Title
Rhythmic Coordination Affects Children’s Perspective-Taking during Online Communication

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors
Wan, Yingjia
Wei, Yipu
Xu, Baorui
et al.

Publication Date
2022

Peer reviewed

Rhythmic Coordination Affects Children's Perspective-Taking during Online Communication

Yingjia Wan (wanyj@psych.ac.cn)
CAS Key Laboratory of Behavioral Science, Institute of Psychology, 16 Lincui Road
Department of Psychology, University of Chinese Academy of Sciences
Chinese Academy of Sciences, Beijing, 100101 China

Yipu Wei (weiyipu@pku.edu.cn)
School of Chinese as a Second Language, 5 Yiheyuan Road
Peking University, Beijing, 100871 China

Baorui Xu (xubaorui@pku.edu.cn)
School of Chinese as a Second Language, 5 Yiheyuan Road
Peking University, Beijing, 100871 China

Michael K. Tanenhaus (mtanenha@ur.rochester.edu)
Department of Brain and Cognitive Sciences, 363 Meliora Hall
University of Rochester, Rochester, NY 14627 USA

Abstract

We examined how rhythmic activities affect children's perspective-taking in a referential communication task with 69 Chinese 5- to 6-year-old children. The child first played an instrument with a virtual partner in one of three coordination conditions: synchrony, asynchrony, and antiphase synchrony. Eye movements were then monitored with the partner giving instructions to identify a shape referent which included a pre-nominal scalar adjective (e.g., *big cubic block*). Participants with awareness of their partner's perspective could, in principle, identify the intended referent before the shape was named when the target contrast (a small cubic block) was in shared ground whereas a competitor contrast was occluded for the partner. Children in the asynchrony and antiphase synchrony conditions, but not the synchrony condition, showed anticipatory looks to the target, suggesting that playing instruments asynchronously or in alternation facilitates perspective-taking, likely by training self-other discrimination and inhibitory control.

Keywords: rhythmic coordination; perspective-taking; referential communication; social development; social cognition

Introduction

Perspective-taking—the distinguishing of one's own knowledge from that of others—plays a crucial role in children's social cognitive development, including enabling effective communication and other forms of social interactions. Most existing work measures children's perspective-taking, or "Theory of Mind", by asking children to explicitly provide answers regarding other people's knowledge or desires (e.g., false-belief tasks, Li et al., 2019; Wimmer & Perner, 1983). However, many real-life social interactions require automatic and continuous perspective-taking (Flavell et al., 1981; Surtees & Apperly, 2012; Surtees et al., 2016). One common measure of implicit perspective-taking manipulates physical co-presence (Clark, 1996) in a referential communication task (Keysar et al., 2000). Interlocutors' views of some potential referents differ, allowing researchers to infer whether listeners can exploit perspective differences, to either disambiguate a target referent which is in the common/shared ground from a competitor that is occluded from the interlocutor (Keysar et al., 2000; Savitsky et al., 2011) or anticipate an otherwise temporarily ambiguous referent. The logic of the manipulation is to manipulate physical co-presence while using instructions with prenominal scalar adjectives, which typically refers to a referent whose contrast should also be in the common ground.

Children are sensitive to perspective differences at a young age. Evidence supports a fundamental change in "Theory of Mind" around four when children are first reported to pass standard false-belief tasks (Flavell et al., 1990; Wellman & Bartsch, 1988). However, children's anticipatory looking behaviors in revised false-belief tasks show some awareness of false beliefs among 2- and 3-year-old children (Clements & Perner, 1994; Garnham & Perner, 2001; Garnham & Ruffman, 2001; Southgate et al., 2007) and even 15-month-old infants (Onishi & Baillargeon, 2005).

