Learning to gesture like locals think: Spontaneous spatial gestures adapt to local norms of spatial thinking

Shervin Nosrati (nosrati.shervin@gmail.com) & Tyler Marghetis (tyler.marghetis@gmail.com)
Cognitive & Information Sciences, University of California, Merced
5200 Lake Rd, Merced, CA 95343

Abstract

When thinking about space, people preferentially adopt a particular frame of reference, anchoring spatial relations either egocentrically (to their bodies) or allocentrically (to external landmarks). These preferences are often shared within a community, raising the question of how private norms for spatial thinking are made public and thus shareable. We tested whether spontaneous co-speech gesture might be one route by which these norms are expressed. Participants (*N* = 62) were exposed to an implicit norm to adopt a particular frame of reference when reasoning spatially. They then solved a spatial puzzle and described their solution. When participants conveyed spatial information in their gestures, they adopted either an egocentric or an allocentric frame of reference. Critically, these gestural frames of reference were congruent with the local norm. Gesture thus reflects socially learned conventions for adopting a spatial frame of reference. We discuss implications for the propagation and perpetuation of implicit cultural norms.

Keywords: spatial cognition, gesture, spatial frames of reference

Introduction

People vary in how they think about spatial relationships. This variation is not random. Within a culture, people tend to converge on shared ways of thinking spatially (Levinson, 1996, 2003; Majid et al, 2004). That is, communities have cultural norms for thinking about space that are implicitly expected and reinforced through social interaction. In some communities people habitually reason about small-scale space *egocentrically*, anchoring spatial relationships to their bodies (e.g. "The coffee is to the left of my computer"). In other communities like the Yupno in Papua New Guinea or speakers of Isthmus Zapotec in Mexico, they reason about spatial relationships *allocentrically*, anchoring them to something external to their bodies, such as cardinal direction or local landmarks (Cooperrider, et al., 2017; Pérez Báez, 2011). Although people can adopt either of these spatial frames of reference, within a community one will be favored over another — even though spatial frames of reference are seldom taught explicitly. How are these spatial norms learned?

One possibility is that people learn how to think about space in socially-normative ways through language, from the ways other community members talk about space. The idea is that established members of a community will use their community's preferred spatial frame of reference when describing spatial relations in speech or sign. New members of a community will habitually encounter such descriptions, which help them learn to think about space in normative ways. This is a form of linguistic relativity (Levinson, 2003; Levinson et al., 2002; Whorf, 1956). Evidence for this mechanism comes from cross-cultural correlations between spatial language and spatial thought (Levinson, 1996; Levinson, 2003; Majid et al, 2004). On this account, repeated adoption of a preferred spatial frame of reference in language shapes preferences for that same spatial frame of reference in thought.

But we know that this cannot be the whole story, since some populations have one norm in language but another in thought (Calderón et al., 2019). Yucatec women in Mexico, for instance, reliably adopt an egocentric frame of reference in language but an allocentric frame of reference in thought (Le Guen, 2011). Communities must have ways beyond language to converge on a shared spatial frame of reference in thought. If language isn't the only way these norms are propagated, by what complementary routes might communities convey their preferred frame of reference?

Here we investigate one possible route: co-speech gesture. According to *gestural relativity*, habitual exposure to gestures that adopt a particular frame of reference influences observers to adopt that frame of reference for their own spatial thinking (Le Guen, 2011; Levinson et al., 2002; Nosrati & Marghetis, 2025).

This account requires two complementary processes: first, speakers' gestures must reliably adopt a frame of reference that is consistent with local norms; and second, observers must reliably adopt this frame of reference in their subsequent thinking. There is recent evidence for this second process. Nosrati and Marghetis (2025) showed that observers will adopt in their own reasoning the frame of reference that a speaker used implicitly in their gestures.

