Exhibition

In the default version of OЯTHO, gameplay proceeds across three stages that vary the control scheme: continuous manual control (Level 1); crosshair placement with a “GO” command (Level 2); and numerical input of coordinate pairs with “GO” (Level 3, an in-game formative assessment). Each level contains a distinct set of mazes (“tracks”) featuring varying obstacles. For the purposes of this study, we launched a randomized version of the game on September 8, 2025, with seven tracks at Level 1, two at Level 2, and two at Level 3, followed by two bonus tracks with Level 1 mechanics that unlock after completing all three levels (see Figure 3).

Figure 3

OЯTHO Tracks



Note. In all tracks, the start is the black dot at the bottom-left of the path (0,0); the end is at the top-right of the path.

Data Collection

System logs include time-stamped touch events, button presses, coordinate values, object positions (sampled at 60 Hz), level/track identifiers, and state changes (e.g., game over).

In selected sessions, two Tobii Glasses 3 eye-trackers and fixed cameras, oriented towards the players and tabletop, capture synchronized gaze, video, and audio. These recordings are collected following prior scheduling and explicit consent.

We focused only on child–child (ages 3–12) dyad data. For RQ1, we conducted an experimental modification of the game from September to January 2026. For RQ2, the game was configured with a default level order and randomized tracks within each level, with data collected from June 2023 to December 2024. Only dyads with at least one record of Level 3 success were included, excluding data after this point. Trials with mid-game restarts were excluded.

Data Analysis

We adopted a stepwise procedure to construct regression models for both RQs. Initial models included all theoretically relevant predictors: game condition sequence (RQ1 baseline model), track features (RQ2 baseline model), track-related contextual factors, time-related measures, and age-related covariates. We then conducted hierarchical regression analyses by incrementally evaluating the impact of each set of predictors on the model to determine which to retain in the final model. The unit of analysis in the study is the dyad, and each row in the regression dataset represents one dyad. Regarding the construction of predictor variables (e.g., time_spent_level1 and time_spent_level2), these were computed by aggregating trial-level logs within each dyad.

RQ1

We conducted an experiment in which we randomized the order of levels that museum visitors encountered OЯTHO to test the impact of exposure to Levels 1 and 2 on performance in the Assessment Level (3).

We used a regression model for predicting fluency with coordinates (measured by speed in Level 3), controlling for the tracks players were assigned in Level 3 and all experiences received beforehand. Following hierarchical regression analyses (see Table S1 & S2), the final statistical model was specified as:

Model 1

$$\text{First\_Success\_level3\_time\_spent} = \beta_0 + \beta_1 \text{ level\_1} + \beta_2 \text{ level\_2} + \beta_3 \text{ levels 1-2} + \beta_4 \text{ levels 2-1} + \beta_5 \text{ level3\_track\_7} + \varepsilon$$

RQ2

To answer RQ2, we built a regression model predicting Assessment Level performance, using the same approach as RQ1, from the number of decisional and sharp-turn challenge points that dyads encountered in Level 1, controlling for the amount of time they spent before the first successful trial and tracks assigned for Level 2 and 3. Due to the design of the tracks and game, decisional area and narrow area appear simultaneously, which leads to high linear correlation between these two variables. Therefore, we only include the decisional region variable for this model.

Following hierarchical regression analyses (see Table S3 & S4), the final statistical model was specified as:

Model 2

$$\text{First\_Success\_level3\_time\_spent} = \beta_0 + \beta_1 \text{ sharp} + \beta_2 \text{ decision} + \beta_3 \text{ level2\_track5} + \beta_4 \text{ level3\_track7} + \beta_5 \text{ time\_spent\_level1} + \beta_6 \text{ time\_spent\_level2} + \beta_7 \text{ failed\_time\_spent\_level3} + \varepsilon$$

Descriptions of the variables in both models are presented in Table 1. For both RQs, data meets linear regression assumptions. Variance Inflation Factor (VIF) test values indicated no multicollinearity (>5). To further ensure the robustness of our findings, all regression models were estimated using robust standard errors (HC3). After applying a log transformation to the target variable, the residuals approximated a normal distribution (see Figure 5), satisfying the normality assumption of linear regression. Regression coefficients were tested against a significance level of 0.05.