Implicit perspective-taking in communication, however, has only been observed after four. In a referential communication task that requires ground information for disambiguation, 5- to 6-year-olds could produce and understand instructions with awareness of others' perspectives (Nadig & Sedivy, 2002). With a similar design, Nilsen and Graham (2009; cf. Fan et al., 2015) found that 4- to 5-year-olds and 5- to 6-year-olds were sensitive to their partner's perspective in communication. When prompted to choose a referent that is ambiguous linguistically, they were able to use information that one of two otherwise ambiguous referents was not visible to the speaker, and hence not the

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society.*

intended target. However, for referential communication tasks that do not require ground information for disambiguation (i.e., the instruction is locally but not globally ambiguous (Heller et al., 2008), it is not yet known whether children still use ground information.

Children's perspective-taking skills are argued to be cultivated through joint actions with shared goals (Moll & Tomasello, 2007). Joint rhythmic activities are an example. In children and adults, joint singing, dancing, and moving in time promote prosocial behaviors (Cirelli et al., 2014; Kirschner & Tomasello, 2010; Rabinowitch & Melzoff, 2017). Many argue that synchrony is the mechanism underlying these positive social effects (e.g., Hove & Risen, 2009; Reddish et al., 2014). Participants making synchronous movements develop a stronger sense of unity and feeling of interdependence, thus becoming more inclined to cooperate, help and share in subsequent social interactions (Kirschner & Tomasello, 2010). Subsequent research found that antiphase synchrony (moving in alternation as opposed to mirroring each other) increased prosociality to the same extent (Cirelli et al., 2014; Wan & Zhu, 2021). Synchronous rhythmic coordination also increases adults' self-reported mentalizing tendencies and skills toward their synchronizing partners, although accuracy of mental state recognition was unaffected (Baimel et al., 2018). Recent developmental research finds that cooperative interaction in problem-solving tasks improves preschoolers' performance on tasks that require representing others' wishes (Jin et al., 2018) and visual perceptions (Li et al., 2019). However, effects of rhythmic activities on spontaneous perspective-taking remain unclear.

Here, we examine how joint rhythmic activities affect children's spontaneous perspective-taking during online language comprehension. The link between synchronous coordination and communication is well-documented. As people converse, syntactic structures and accents become more similar (Branigan et al., 2000; Giles et al., 1992), body movements more synchronized (Shockley et al., 2003; Chartrand & Bargh, 1999), and eye movements more coordinated (Richardson et al., 2007).

Given the positive social effect of synchronous rhythmic activities and the link between coordination and language comprehension, synchronous music-making could facilitate perspective-taking during communication. Alternatively, synchronous behavioral matching could inhibit perspective-taking. Perspective-taking often involves overcoming the "curse of knowledge", that is, inhibiting the interference of one's own conflicting view (Friedman & Leslie, 2005). Children below the age of 4.5 to 5 have difficulty passing false-belief tasks (Ghrear et al., 2021; Wellman et al., 2001). One likely contributing factor is their still-developing inhibitory control (Bernstein et al., 2007; Coolin et al., 2015; Lagattuta et al., 2014; Nilsen & Graham, 2009). While interlocutors generally use perspective information in language comprehension and production (Brown-Schmidt et al., 2008; Hanna & Tanenhaus, 2004; Hanna et al., 2003; Heller et al., 2008; Nadig & Sedivy, 2002), they at times make egocentric errors (Keysar et al., 2000). One explanation is that those with poor inhibitory control have difficulty inhibiting perspective-inappropriate interpretations (Brown-Schmidt, 2009). Further research demonstrated that inhibition training could improve spontaneous perspective-taking in language comprehension in adults (Santiesteban et al., 2012) and children (Kampis et al., 2021). Participants in the inhibition-of-imitation training group were asked to do the opposite of what a person on the screen did, and they outperformed a group trained to imitate in a subsequent referential communication task. The result is attributed to an increase in self-other control during the inhibition-of-imitation training, leading to better inhibition of the self-perspective in later communication. Synchronous behavioral matching, on the other hand, may increase self-other overlap and thus lead to difficulty differentiating one's own and others' views. Therefore, rhythmic activities involving inhibiting imitation (i.e., making sounds asynchronously or oppositely) might facilitate perspective-taking in language comprehension more than synchronous performance.

