MAKING AR AND VR SHAREABLE: MIRRORING AS A COORDINATION ARTEFACT

Beatrice Bruno Franco[1], Marzia Garzetti[1,2], Ornella Robutti[1]

[1]Università di Torino,

[2] University of Applied Sciences and Arts of Southern Switzerland (SUPSI)

AR/VR research in mathematics education has highlighted new possibilities for embodied interaction with mathematical objects. Less is known about how collaboration is organised in asymmetric classroom settings where only one student has immersive access to AR/VR and peers rely on a mirrored display. This qualitative case study examines mirroring, the real-time projection of the headset user's view onto a shared screen, in a Grade 8 task on translations in the prototype Transforming Math (TM). Data were analysed through multimodal semiotic analysis grounded in the semiotic bundle to trace joint attention and joint action. Findings shows that mirroring stabilises shared referents, enables request, execution – monitoring - coordination, supports conjecture checking, and fosters shared semiotic resources.

INTRODUCTION

Research on virtual and augmented reality (AR/VR) in mathematics education has highlighted new possibilities for embodied interaction with mathematical objects and for coordinating attention around shared digital artefacts (Bulut & Borromeo Ferri, 2023; Walkington et al., 2024). However, a key gap concerns how collaboration is organised when access to the immersive environment is asymmetric, that is, when some student can directly perceive and manipulate the virtual scene through the headset, while peers must coordinate the activity from outside the AR/VR environment. In such settings, the interactional infrastructure that enables participants to see together and act together becomes critical yet remains underexplored. This paper addresses this gap by focusing on a hybrid configuration in which one student wears the headset and manipulates virtual objects, while peers observe the activity through *mirroring*, which projects the headset user's first-person visual field onto a shared screen in real time. Rather than simply duplicating the virtual scene, mirroring produces a shared *mirrored display* that makes the headset user's actions, selections, and on-screen feedback publicly available to the group, thereby creating a common hybrid environment for reference, coordination, and negotiation. Although prior work has largely examined how individual learners interact with virtual objects, much less is known about how collaborative work is organised around mirrored displays, and in particular how mirroring shapes students' joint attention and joint actions around emerging mathematical objects (Shvarts & Abrahamson, 2019; Böckler & Sebanz, 2013). We investigate these issues in a Grade 8 activity using the AR/VR prototype Transforming Math (TM), designed to support exploration of geometric

2026. In H. Krzywacki, M. Hähkiöniemi, E. Haataja, M.S. Hannula & P. Hästö (Eds.). *Proceedings of the 49th Conference of the International Group for the Psychology of Mathematics Education: Research Reports*, pp. 227-234. PME.

transformations in three dimensions. The purpose of the study is to examine whether and how mirroring contributes to the collective construction of mathematical meaning during collaborative work. To analyse interaction processes in this hybrid environment, we adopt a multimodal perspective grounded in the notion of semiotic bundle (Arzarello et al., 2009), which allows us to trace how speech, gaze, gestures, and the digital signs made available through mirroring are dynamically coordinated as students progressively construct and negotiate mathematical objects and relations.

THEORETICAL FRAMEWORK

Recent research has investigated learning as a collaborative, embodied process unfolding in designed environments, where digital artefacts structure the shared space for action and perception (Shvarts & Abrahamson, 2019). Within such environments, where attention to content is conceived as perceptual, embodied and distributed through participants, Shvarts & Abrahamson (2019) draw on the construct of joint attention and joint action to analyse collaborative processes of learning. *Joint attention* is understood as the process through which multiple individuals orient toward and interpret the same visual display, including emerging mathematical objects (Shvarts & Abrahamson, 2019), while *joint action* refers to the coordination of participants' actions within a shared activity space (Böckler & Sebanz, 2013).