Table 1

Summary of Variables

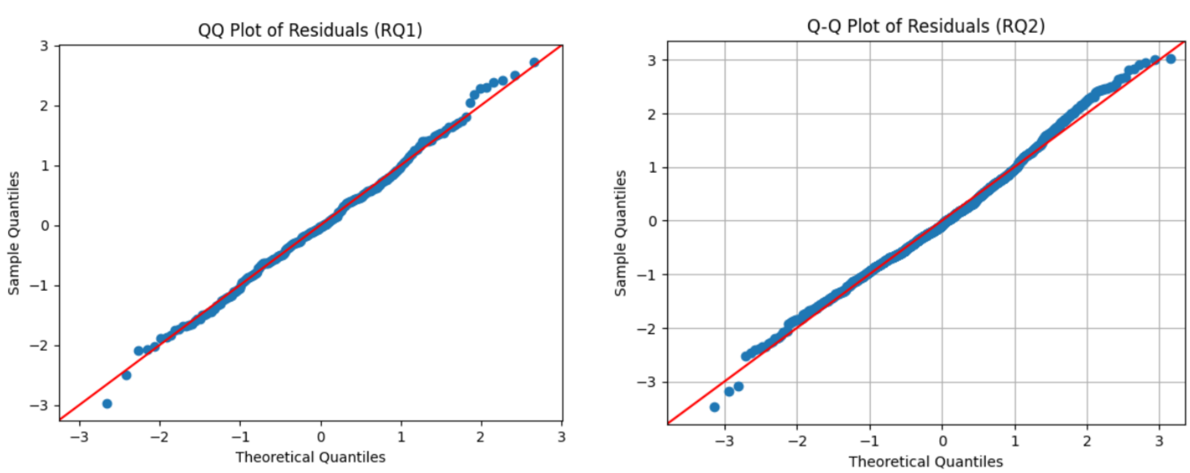
Research question	Variable	Type	Scale	Description
1	level 1[a]	Categorical	0, 1	Players played only Level 1 before Level 3
1	level 2[a]	Categorical	0, 1	Players played only Level 2 before Level 3
1	levels 1-2[a]	Categorical	0, 1	Players played Level 1, then Level 2 before Level 3
1	levels 2-1[a]	Categorical	0, 1	Players played Level 2, then Level 1 before Level 3
1 & 2	First_Success_level3_time_spent[b]	Continuous	4237-414197[e] 3153-215198[f]	Time taken to achieve first successful trial in Level 3 quantitative coordinate task
1	time_spent_level1[b]	Continuous	6756-1019500[f]	Time spent on Level 1 trials before first successful trials in Level 3
1	time_spent_level2[b]	Continuous	8959-1228887[f]	Time spent on Level 2 before first successful trials in Level 3
1	failed_time_spent_level3[b]	Continuous	0-458429[f]	Time spent on failed trials before first successful trials in Level 3

2	sharp	Continuous	0-5	Number of sharp turn regions player encountered
2	decision	Continuous	0-36	Number of branch regions player encountered
2	level2_track5[c]	Categorical	0, 1	Randomly assigned track players encountered when playing Level 2
2	level3_track7[d]	Categorical	0, 1	Randomly assigned track players encountered when playing Level 3

Note. [a]0 = Has neither Level 1 nor Level 2 trials before Level 3. [b]Measured in milliseconds. [c]0 = Assigned track 4 for Level 2, 1=Assigned track 5 for Level 2. [d]0 = Assigned track 6 for Level 3, 1=Assigned track 7 for Level 3. [e]Research Question 1. [f]Research Question 2.

Figure 5

Q–Q plots of residuals for Models 1 and 2



RQ3

To answer RQ3, we conducted multimodal qualitative analysis of audio–video data of speech, movement, and eye gaze. We looked for instance of dyads navigating challenging features and selected one representative case dyad whose interaction exemplified the typical coordination patterns observed across dyads, namely, negotiating reference to spatial relations and coordinating actions during moments of increased task difficulty.

Results

Quantitative Results

For RQ1, Table 2 presents completion times across sequencing conditions. For RQ2, all dyads followed the same sequence, a fixed sequence of Level 1, Level 2, and Level 3, and successfully completed Levels 1 and 2, resulting in a single condition. The mean total completion time was 243.79 seconds (SD = 157.86; N = 1,218).