To test the above hypotheses, we first manipulated the type of rhythmic coordination a child had with a computer partner, and then examined the child's spontaneous perspective-taking in a referential communication task. In the music task, the child played the drum synchronously, asynchronously, or antiphase-synchronously with the partner.

For the referential communication task, we used a modified screen-based version of Heller et al.'s (2008) design and monitored children's eye movements as they heard referring expressions with pre-nominal scalar adjectives (e.g., big/small). A pre-nominal scalar adjective assumes a contrast. Thus, *pick up the big cube* would be uttered only when there is also a small cube. When there are only two cubes, the listener can identify the referent after hearing *big*. However, if there were also a big and a small blue triangle, then the listener would need to hear the shape name. Now consider what would happen if the small triangle was visible to the listener but not the speaker. A listener who was using perspective information could identify the intended referent upon hearing *big,* whereas a listener who was not aware that the speaker couldn't see the small triangle would not be able to identify the referent until the shape was named. Thus, the timing of referent identification, as indexed by eye-movements, indicates whether or not the child was spontaneously using differences in perspective to rapidly disambiguate a temporarily ambiguous referring expression — that is an expression which would soon be unambiguous (unlike Nadig & Sedivy, 2002) even without using perspective differences (see detailed design in the Method).

An important difference between this current experiment and the previous ones is the use of a more complex display setting: six objects were presented on a nine-grid shelf, in contrast to four objects in four grids used in Heller et al.'s and Nadig and Sedivy's experiments. According to Apperly et al. (2010), complex displays with more objects may give rise to more egocentric behaviors. The complex display in this current study provides an opportunity to explore the upper limit of children's perspective-taking skills.

We first tested whether 5- to 6-year-old children use ground information spontaneously in communication. If the speaker uses a scalar adjective *big* to refer to the target, it is assumed that the speaker knows that there is a small counterpart that forms the contrast. If the children can spontaneously use ground information, at the time of hearing a scalar adjective, they should show more anticipatory looks to the target, which has a size contrast object visible to both sides, in comparison to a competitor, which also fits the description of *big* but has its size contrast excluded from the view of the interlocutor (see example in Figure 3), give that it would be infelicitous for the interlocutor to describe an object with a scalar adjective if its contrast is not visible.

Our second goal was to examine how different types of coordination in rhythmic coordination influence spontaneous perspective-taking. Asynchronous coordination could exercise inhibitory control and enhance self-other distinction. Antiphase synchronous coordination has prosocial effects that are equivalent to synchrony, and it also involves inhibiting the tendency to mimic the partner. Thus, if inhibition is important, then asynchronous and antiphase synchrony should pattern together.

Pilot work established that 5- to 6-year-olds, but not younger children were capable of completing both tasks with minimal guidance, so we chose this age group.

Method

Study Design

We used a 3 × 2 mixed design. The between-participant variable was coordination condition (synchrony vs. antiphase synchrony vs. asynchrony); the within-participant variable was ground (shared vs. privileged). Participants worked with the same computer partner to complete the manipulation and test tasks, but they were told before the experiments that the partner was another child from a different kindergarten.

Participants

The participants included 69 children from a kindergarten in Nanjing, China. Data from 14 children were excluded due to calibration failure (n = 8) or poor quality of eye movement data (n = 6). The final sample consisted of 55 children (mean age = 6.2, range: 5.7-6.6, 27 females). Kindergarten and parents provided consent. Ethical approval was obtained from the first author's host institution. Children were randomly assigned to one of the three coordination groups. Each participant was compensated with a cartoon eraser.

Apparatus

We used an EyeLink Portable Duo (SR Research), sampling at 1000 Hz. Tasks were controlled and recorded by Experiment Builder (version 2.3.38, SR Research). We also provided participants with a chin rest for better head positioning.