As for the first process, there are prima facie reasons to believe that gesture might be better suited for conveying spatial relationships than speech. While speech must express spatial information sequentially in discrete symbolic units, gesture can represent spatial relationships holistically by enacting a layout. Direct evidence for this process, however, is surprisingly limited. Psychological studies have established that gestures convey rich spatial information to observers (McNeill, 1992). Anecdotally, storytellers have been reported to reliably adopt culturally-normative spatial frames of reference in their gestures (e.g., Haviland, 1993).

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

And Zapotec-Spanish bilinguals describing spatial relationships will consistently adopt the same frame of reference in their gestures, regardless of the language they are speaking (Marghetis et al., 2020).

People's gestures, therefore, adopt a frame of reference when conveying spatial relationships, and that frame of reference is sometimes correlated with local cultural norms for reasoning about space. We lack causal evidence, however, that local norms of thinking are actually driving the adoption of a particular frame of reference in gesture. Both local norms of thinking and individual gestural production could be shaped by some other confounding influence, such as the presence of salient visual landmarks (Li & Gleitman, 2002). It thus remains unclear whether the frames of reference that speakers adopt in gesture are flexibly conforming to community norms of spatial thinking. If so, this would establish gesture as a viable route by which local norms for thinking about space are made visible, transmitted, and preserved within a community.

To test this, we manipulated the local norm for thinking spatially (i.e., egocentrically or allocentrically). By 'local norm,' we mean a convention instantiated within a brief social interaction, rather than a stable, community-wide culture norm. We then tested whether people adopted this normative frame of reference in their gestures while describing spatial relations. We hypothesized that people would adjust their gestural frame of reference to conform to local norms of spatial thinking.

Methods

Participants

Students at an American research university participated in return for partial course credit (*N* = 62; gender: 47 female, 13 male, 2 non-binary; age: *M* = 20 years, range [18, 30]). Most reported English as their primary language (*n* = 50), followed by Spanish (*n* = 8). A third of participants reported proficiency in Spanish (34%), with a few reporting proficiency in Tagalog, Cantonese, and other languages. Both English and Spanish have been characterized as languages that preferentially use an egocentric frame of reference to describe space (Levinson, 2003; Calderón et al., 2019). Informed consent was obtained prior to the start of the study. One participant with prior knowledge of the study was excluded; there were no other exclusions.

Overall Design

In a between-groups experiment, we manipulated the implicit, local norm for spatial reasoning (egocentric vs. allocentric) on two rotation-based spatial reasoning tasks: a Cup Task and a Map Task (Fig. 1, top vs. bottom). For each task, we first established the local norm during a Training Phase during which participants observed and then replicated an experimenter's solution strategy, with corrective feedback (Fig. 1, left). This was followed by a Test Phase where participants solved the puzzle on their own and then described their solution, during which we recorded their spontaneous spatial gestures (Fig. 1, right).

Procedure for the Training and Test Phases

The experiment took place in an indoor room without exterior windows. Two tables were placed on opposite sides of the room (Fig. 1). Two experimenters were present: the 'Director' provided instructions; the 'Player' solved the task alongside participants.

During the initial **Training trials** for each task (*n* = 4), the Director demonstrated the solution on one table (i.e., moved a ball between cups, or pushed a cube out of a maze) while the Player and the participant watched (Fig. 1, left). All three then rotated 180°, and the Director asked them to solve it the "same way" on the other table (Fig. 1, right). Note that, after this rotation, egocentric and allocentric strategies generate different solutions: an egocentric strategy would maintain the spatial relationships from the observer's viewpoint, while an allocentric strategy would maintain the spatial relationships relative to external coordinates.

On the first Training trial, only the Player solved the puzzle. They used the assigned frame of reference and thus demonstrated the local norm (i.e., egocentric vs allocentric). The participant observed. On subsequent Training trials, both the participant and the Player solved the puzzle, and the participant received feedback on their solution. Specifically, if the participant's response did not conform with the local norm (i.e., egocentric vs allocentric), the Director told them they had solved the puzzle incorrectly and asked them to follow the Player's lead. The frame of reference was never referenced explicitly, only conveyed implicitly through the Player's behavior. Thus, the Player's spatial frame of reference was conveyed implicitly through their responses; the fact that this was *normative* was conveyed by the Director's corrective feedback.