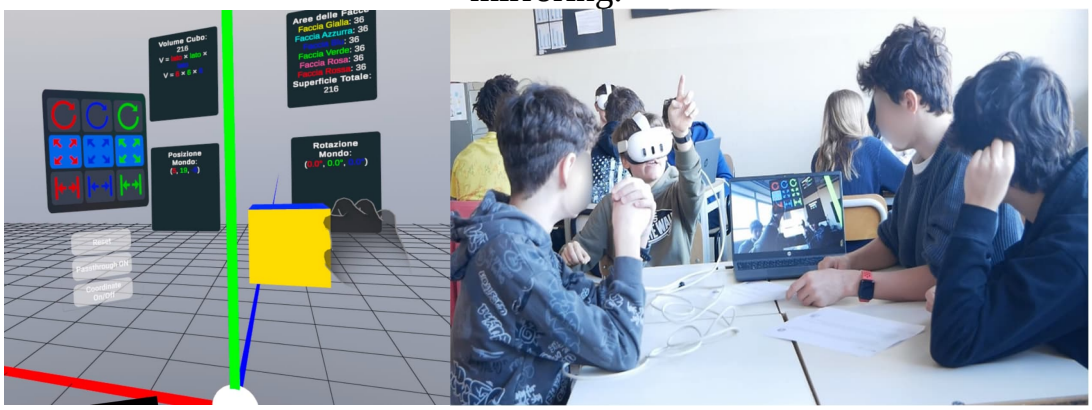
Building on these perspectives, our study focuses on a designed digital learning environment constructed through AR and VR. The environment is accessed through a headset and coupled with mirroring, which projects the headset user's visual field onto a shared mirrored display, thus structuring a shared hybrid environment. To examine whether and how joint attention and joint action emerge in this hybrid environment, and how they are supported by mirroring, we adopt a semiotic analysis grounded in the notion of semiotic bundle (Arzarello et al., 2009). Within the semiotic bundle, collective construction of mathematical meaning is conceived as a semiotic interweaving that involves a plurality of signs and resources, such as speech, gestures, gaze, body movements and orientations, and the digital signs and overlays provided by interactions with artefacts. These signs and resources are examined in terms of their dynamic coordination over time, through which mathematical objects and relations are progressively constructed and negotiated. The semiotic bundle construct makes it possible to study how different semiotic systems are coordinated and transformed during activity (Arzarello et al., 2009). Digital artefacts are thus not treated as neutral tools, but as active participants in the genesis and evolution of mathematical signs (Arzarello et al., 2015). In this context, *mirroring* is conceptualised as a coordination artefact that shapes the semiotic bundle that is crucial for understanding how students orient and share their attention toward relevant objects and actions, and thus for analysing coordination between participants.

In what follows, we use the semiotic bundle as an analytical lens to investigate the following research question: *what role does mirroring play in TM in supporting students' joint attention and joint action in the collective construction of mathematical meaning?*

METHODS

The analysed example is drawn from an activity carried out using the prototype Transforming Math (TM), developed by Mosé Colangelo within the research team at the University of Turin, coordinated by Ornella Robutti, as part of the HORIZON WIDERA TransEET project (coordinated by Chronis Kynigos). TM is an AR/VR environment designed to support the exploration of geometric transformations (translation, rotation, and dilation) in 3D space through a Cartesian coordinate system, with axes colour-coded in red (x-axis), green (y-axis), and blue (z-axis), and a manipulable cube (Figure 1, left). Transformations are performed via a control panel (Figure 1, left) and visualised in real time, alongside changes in numerical properties (coordinates, edge lengths, face areas, volume and rotation angles) (Figure 1).

Figure 1: On the left TM 's VR interface. On the right: spatial arrangement of students using the TM and the shared view of the TM environment through mirroring.



The analysed activity involves four eighth-grade students (A, L, P and R) engaged in a collaborative mathematics task specifically designed for TM. The tasks guided students' exploration and attention toward identifying the key characteristics of these transformations, with a particular focus on their invariants. Students worked in the hybrid environment: one student at a time interacted with the AR/VR environment through the headset, while the other students of the group collaborated by observing and discussing the activity via the mirrored display (Figure 1, right).

Collected data consist of external video recording, screen capture of the mirrored display, audio recording, and students' worksheets. The present analysis focuses on a 20-minute segment in which students work on translation.