Procedures

Manipulation Phase – Percussion game. Children played a two-player percussion game with a computer partner (whom participants believed to be another child). Children were presented with an interface including two avatars: the avatar in yellow on the left side represented the participant, while the avatar in pink on the right side was the partner. Figures of musical instruments fell off from the top of the screen, hit a horizontal line at about two-thirds of the vertical line that separated the screen, and fell into the bag at the bottom of the screen (see Figure 1). When the figures fell onto and crossed the horizontal line, the half of the line on the participant's side turned red and the other half on the partner's side turned green. The instrument that the participant played in the experiment was a snare drum that required pressing a key to make a drum sound, while the computer partner's instrument was a cymbal that played a sound automatically when hitting the line. The participants were instructed to press the button every time the drum figure hits the line.

Prior to the main experiment, children were provided with a practice section of (1) familiarizing themselves with the instruments and button-pressing; (2) practicing percussion on the drum at a speed of 30 beats per minute (henceforth bpm), 40 bpm, and 50 bpm without the presence of the partner; (3) watching the intended duet performance in their experimental condition, with the presence of the partner on the screen; and (4) practicing the drum part in a two-player setting (the partner's avatar is present but does not play, indicating that the partner has not entered the game yet). Children were then presented with a "waiting for the other player to join" interface designed to make the game more realistic. Once the partner "entered", the main experiment began.

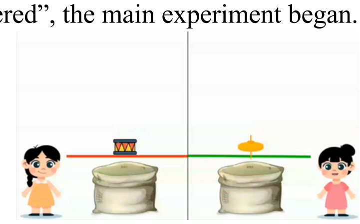


Figure 1: Example display of percussion game (the two-player setting for female participants).

We manipulated the rhythmic patterns between the two musical instruments by adjusting the cymbal's time or speed of falling. In the synchrony condition, the cymbal and the drum fell off at the same starting time and speed; in the antiphase synchrony condition, the cymbal fell at the same speed as the drum, but the start-time is half a beat later, so the dyad played in alternation; in the asynchrony condition, the cymbals fell at a speed 30% faster than the drum.

In the main experiment, participants were required to complete four sessions at a speed of 45 bpm, 50 bpm, 55 bpm, and 60 bpm sequentially in each experimental condition (rhythmic pattern). The duration of each session was one minute. The participants' eye movements and button-pressing time were recorded for analysis.

Test phase: Online referential communication task. This perspective-taking task consisted of 16 experimental trials and 16 filler trials. The first two trials were always filler trials, while the rest of the trials were randomized. Each trial paired (a) an auditory instruction (see Figure 2) with (b) a visual scene including two cartoon avatars and a nine-grid shelf (see Figure 3). Before the main task, participants learned about the perspective difference, the shapes of blocks, and the use of a mouse in a warm-up session.

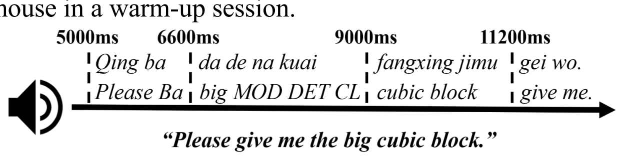
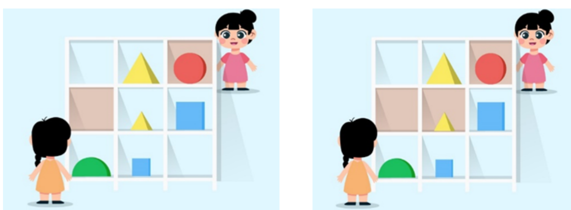


Figure 2: Example auditory instruction (dashed lines indicate the sound onset of speech segments).



Shared-ground condition *Privileged-ground condition*

Figure 3: Example displays of two ground conditions in the online perspective-taking task (for girls) paired with the instruction n Figure 2. Four areas of interest were coded: target (the big blue cubic block), competitor (the big yellow triangle block), target-contrast (the small blue cubic block), and competitor-contrast (the small yellow triangle block).

After a 5-second preview, participants heard pre-recorded instructions. The audio source was from an adult female speaker of Mandarin Chinese, and the pitches were adjusted to simulate a child's voice using *Audacity* (version 3.0.0). To prevent the co-articulation effect on auditory language processing (Magnuson et al., 2003), each part of the auditory instructions was recorded independently and then assembled to make a complete sentence. The auditory sentences were double-checked with native speakers on naturalness.