During subsequent **Test trials** (*n* = 4), the roles changed. Now, the participant first solved the puzzle on their own while the Player was turned so they could not see the participant's solution. The participant and Player then turned to face each other, and the Director asked the participant to describe their solution to the Player. The Director then asked the Player to rotate to the second table to solve it "the same way." A barrier was placed between the tables so participants could not see the Player's solution.

Participants' descriptions were videorecorded and also coded online by the Director. On each Test trial, the Director noted whether the participant used their hands or head to convey the motion, and if so, the primary direction of these gestures. Participants were not encouraged to gesture.

After completing both tasks, participants answered a series of demographic and individual difference questions (e.g., gender, age, language use, urban/rural environments they grew up in, and ability to judge cardinal directions). For lack of space we don't analyze these variables here.

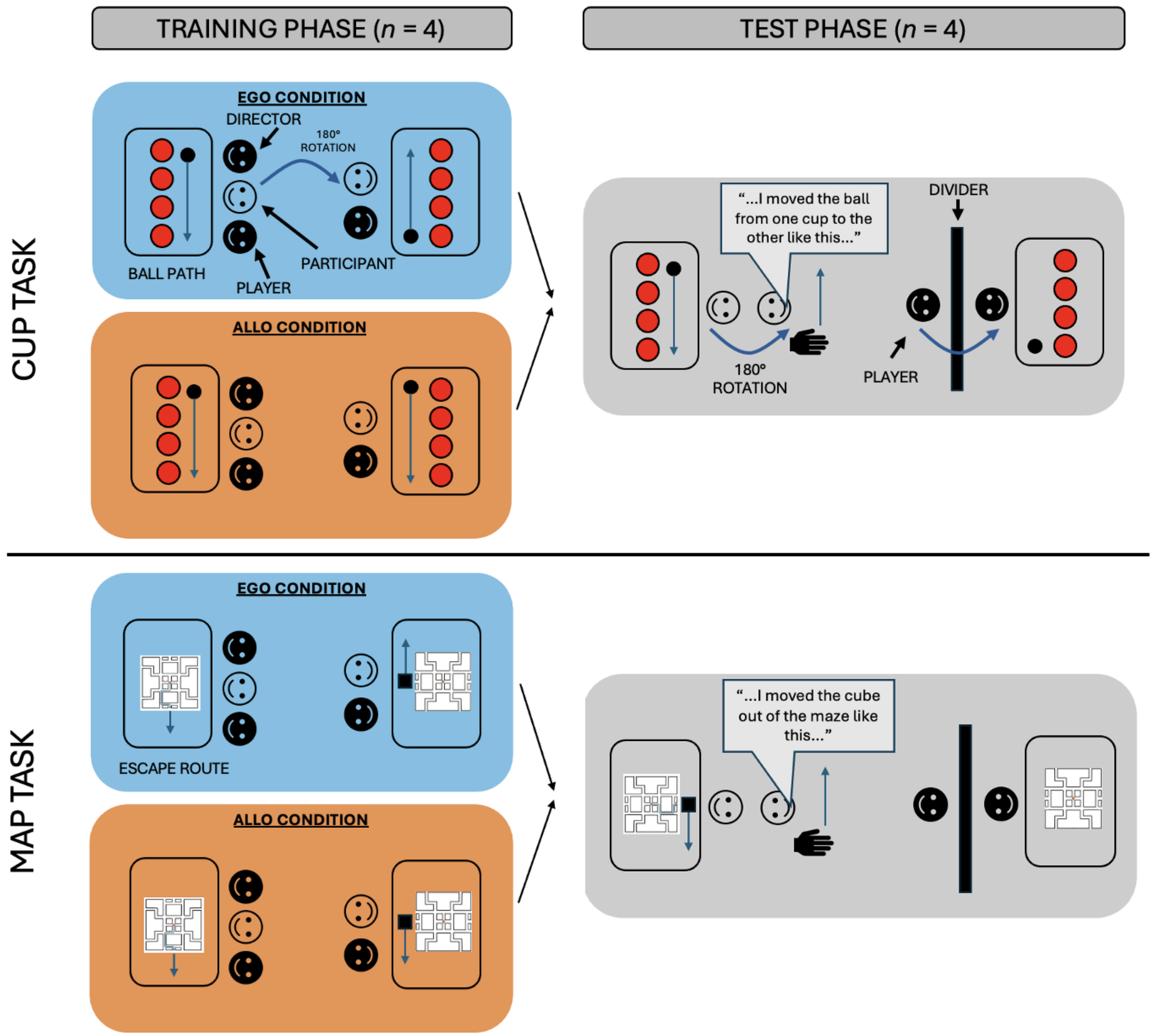

Fig. 1: Study procedure. Participants completed Training and Test phases for each task (Cup and Map). During Training (left), participants were exposed to a local norm for spatial reasoning (egocentric vs allocentric), manipulated between-subjects. The participant and a second experimenter (the "Player") watched the Director solve the puzzle on one table, then rotated 180° and solved it themselves "the same way" on a second table. The Player modeled the use of the assigned frame of reference, and the Director provided feedback to participants if their solution did not conform to the assigned frame of reference. During Test trials (right), participants first solved the spatial puzzle on their own, then rotated 180° to describe their solution to the Player. The Director recorded the Participants' gestures during their descriptions.

Procedure for Cup and Maze tasks