Table 2

Completion time (s) under different sequencing conditions in RQ1

Sequential conditions	N	Mean (seconds)	SD (seconds)
Only_level_1	31	132.4	83.1
Only_level_2	36	184.8	76.8
Level_1_then_2	48	217.3	96.2
Level_2_then_1	26	245.0	139.5
Only_level_3	113	108.2	96.4
Total	254	156.6	110.1

Model 1

As shown in Table 3, players who experienced the Level 1–Level 3 sequence of levels (sensorimotor coordination in Level 1, followed by the quantitative coordinate task in Level 3) completed the Cartesian coordinate level in about 60% *(e^{-0.507} ≈ 0.602)* of the time of those who did not play Levels 1 and 2, indicating approximately 40% faster performance. This result was statistically significant, controlling for the track they received on Level 3. No other level sequences showed a statistically significant impact on performance.

Table 3

Multivariate Linear Regression for RQ1

Variable	Regression Coefficient	Standard Error	*p*	95% CI
Constant	-1.159	0.871	0.183	-2.865 ~ 0.548
Level 1	-0.507	0.141	0.000**	-0.784 ~ -0.230
Level 2	-0.151	0.131	0.249	-0.407 ~ 0.106
Levels 1-2	-0.129	0.103	0.212	-0.331 ~ 0.073
Levels 2-1	-0.043	0.136	0.753	-0.309 ~ 0.223
Level3_track_7	1.137	0.082	0.000**	0.977 ~ 1.298

Note. Total $N = 254$ CI = confidence interval.

Model 2

As shown in Table 4, for each decision point encountered by a dyad, their estimated mean time spent on the first successful trial of Level 3 was approximately 1.9% shorter *(e^{-0.019} ≈ 0.981)*, controlling for Level 2 and 3 track assignment and time spent in each level.

Table 4

Multivariate Linear Regression for RQ2

Variable	Regression Coefficient	Standard Error	*p*	95% CI
Constant	9.855	0.040	0.000**	9.778 ~ 9.933
Level2_track_5	-0.066	0.034	0.048*	-0.132 ~ -0.001
Level3_track_7	1.123	0.033	0.000**	1.059 ~ 1.188
sharp	0.07	0.036	0.052	-0.001 ~ 0.141
decision	-0.019	0.006	0.002**	-0.031 ~ -0.007
time_spent_Level1	0.000	0.000	0.004**	0.000 ~ 0.000
time_spent_Level2	0.000	0.000	0.002**	0.000 ~ 0.000
failed_time_spent_level3	0.000	0.000	0.088	0.000 ~ 0.000

Note. Total $N = 1218$. CI = confidence interval.

Qualitative Results

To investigate the moment-to-moment emergence of coordinate perception, we analyzed a representative case study of a player dyad. While this specific dyad consisted of a mother and child—a composition necessitated by the museum's caregiver protocols for recorded sessions—their interaction illustrates the universal coordination mechanisms triggered by the game's design.

Figure 6a shows the first track encountered by the dyad, positioned respectively as the *x*- and *y*-players. After two seconds of *x*-dominant adult action, the child joins in, taking control of the *y*-axis slider, and soon utters an elongated "ooh!"—a vocalization we interpret as marking her eager realization of her role in the game mechanics.

Figure 6

Adult–Child Collaborative Joint Action—Playing OЯTHO Level 1

a. *b.*

Note. a. An OЯTHO Level 1 track with highlighted locations bearing structural features of interest from Model 2. b. Adult and child discussing their prospective joint actions as they encounter this particular location.

At first, their motion is slow and somewhat erratic, as though testing how their individual proximal inputs affect the distal behaviors of the shared object. The ball drifts unevenly through the track, apparently reflecting their ongoing attempts to co-establish a constant pace. Gradually, their movements become more evenly paced, and their path runs evenly between the track boundaries.

The ball presently reaches a bifurcation in the track that requires a joint decision (see Figure 6a). The players pause. The child, gazing toward the narrow connector branch asks, “And here?” The mother asks back, “Where do we have to go?” Both players then shift their gaze toward the track’s end point and point toward it. The child answers, “To that circle.” The mother proposes taking the shorter route through the narrow connecting branch (see Figure 6b), the child consents, and the pair resumes their coordinated manipulation of the ball along that branch.

A few seconds later, as they turn into the narrow branch, the child says, “You have to go a bit lower.” This instruction drew on the child’s egocentric frame of reference (see Figure 7a). For the mother—standing perpendicularly—“lower” meant moving left on the screen (see Figure 7b). Nevertheless, the mother replied “Okay” and adjusted her movement. The child then added, “and turn,” a phrase that again highlights how communicating intentions occasions opportunities for establishing neutral spatial terminology.

Figure 7

The Ambiguous Meaning of "Lower" Through Two Players' Orthogonal Points-of-View

a. b.