The analysis is anchored in the constructs of *joint attention* and *joint action* (Shvarts & Abrahamson, 2019; Böckler & Sebanz, 2013), which we examine through Arzarello's notion (2009) of *semiotic bundle* to trace the coordination of semiotic signs and resources in interaction. Within the semiotic bundle, particular attention is paid to gestures, verbal interaction and gaze involving the mirrored display. The analysis was indeed conducted to identify recurrent roles of mirroring within the task by drawing on a semiotic analysis of coordination that operationalised and traced joint attention and joint action. *Joint attention* is identified through a semiotic analysis that reveals the coordination of gazes, gestures and interactions conveyed by mirroring. Coordination can be detected, for example, starting from a gesturing hand, or shared discursive treatment of the same focal content, including uptake, reformulation, and contestation anchored to what is visible on the shared display. *Joint action* is operationalised through recurring coordination actions between the headset user (HU) and the non-headset user (NHU) students: NHU students request specific actions, which the HU executes in TM, while the group monitors outcomes on the mirrored display and guides subsequent moves through prompts, confirmations, or corrections. In the following section, we analyse the *different roles of mirroring* identified through the semiotic analysis, with particular focus on how they support joint attention and joint action.

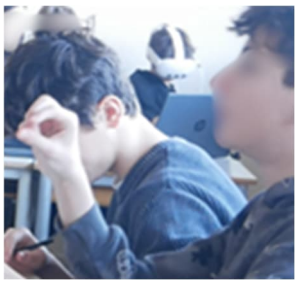
RESULTS

The results are presented through a set of illustrative episodes selected from the 20-minute segment because they instantiate the semiotic coordination practices through which joint attention and joint action are achieved (i.e., systematic co-orientation of gaze and gesture toward the mirrored display; request, execution, monitoring sequences linking talk, action and on-screen feedback). Rather than offering an exhaustive account of the whole segment, we focus on moments where mirroring becomes consequential for the collective construction of mathematical meaning, since these episodes allow us to make explicit how the mirrored display enters the semiotic bundle and supports coordination across participants. In what follows, we organise these episodes around four recurrent roles of mirroring identified in the analysis: (i) in *coordinating actions* between the HU and NHU students; (ii) in the *sharing of gestures and movements* within the group, understood as resources for attributing meaning to mathematical objects; (iii) in the *verification of hypotheses*, often in the form of requests to perform actions addressed by NHU to the HU; and (iv) in the *redistribution of epistemic roles* in both directions, both when the HU enables the group to explore the digital environment through mirroring, and when the NHU students orient the HU's selections and actions. These roles do not all refer to the same analytical level: some concern actions (i-ii), while others concern the epistemic purposes for which such actions are mobilized (iii-iv). In the following four illustrative episodes are reported.

To illustrate (i), we consider the episode in which R wears the headset for the first time and attempts to move the cube without having selected the translation buttons (Figure

1). Here, the mirrored display, as a coordination artefact, makes the selection state and the headset user's attempts visible to all, enabling peers to issue timely prompts and corrections grounded in what is jointly observable. The missing selection, visible through the mirrored display, is immediately noticed by peers monitoring the screen: P says, "No, you have to click! R you didn't click," while R replies, "wait a second, I'll reset," attempting to select the button but failing to do so. The interaction thus takes the form of fine-grained regulation of R's action based on what peers can see on the mirrored display. Verbal prompts such as "NO, you need to be more confident", "More decisive", "No, you need to be a little higher… a little lower…", accompanied by gestural indications of the height at which to hold the hand in order to successfully select the button. Here, joint attention to the mirrored display encourages joint action, a sequence of requests - executions - monitoring: peers' instructions prompt R's attempts to correctly use TM, which are adjusted in real time based on the mirrored feedback.