For the Cup task, both tables had four cups in a row and a small ball to move between cups. The task consisted of trying to reproduce on the second table the path and manner of the ball's motion on the first table. During the Training Phase, the Director used preset trajectories for the ball, with the primary direction — leftward or rightward — manipulated between trials (Fig. 2A). During the Test Phase, the participant was free to improvise the ball's trajectory from one of the outermost cups to the other.

For the Map task, the tables had identical mazes that were symmetric under 180° rotation The map on the first table had a prescribed route marked with a blue line. The task consisted of trying to reproduce on the second table the route used to escape the maze on the first table. The routes consisted primarily of motion along the lateral axis, with the primary direction — leftward or rightward — manipulated between trials (Fig. 2B).

Critically, both tasks can be solved in two ways: egocentrically or allocentrically. The egocentric solution involves anchoring spatial relations to the observer's own body. For instance, participants could reproduce a map escape route using the same sequence of left and right turns, from their perspective. The allocentric solution involves anchoring spatial relations to external coordinates. For instance, participants could reproduce a map escape route using movements anchored to the room's layout, moving away or toward a salient wall. While neither solution is "wrong," the Training Phase conveyed implicitly that one frame of reference was the normative solution.

Task order was fixed (Cup Task then Map Task). For each task, we used a fixed list of items and counterbalanced the item order across participants (Fig 2).

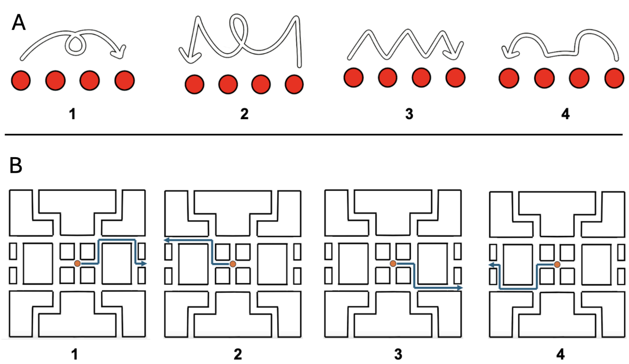

Fig. 2: **Items for the Cup and Map tasks.** (A) Paths used in the Training trails for the Cup task. Red circles represent plastic cups; white arrows indicate motion trajectories. (B) The escape routes in the Map task. On the first table, the red origin and blue escape route were printed on the map; on the second table, only the red origin was printed, so the escape route had to be inferred.