Note. a. The child's point of view as she refers to "lower." b. The mother's contemporary view. Red circles, indicating foveal vision, obscure the white ball.

At the next branching off of a narrow passage (see Figure 6a, near the end point), the pair did not verbalize their decision, as though further discussion was no longer necessary. Instead, they applied the same strategy as before, silently choosing the shorter route.

Discussion

RQ1 aimed to evaluate the game's pedagogical design rationale, namely whether first engaging in sensorimotor entrainment—that is, figuring out how to maneuver the ball together, each player using their axial motor control—later contributes to adopting formal mathematical nomenclature, that is, using pairs of axial values signifying the location of points on the Cartesian plane. Supporting our design conjecture, our quantitative investigation shows that players who first developed motor-control skills later performed better with the symbol-control game mechanics than those who began directly with the symbol-control game mechanics. This finding supports claims by learning scientists who, drawing on radical-constructivist and enactivist theories of cognition, view mathematics learning as developing new perceptual structures guiding action (Abrahamson, 2014; Steffe & Kieren, 1994). Yet what exactly is it about sensorimotor entrainment that gives rise to mathematical fluency?

RQ2 probed figural properties of the gaming tasks, specifically the track's structural elements, in an attempt to implicate dyadic experiences pivotal for developing coordinate perception. Quantitative analyses revealed that encountering "decision points," track bifurcations that constrain continuous co-control (highly correlated with challenging narrow track segments) is correlated with greater success, later, on the symbol-based game mechanics. Still, what takes place among the players at these decision points that gives rise to each of them better transitioning from sensorimotor entrainment to mathematical discourse?

Turning to RQ3, we next sought to demonstrate how our quantitative findings play out in moment-to-moment dyadic interactions. Specifically, we looked to understand *how* encountering the would-be challenging track features contributes to the development of coordinate perception. The focal episode showed how arrival at "decision points" instigates dyadic discussion, where players are spurred to communicate their intentionality, present dilemmas, weigh tradeoffs, and deliberate over prospective joint action. These discussions recruit multimodal semiotic resources of speech, gaze, gesture, and action to forge new speech referents, such as establishing frames of reference supporting coordinated action (e.g., debating the relational specifier "lower"). It would appear that rendering immersive action explicit, by articulating implicit qualities of perceptual orientation, could explain why our quantitative analysis found that encountering decision points in Level 1 predicts greater success rates on Level 3. Namely, negotiating joint action bootstraps the development of coordinate perception from sensorimotor entrainment to mathematical reasoning. These cognitively auspicious effects of “languaging” perception have been previously witnessed in empirical evaluations of mathematics learning activities (e.g., Steffe & Kieren, 1994). More generally, perturbing action may catalyze cognitive development (Koschmann et al., 1998).

Limitations

Whereas our quantitative analysis focuses on child–child dyads, the qualitative case study features a mother–child pair, due to ethics-board protocols. We emphasize that these qualitative data are illustrative; they serve to demonstrate coordination mechanisms inherent to the game’s design, such as modulating pace or negotiating routes at bifurcations.

Conclusions

Given auspicious conditions, children can develop new mathematical notions through playing interactive games without prescriptive instruction or explicit guidance (Kamii & DeClark, 1985). So doing, children learn best when they engage first in non-symbolic activities and only later in manipulating numerical signifiers (Nathan, 2012). This is true for collaborative learning, too, where moments of collective breakdown, negotiation, and resolution promote conceptual understanding (Koschmann et al., 1998), often through “languaging” action into new ontologies (Maturana & Varela, p. 211; Steffe & Kieren, 1994, p. 723) to enhance coordinative inter-regulation. As such, this study has both corroborated the action-based genre of the embodied design framework (Abrahamson, 2014; Alberto et al., 2022) and extended it to the configuration of dyadic collaboration. Namely, conditions can be set so that two children at play spontaneously develop new perceptual orientations relevant to targeted mathematical practices, such as coordinate perception, as their expedient means of solving emergent motor-control problems.

Thus, new cognitive structures emerge from recurrent sensorimotor patterns that enable action to be perceptually guided (Varela et al., 1991).

The OЯTHO project arguably creates auspicious conditions for large-N research investigating neuro-cognitive processes of individual learning in joint action (de Gelder, 2023).

Finally, as interactive digital technologies become generally accessible, we envisage schools adopting discovery-based motor-control activities such as OЯTHO. Future research could examine how best to integrate these activities into existing curricular modules, in particular evaluating the activities' potential pedagogical function as cognitive gateways to conceptual units across digital and non-digital media.