In a visual scene, the avatar of the participant faces the shelf, whereas the partner stands on the other side of the shelf. Grids with light brown shadows blocked the partner's view. There were six blocks in each visual scene, four of which were visible to both avatars in the privileged condition and five in the shared condition. In the privileged condition, the competitor-contrast was only visible to the participants; while in the shared condition, the competitor-contrast was visible to both sides. We randomized the grid positions of the target objects and ensured that the target object appeared in the same position no more than three times; the positions of other objects were randomized, but the number of times they appeared in the same position was not controlled. The size, color, and shape of all objects are balanced across trials.

Results

Percussion game. Performance in the percussion game was rated on how accurately they played to the designed rhythm in the four main sessions. A button press was counted as accurate if it took place when the drum figure was passing through the horizontal line (a 550s time interval). While accuracy in synchrony and anti-phase synchrony conditions is higher than that of the asynchrony condition, children in all conditions performed reasonably well. (synchrony: 95.31%; asynchrony: 90.07%; antiphase synchrony: 95.92%).

Referential communication task. Children on average answered 99% of the experimental trials correctly. Figure 4 exhibits the change of proportion of looks to the target (the big blue cubic block in Figure 3), competitor (the big yellow triangle block), and target contrast (the small blue cubic block) under the two ground conditions (privileged vs. shared) over time. Under the privileged condition, the proportion of looks to the target (the red concrete line) diverges from that to the competitor (the blue concrete line) much earlier compared to the shared condition (the red dot-dash line vs. the blue dot-dash line), indicating an effect of ground.

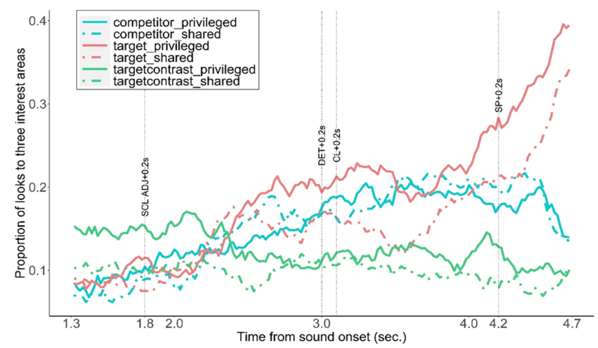


Figure 4: Proportion of looks to three interest areas under two ground conditions (privileged vs. shared). Four vertical dotted lines represent the onset (200ms added) of the scalar adjective, determiner, classifier, and shape adjective.

The critical time window for analysis (Figure 5) is from 1.8s (the onset of scalar adjective plus 200ms) to 4.2s (the onset of the shape adjective plus 200ms). The ground effect and the influence of coordination on the ground effect are evaluated in terms of target ratio[1] and target-set ratio[2]. The reason to include a measure of target-set ratio is twofold. First, when processing a scalar adjective (e.g., 'big'), children tended to continuously look at both the target (e.g., a big cubic) and the contrast of the target item (e.g., a small cubic),

[1] Target ratio = proportion of looks to the target / (proportion of looks to the target + proportion of looks to the competitor).

[2] Target-set ratio = proportion of looks to the target-set / (proportion of looks to the target-set + proportion of looks to the competitor-set). The target-set includes the target and the target-contrast; the competitor-set includes the competitor and the competitor-contrast.