Analysis

Our primary dependent variable was the spatial frame of reference adopted in participants' gestures during their Test Phase descriptions. For each Test trial, we determined whether the participant's gesture conformed to an egocentric frame of reference, an allocentric frame of reference, or to neither frame of reference. For example, on the Cup task, if a participant moved the ball rightward, and then after rotating they gestured rightward, then their response was egocentric. By contrast, if a participant moved the ball from the cup closest to the wall to the cup farthest from the wall, and then after rotating their gesture conveyed a motion away from the wall, then their response was allocentric. See Figure 3 for examples of allocentric and egocentric gestures.

Results

Spontaneous adoption of spatial frames of reference in gesture

We first established that participants adopted a frame of reference to describe motion in their gestures. Across 496 trials (62 participants × 8 critical trials), participants almost always gestured while describing their solutions (86% of trials). On only a tiny minority of trials did participants gesture in a way that was inconsistent with both egocentric and allocentric frames of reference (2% of trials). Overall, therefore, participants reliably adopted either an egocentric or an allocentric frame of reference in their gestures (84% of all trials; Fig. 4), most commonly using an egocentric frame of reference (56% of trials). This Egocentric bias was most pronounced in the Maze task (mean proportion of trials: M = .66 ± .03 SEM).

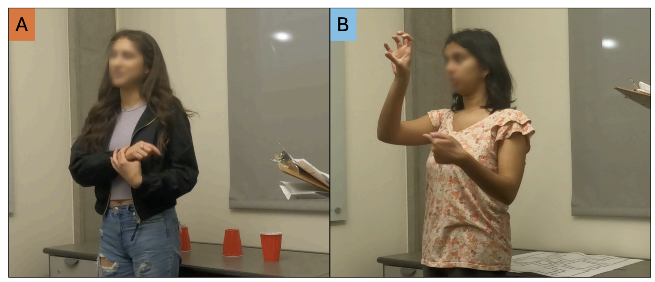

Fig. 3: Frames of reference in gesture. (A) The participant adopted an allocentric frame of reference while describing how she had moved a ball. Before rotating, while facing the table behind her, she had moved the ball from her left to her right. After rotating, she gestured from her right to her left, thus reproducing the motion from an external, allocentric perspective. (B) The participant adopted an egocentric frame of reference while describing how she had escaped a maze. Before rotating, while facing the table behind her, she had escaped by moving to her right. After rotating, she pointed to her right with one hand and reenacted the rightward motion with her other hand, thus reproducing the motion egocentrically, from her perspective.

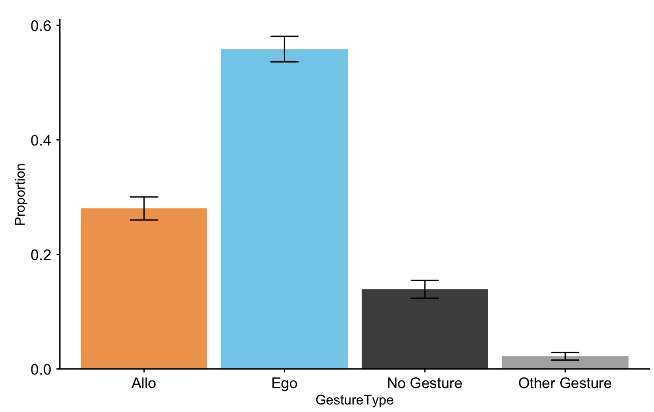

Fig. 4: Participants' gestures during motion descriptions. (allo = allocentric; ego = egocentric)

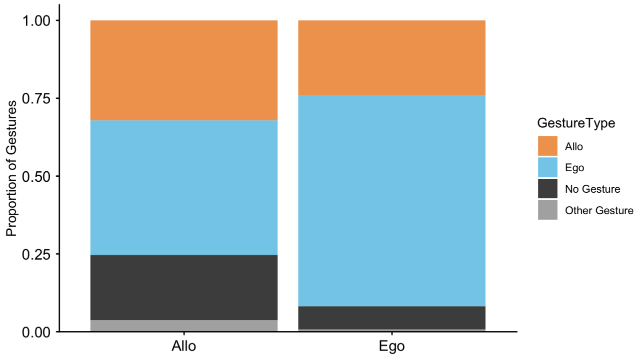

Fig. 5: Gesture production across conditions. Proportion of gesture responses among participants exposed to the allocentric norm (left) and the egocentric norm (right).