References

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

Abrahamson, D., & Wilensky, U. (2007). Learning axes and bridging tools in a technology-based design for statistics. *International Journal of Computers for Mathematical Learning, 12*(1), 23–55. https://doi.org/10.1007/s10758-007-9110-6

Alberto, R., Shvarts, A., Drijvers, P., & Bakker, A. (2022). Action-based embodied design for mathematics learning: A decade of variations on a theme. *International Journal of Child-Computer Interaction, 32*, 100419. https://doi.org/10.1016/j.ijcci.2021.100419

de Gelder, B. (2023). Social affordances, mirror neurons, and how to understand the social brain. *Trends in Cognitive Sciences, 27*(3), 218–219. https://doi.org/10.1016/j.tics.2022.11.011

Goldstone, R. L., Landy, D. H., & Son, J. Y. (2010). The education of perception. *Topics in Cognitive Science, 2*(2), 265–284. https://doi.org/10.1111/j.1756-8765.2009.01055.x

Hutto, D. D., & Abrahamson, D. (2022). Embodied, enactive education: Conservative versus radical approaches. In S. Macrine & J. Fugate (Eds.), *Movement matters: How embodied cognition informs teaching and learning* (pp. 39–52). MIT Press.

Kamii, C. K., & DeClark, G. (1985). *Young children reinvent arithmetic: Implications of Piaget's theory*. Teachers College Press.

Kita, S., Alibali, M. W., & Chu, M. (2017). How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological Review, 124*(3), 245–266. https://doi.org/10.1037/rev0000059

Koschmann, T., Kuuti, K., & Hickman, L. (1998). The concept of breakdown in Heidegger, Leont'ev, and Dewey and its implications for education. *Mind, Culture, and Activity, 5*(1), 25–41. https://doi.org/10.1207/s15327884mca0501_3

Maturana, H. R., & Varela, F. J. (1992). *The tree of knowledge: The biological roots of human understanding*. Shambala Publications. (Originally published in 1987).

Mechsner, F., Kerzel, D., Knoblich, G., & Prinz, W. (2001). Perceptual basis of bimanual coordination. *Nature, 41*(6859), 69–73. https://doi.org/10.1038/35102060

Nathan, M. J. (2012). Rethinking formalisms in formal education. *Educational Psychologist, 47*(2), 125–148. https://doi.org/10.1080/00461520.2012.667063

Reed, E. S., & Bril, B. (1996). The primacy of action in development. In M. L. Latash & M. T. Turvey (Eds.), *Dexterity and its development* (pp. 431–451). Lawrence Erlbaum Associates.

Schmitz, L., Vesper, C., Sebanz, N., & Knoblich, G. (2017). Co-representation of others' task constraints in joint action. *Journal of Experimental Psychology: Human Perception and Performance, 43*(8), 1480–1493. https://doi.org/10.1037/xhp0000403

Sebanz, N., & Knoblich, G. (2021). Progress in joint-action research. *Current Directions in Psychological Science, 30*(2), 138-143. https://doi.org/10.1177/0963721420984425

Shvarts, A. (2018). Joint attention in resolving the ambiguity of different presentations: A dual eye-tracking study of the teaching-learning process. In N. Presmeg, L. Radford, W.-M. Roth, & G. Kadunz (Eds.), *Signs of signification: Semiotics in mathematics education research* (pp. 73–102). Springer.

Shvarts, A., & Abrahamson, D. (2023). Coordination dynamics of semiotic mediation: A functional dynamic systems perspective on mathematics teaching/learning. *Constructivist Foundations, 18*(2), 220–234. https://constructivist.info/18/2/220.shvarts

Steffe, L. P., & Kieren, T. (1994). Radical constructivism and mathematics education. *Journal for Research in Mathematics Education, 25*(6), 711–733. https://doi.org/10.5951/jresematheduc.25.6.0711

Török, G., Pomiechowska, B., Csibra, G., & Sebanz, N. (2019). Rationality in joint action: Maximizing coefficiency in coordination. *Psychological Science, 30*(6), 930–941. https://doi.org/10.1177/0956797619842550

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. MIT Press.

Wei, H., Rogier, B., & and Drijvers, P. (2025). A digital-embodied design for functional thinking in the classroom. *Mathematical Thinking and Learning*, 1–31. https://doi.org/10.1080/10986065.2025.2493375