Figure 2: simulated gesture



To illustrate (ii), we focus on the emergence of *simulated gestures*, understood as gestures that reproduce the movements required to perform actions in TM (as observed on the mirrored display), produced by NHU students (Figure 2). The task requires activating only the green and red directions (respectively y and x), and the mirrored display shows that the cube can move only on one plane: the red/green (x, y). P, R, and A watch on the screen what L – with the headset – is doing, while L verbalises the allowed movements: "I can move it up or down and then to the right and left." On the shared mirrored display, R states, "It moves in all directions," and P reformulates, "*All directions* means in 3D"; A challenges this with "NO it's in 2D," anchoring their stance to what everyone can see on the screen and cueing toward it with their face. At this juncture, the first simulated gesture of the entire segment appears: P has not yet used the headset and has been gazing at the mirrored display and L's actions. He simulates 'grasping' the cube by pinching his index finger and thumb together, using this action as an argumentative resource, saying, "eh no! You can't move it also like this" (Figure 2). The gesture proceeds along the imagined blue axis (z-axis), which at that moment is not activated, thereby making explicit the difference between translation in 2D and in 3D. Across the analysed segment, this simulated grasp gesture (index finger and thumb in contact) is then taken up systematically by all group members to recall,

anticipate, or check movements and constraints of the TM environment, occurring overall approximately 20 times in 21 minutes. Simulated gesture complements the mirrored display by re-enacting action possibilities; its repeated uptake indicates that the group shares and stabilises it as a common semiotic resource for coordinating interpretation ("2D/3D") and joint action around the same display.

To illustrate (iii), we consider the episode in which the group works in a checking phase rather than an exploratory one: students first construct a conjecture about the signs of the axis coordinates (positive/negative) without acting in TM, and only afterwards address operational requests to the HU to test the conjecture (by moving the cube along selected axes) to the HU to empirically test and validate their idea. R, the HU, moves the cube to the right and left (red x-axis) and observes how the coordinates change in sign, while the peers follow the changes in the values displayed on the screen. L proposes "If you move it up…, it's positive", accompanying the utterance with a simulated grasp gesture (thumb and index finger very close) that supports the formulation of the conjecture by blending with the directional indication along the vertical axis; L then proposes the opposite case: "if you move it down, it's negative", with an analogous downward gesture and gaze following the hand. The verification dynamic becomes explicit when the hypothesis concerns the blue z-axis and the away/towards movement: L begins, "If you move it away it's positi…," A counters, "Yes, if you move it away it should be negative," and L replies, "No, at most it's the other way around! If you move it close it's negative." At this point, the mirrored display, as the coordination artefact, provides the common evidential field (selection states and coordinate changes) through which the group can jointly evaluate competing conjectures. To decide, L proposes to R, "Try to do the same thing with the blue"; R selects the blue button and deselects the red, and the group observes the coordinate variation in relation to translation actions along the blue axis (z-axis). This episode highlights a sequence of requests - executions - monitoring in which competing hypotheses about positivity and negativity are resolved by aligning speech, simulated gestures and the mirrored display of selection states and coordinate changes.

While (iii) relates to epistemic role redistribution, we treat it separately because TM is used explicitly as a tool for empirical control. Here we focus on (iv) the redistribution of epistemic roles in two directions. On the one hand, NHU may take on a "director" role, orienting selections and actions: for instance, L guides R's work through specific prompts ("Select the second row," "the red one," "and try to move it"), supported by observing the mirrored display and by instructional gestures. On the other hand, in the initial exploratory phase, the HU makes available to the group the affordances and constraints of the digital environment: after P asks to activate two directions ("so… select the red and the green […] What can you move?"), L responds by combining talk and action ("I can move it up or down and then to the right and left") while carrying out the movements, and P, R, A watch on the screen what L does. These role shifts

show how mirroring redistributes epistemic work: directing relies on visible on-screen states to specify and evaluate joint actions, while demonstration turns individual exploration into a shared resource through peers' co-orientation (joint attention) and uptake on the mirrored display.