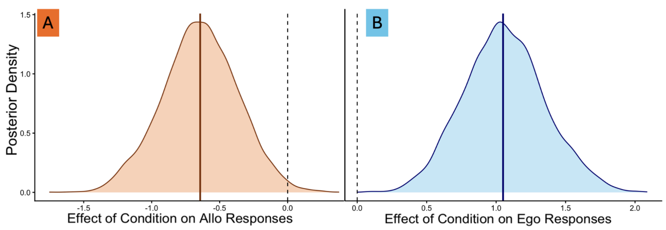

Fig. 6: Local norms shaped the frame of reference used in gesture. Posterior distributions of the effect of the egocentric (vs. allocentric) norm on the log-odds of gesturing (A) allocentrically and (B) egocentrically. (dashed vertical lines = zero effect; solid vertical lines = means of the posterior distributions)

Local norms shaped participants' spatial frame of reference in gesture

We next investigated if participants adapted their gestures to the local norm for spatial thinking (Fig. 5). As predicted, participants' gestures were shaped systematically by the local norms modeled during training. Participants in the Egocentric condition were much more likely to gesture about spatial relationships egocentrically (proportion of trials: *M* = .68 ± .03 SEM) than allocentrically (*M* = .24 ± .03 SEM). Participants in the Allocentric condition, by comparison, produced fewer egocentric gestures (*M* = .46 ± .03 SEM) and more allocentric gestures (*M* = .32 ± .03 SEM), although they still exhibited an overall egocentric bias. Rather than fully switching participants' frame of reference, therefore, the allocentric norm attenuated a robust egocentric default.

The effect of the local spatial norm on participants' spontaneous gesture was supported by a multilevel Bayesian logistic model (Fig. 6). Participants in the Egocentric condition were reliably more likely to produce egocentric gestures (b = 1.06, 95% Bayesian Credible Interval = [0.50, 1.66]) and reliably less likely to produce allocentric gestures (b = -0.64, 95% Bayesian Credible Interval = [-1.2, -0.11]), compared to those in the Allocentric condition. Note that the 95% Bayesian credible intervals do not include 0, indicating strong evidence that participants adjusted their gesture to conform to the local norms of spatial thinking.

The effect of implicit exposure to the spatial norm persisted for multiple trials (Fig. 7). During the Training phase, participants quickly inferred the implicit norm for adopting either an egocentric or allocentric frame of reference (Fig. 7, left of the dashed vertical lines). By the final training trial, most participants in the Egocentric condition solved the puzzle egocentrically, for both the Cup and Map task, while participants in the Allocentric condition were either split equally between frames of reference (Cup task) or preferred the egocentric solution (Map task). Once training was complete, participants received no more cues to use a specific frame of reference. Despite the lack of ongoing cues, participants persisted in gesturing in ways that reflected the implicit norm (Fig. 7, right of the dashed vertical lines). In the Cup task, participants in the Egocentric condition were more likely to gesture egocentrically than allocentrically, and vice versa for those in the Allocentric condition. In the Map task, there was an overall bias toward egocentric gesturing, but this was attenuated among participants in the Allocentric condition. Thus, brief exposure to an implicit spatial norm shaped participants' subsequent spatial gestures even multiple trials after they had stopped receiving any cues to the norm.