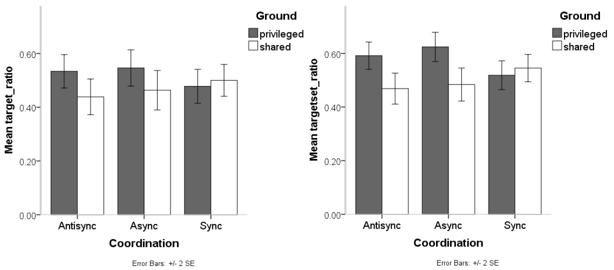


Figure 5: Mean target ratios (left) and target-set ratios (right) for the critical time window in privileged and shared conditions across coordination groups.

unlike adults who tend to focus on the target and only briefly attend to the contrast. Second, as children's fixations on the screen were much more sparsely distributed than adults, the trials where they were looking at neither the target nor the competitor during the critical time window took up 20.7% of the total trials, whereas cases in which participants did not look at the target-set or the competitor-set only took up 7.3% of the trials. Adding target-set ratio as a measure could include more observations, and thus, increase the statistical power of the analysis. Figure 5 shows the target ratios and target-set ratios under different conditions.

We analyzed the target ratio and target-set ratio as a function of ground and coordination using a linear mixed-effect model with lme4 (Bates et al., 2015) in R (R Core Team, 2021). The two analytical models include respectively target ratio and target-set ratio as dependent variables, ground, coordination, ground*coordination as fixed effects, and random intercepts of subject and item as random effects. Children's age in months and accuracy of performance in the music task are added to the models as covariates. Table 1 and Table 2 present the regression coefficients of the two models (reference levels: shared ground and antiphase synchrony coordination). There was no difference in accuracy or reaction time across conditions.

Table 1: Summary of the analysis on target ratio

Variable	β	*SE*	*t*	*p*	β 95%CI
Covariates					
Month	.001	.004	0.30	.77	[-.01, .01]
Accuracy	-.04	.32	-0.12	.90	[-.66, .59]
Fixed effects					
Ground (privileged vs shared)	.10	.04	2.47	.01*	[.02, .19]
Coordination (sync vs anti-sync)	.06	.04	1.42	.16	[-.02,.14]
Coordination (async vs anti-sync)	.03	.05	0.63	.53	[-.06, .12]
Ground (privileged vs shared) *Coordination (sync vs anti-sync)	-.12	.06	-1.95	.05	[-.23, .00]
Ground (privileged vs shared) *Coordination (async vs anti-sync)	-.03	.06	-0.45	.66	[-.15, .09]
Random effects	Variance		*Std. Dev.*		
Subject	.00		.00		
Item	.02		.14		
Fit of goodness			R^2 (conditional)= .17		

* $p < .05$, ** $p < .001$

Table 2: Summary of the analysis on target-set ratio

Variable	β	*SE*	*t*	*p*	β 95%CI
Covariates					
Month	-.002	.003	-0.82	.42	[-.01, .00]
Accuracy	.15	.26	0.56	.58	[-.36, .65]
Fixed effects					
Ground (privileged vs shared)	.13	.03	3.60	<.001**	[.06, .19]
Coordination (sync vs anti-sync)	.07	.03	1.98	.05	[.00, .14]
Coordination (async vs anti-sync)	.03	.04	0.74	.46	[-.05, .10]
Ground (privileged vs shared) *Coordination (sync vs anti-sync)	-.14	.05	-2.90	.004*	[-.23, -.05]
Ground (privileged vs shared) *Coordination (async vs anti-sync)	-.003	.05	-0.06	.95	[-.10, .09]
Random effects	Variance		*Std. Dev.*		
Subject	.00		.01		
Item	.02		.15		
Fit of goodness			R^2 (conditional) = .23		

* $p < .05$, ** $p < .001$

The effect of ground is significant in both target ratio and target-set ratio measures given the current reference level (antiphase synchrony coordination). Higher target ratio and target-set ratio are found for the privileged ground condition, in contrast to the shared ground condition (target ratio: $\beta = .10$, $SE = .04$, $t(724) = 2.47$, $p = .01$, 95%CI [.02, .19]; target-set ratio: $\beta = .13$, $SE = .03$, $t(725) = 3.60$, $p < .001$, 95%CI [.06, .19]). The asynchrony condition does not differ from the antiphase synchrony condition in terms of the ground effect (target ratio: $\beta = -.03$, $SE = .06$, $t(724) = -.45$, $p = .66$, 95%CI [-.15, .09]; target-set ratio: $\beta = -.003$, $SE = .05$, $t(746) = -.06$, $p = .95$, 95%CI [-.10, .09]), suggesting a similar ground effect in these two coordination groups.