DISCUSSION AND CONCLUSIONS

This study asked what role mirroring plays in TM in supporting students' joint attention and joint action in the collective construction of mathematical meaning. Our analysis suggests that mirroring contributes to both processes by making the headset user's perceptual field publicly available and therefore usable as a shared space for reference, coordination, and negotiation. In the episodes discussed, mirroring recurrently supports joint attention by stabilising common focal points (e.g., axes, selection states, coordinate values) toward which gaze, talk, and gesture can converge (Shvarts & Abrahamson, 2019), and supports joint action by enabling request, execution, monitoring sequences through which NHU students guide, refine, and evaluate the HU's operations and vice versa (Böckler & Sebanz, 2013). The semiotic bundle allows us to highlight how these coordination processes become visible as the dynamic interweaving of multimodal resources (Arzarello et al., 2009), structuring the collective construction of mathematical meaning. A particularly indicative result concerns the emergence and circulation of simulated gestures, first produced by students *not* wearing the headset. This points to mirroring as more than a technical duplication of the headset view: it can function as a coordination artefact that supports the collective construction of shared semiotic resources (Arzarello et al., 2015) and consequently mathematical meaning. By observing actions and constraints through the mirrored display, students generate embodied signs that complement the visual feed and become reusable means to anticipate, contest, and stabilise meanings (e.g., distinguishing 2D/3D movement). In this sense, mirroring appears to enrich the group's multimodal repertoire and to foster collective construction of mathematical meaning that is not reducible to the headset user's individual experience but distributed across participants through shared access to the activity space. The episodes also suggest, consistently with other research on technology role in cognitive activity (Arzarello et al., 2002), that mirroring may support different functions of TM across time: in some moments TM is appropriated mainly for exploration and meaning construction, while in others it is recruited to test conjectures already formulated through interaction (a shift that resonates with the distinction between ascending and descending control). Although we do not claim generality beyond the analysed segment, this observation highlights how mirroring can help the group re-purpose the same environment for exploration, coordination, and validation, depending on emerging goals.

These findings have implications for the design of AR/VR learning environments aimed at collaborative mathematical activity. When immersive interaction is coupled with mirroring, the environment can become more readily "shareable": non-immersed

peers can participate in directing actions, monitoring feedback, and developing shared semiotic resources grounded in what is visible on the display. This suggests that mirroring should be treated as a pedagogically relevant feature, designed to make key states and changes legible to observers (e.g., selections, invariants, numerical feedback), so that the hybrid environment can support collective construction of mathematical meaning rather than parallel individual experiences. Future work could examine how variations in mirroring design (e.g., resolution of interface cues, layout, latency) shape these coordination processes, and how task prompts and teacher moves may further leverage mirroring to sustain joint attention, joint action, and shared semiotic development.

References

Arzarello, F., Olivero, F., Paola, D., & Robutti, O. (2002). A cognitive analysis of dragging practices in Cabri environments. *ZDM*, *34*, 66–72. https://doi.org/10.1007/BF02655708

Arzarello, F., Paola, D., Robutti, O., & Sabena, C. (2009). Gestures as semiotic resources in the mathematics classroom. *Educational Studies in Mathematics*, *70*, 97–109. https://doi.org/10.1007/s10649-008-9163-z

Arzarello, F., Robutti, O., & Thomas, M. (2015). Growth point and gestures: Looking inside mathematical meanings. *Educational Studies in Mathematics*, *90*, 19–37. https://doi.org/10.1007/s10649-015-9611-5

Böckler, A., & Sebanz, N. (2013). Linking joint attention and joint action. In J. Metcalfe & H. Terrace (Eds.), *Agency and joint attention* (pp. 206–215). Oxford University Press.

Bulut, M., & Borromeo Ferri, R. (2023). A systematic literature review on augmented reality in mathematics education. *European Journal of Science and Mathematics Education*, *11*(3), 556–572. https://doi.org/10.30935/scimath/13124

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*. https://doi.org/10.1016/j.lcsi.2019.05.003

Walkington, C., Nathan, M. J., Huang, W., et al. (2024). Multimodal analysis of interaction data from embodied education technologies. *Educational Technology Research and Development*, *72*, 2565–2584. https://doi.org/10.1007/s11423-023-10254-9