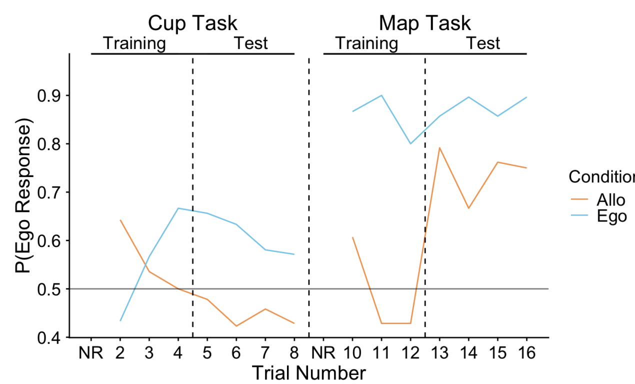

Fig. 7: Time series of egocentric responses for the Cup (trials 1 to 8, left) and the Map (trials 9 to 16, right) tasks. Vertical dashed lines indicate the transition from Training to Test. During Training, participants' responses consisted of their puzzle solutions; during the Test phase, their responses consisted of their gestural descriptions. (Solid horizontal line = random choice of a frame of reference. NR = no response by participants on the first training trial.)

Discussion

We set out to investigate whether speakers systematically adjust their gesture to conform to local norms of spatial thinking. Overall, participants reliably adopted either an egocentric or an allocentric frame of reference in their gesture. Participants exhibited an overall bias toward the egocentric frame of reference, consistent with the preferential use of egocentric encoding in English and Spanish, the participants' dominant languages. Critically, however, participants were sensitive to unspoken local norms of spatial thinking. Those who were exposed to an allocentric norm were more likely to adopt an allocentric frame of reference in their spontaneous gestures. Our data do not distinguish whether the local norm caused changes to participants' spatial representations, communicative choices, or both. Whatever the case, these results demonstrate that gesture can function as a public channel through which locally reinforced conventions can be expressed and potentially transmitted.

Gestural relativity: private thought, public gesture

Reasoning about spatial relationships is mostly a private affair — that is, until we ask someone to pass the salt, drive us home, or otherwise think about space in ways that coordinate with our own thinking. One consequence of the privacy of spatial thinking is that two people may unknowingly be thinking about space quite differently, even as they speak or think together. Even within a community, there may be a great deal of hidden variation in conceptions of space, just as there is hidden variation in other aspects of private thought (Lupyan et al., 2023).

Here, we have demonstrated one route by which these private thoughts become public and thus coordinated. Participants' spontaneous co-speech gesture revealed publicly the frame of reference they had adopted privately. Gesture thus offers a signal with which others can coordinate their own thinking. When one's private thinking has been shaped by community norms, gesture can externalize those norms and make them explicit. Given that observers adopt the spatial frame of reference used by speakers in their gestures (Nosrati & Marghetis, 2025), co-speech gesture may help cultivate shared norms, aligning the spatial reasoning of a community.

Recall that, according to one version of linguistic relativity, habitual exposure to spatial language can shape spatial thought, thus creating shared norms of thought within a culture. Our results suggest that gesture may offer a complementary mechanism. According to gestural relativity, habitual exposure to spatial frames of reference in gesture can shape observers' spatial thought. Nosrati & Marghetis (2025) demonstrated that observers are, indeed, influenced by speakers' spatial gestures. Here we have shown that those gestures reflect implicit norms of spatial thinking. Together, these results suggest that gestural relativity is a viable account of cross-cultural variation and within-culture consistency in spatial frames of reference. The present study, however, only demonstrates mechanistic plausibility under controlled conditions; longer-term cultural stabilization likely requires repeated and ongoing interaction.

An ecology of influences on spatial cognition

Although participants adapted their gesture to local norms of spatial thinking, this adaptation was not without constraint. Across conditions, there was an overall bias toward an egocentric frame of reference.

There may be a few reasons for this egocentric bias. It's consistent with the preferential use of egocentric descriptions in English and Spanish, the primary languages spoken by our sample. Brief exposure to local norms may have been insufficient to fully override spatial biases established by a lifetime of language use. Task features may have also constrained the adoption of an allocentric frame of reference. For instance, there may not have been sufficient access to salient landmarks to anchor an allocentric coordinate system (cf., Li & Gleitman, 2002). In our study, concrete pillars and a nearby whiteboard were the most salient landmarks, but visible access to the sky or local topography may have further encouraged an allocentric strategy. Even in a physical space with modest landmarks, however, the allocentric norm shifted participants' gesture.

Moreover, the tasks themselves appeared to shape participants' spatial thinking, with more egocentric reasoning on the Map task. This could reflect an association between maps and personal navigation, prompting participants to imagine themselves navigating the maze. Or the Map task may have been more challenging, which has been argued to increase reliance on one's habitual frame of reference (Levinson et al., 2002). The spatial strategy that somebody adopts may reflect a combination of influences from the task itself, a lifetime of habitual language use, and other cues to local norms of reasoning.

These converging influences on spatial thinking may, over time, establish habits of thought that are stable and shared. The present findings demonstrate that short-term exposure to local norms can shape subsequent thought and gesture. On longer timescales, these and other influences may conspire to establish shared cultural norms within an entire culture. Cultural conventions for spatial thinking, therefore, need not rely on explicit instruction. Implicit social learning, linguistic norms, regularities in gesture, and material artifacts may all conspire together to stabilize and coordinate a community's spatial thinking. Spatial cognition is not a monolith but a distributed ecosystem of habits, strategies, and practices.

Conclusion

This study demonstrates that spontaneous gestures flexibly adapt to locally induced spatial conventions. While gestures are often thought to be idiosyncratic, personal, and ad-hoc, here we have shown that, like language, they are often normative. People spontaneously encoded spatial frames of reference in gesture and did so in ways that reflected implicit local norms. Gesture is thus a viable candidate for communicating implicit spatial norms to those around us. By making implicit norms public through gesture, a community can induct new members into shared ways of thinking about space — and perhaps other conceptual domains, such as time and number. Future work should explore the interrelatedness of spatial gesture and spatial thought on longer timescales and in other cultural contexts.

References

Calderón, E., De Pascale, S., & Adamou, E. (2019). How to speak “geocentric” in an “egocentric” language: A multimodal study among Ngigua-Spanish bilinguals and Spanish monolinguals in a rural community of Mexico. *Language Sciences*, 74, 24-46.

Cooperrider, K., Slotta, J., & Núñez, R. (2017). Uphill and downhill in a flat world: The conceptual topography of the Yupno house. *Cognitive Science*, 41, 768–799.

Haviland, J. B. (1993). Anchoring, Iconicity, and Orientation in Guugu Yimithirr Pointing Gestures. *Journal of Linguistic Anthropology*, 3(1), 3–45.

Le Guen, O. (2011). Speech and Gesture in Spatial Language and Cognition Among the Yucatec Mayas. *Cognitive Science*, 35(5), 905–938.

Levinson, S. C. (1996). Language and space. *Annual Review of Anthropology*, 25(1), 353-382

Levinson, S. C. (2003). *Space in language and cognition: Explorations in cognitive diversity*. Cambridge University Press.

Levinson, S. C., Kita, S., Haun, D. B. M., & Rasch, B. H. (2002). Returning the tables: Language affects spatial reasoning. *Cognition*, 84(2), 155–188.

Li, P., & Gleitman, L. (2002). Turning the tables: Language and spatial reasoning. *Cognition*, 83(3), 265-294.

Lupyan, G., Uchiyama, R., Thompson, B., & Casasanto, D. (2023). Hidden Differences in Phenomenal Experience. *Cognitive Science*, 47(1), e13239.

Majid, A., Bowerman, M., Kita, S., Haun, D. B. M., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Sciences*, 8(3), 108–114.

Marghetis, T., McComsey, M., & Cooperrider, K. (2020). Space in Hand and Mind: Gesture and Spatial Frames of Reference in Bilingual Mexico. *Cognitive Science*, 44(12), e12920.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago press.

Nosrati, S., & Marghetis, T. (2025). Gestural Relativity of Spatial Cognition: Speakers' co-speech gestures shape listeners' spatial frame of reference. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 47).

Pérez Báez, G. (2011). Spatial frames of reference preferences in Juchitán Zapotec. *Language Sciences*, 33, 943–960.

Whorf, B. L. (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. (J. B. Carroll, Ed.). Cambridge, MA: MIT Press.