However, the difference in target-set ratio between the privileged and shared condition is smaller in the synchrony compared to the antiphase synchrony condition, as evidenced by an interaction of coordination and ground ($\beta = -.14$, $SE = .05$, $t(743) = -2.90$, $p = .004$, 95%CI [-.23, -.05]). The interaction effect in target ratio is in the same direction and marginally significant ($\beta = -.12$, $SE = .06$, $t(724) = -1.95$, $p = .05$, 95%CI [-.23, .00]). When the synchrony coordination is set as the reference level of the model, ground effects also differ from the asynchrony condition: There is a significant interaction between ground and coordination in target-set ratio ($\beta = .14$, $SE = .05$, $t(744) = 2.81$, $p = .005$, 95%CI [.04, .23]). Re-setting the reference level to the synchrony coordination demonstrates that ground effect is insignificant (target ratio: $\beta = -.01$, $SE = .04$, $t(724) = -.26$, $p = .797$, 95%CI [-.09, .07]; target-set ratio: $\beta = -.01$, $SE = .03$, $t(741) = -.43$, $p = .668$, 95%CI [-.08, .05]) in the synchrony condition.

In sum, the ground effect measured by the target ratio and target-set ratio is significant for both the antiphase synchrony group and the asynchrony group, with no significant difference between these two groups. The synchrony condition, however, does not show a ground effect.

Discussion

We demonstrated perspective-taking for 5- to 6-year-old children in a referential communication task where the use of ground information allows for earlier identification of a referent as indexed by anticipatory looks to the target. Importantly, the ground effect was influenced by the type of rhythmic coordination prior to the referential communication task. Children in asynchrony and antiphase synchrony conditions showed more anticipatory looks when the competitor's size contrast was occluded from the partner's view, whereas children in the synchrony condition did not use perspective information to anticipate the target.

The advantage of asynchrony over synchrony in facilitating perspective-taking in real-time language processing contrasts with the positive social effects of synchrony reviewed earlier. One possibility is that synchronous experience emphasizes unity and blurs self-other distinction, leading people to overestimate similarity and overlook perspective differences in their views.

The asynchrony vs. synchrony comparison is consistent with previous studies in which anti-mimicry but not mimicry improves perspective-taking (Santiesteban et al., 2012). Synchrony, like mimicry, is a default mode of coordination requiring limited inhibition, whereas moving out of synchrony requires self-control (Finkel et al., 2006; Rauchbauer et al., 2020). Indeed, in the manipulation phase, children's initial tendency was to play synchronously. Children in the asynchrony and the antiphase synchrony conditions initially struggled to play their parts as instructed. In Santiesteban et al. (2012), anti-mimicry enhanced performance in the referential communication task but not another mentalizing task that does not require spontaneously representing conflicting views. This suggests that exercise of inhibitory control in joint activities could mediate its social effect. Future research should examine how rhythmic activities affect children's perspective-taking in tasks with varying degrees of requirement for inhibitory control.

As hypothesized, children in the antiphase synchrony condition used perspective-taking in reference resolution, outperforming the synchrony group and performing as well as the asynchrony group. Thus, antiphase synchrony appears to have both synchrony's effect of promoting prosociality (Cirelli et al., 2014) and asynchrony's effect of facilitating perspective-taking, which could enhance team performance while maintaining self-other boundaries.

The coordination findings contribute to a growing body of research challenging the "similar is better" view by providing evidence that interpersonal coordination with degrees of freedom may be more beneficial to interpersonal interaction (Fusaroli et al., 2012; Abney et al., 2015; Wallot et al., 2016). According to the interpersonal synergies perspective, coordination should be regarded as a dynamic soft-ensemble that maintains unity in task-relevant aspects but allows variability in task-irrelevant areas (Riley et al., 2011). Take joint music activities as an example. Participants could follow the same tempo but play different instruments with different notes and rhythmic patterns, with the goal of complementing each other's performance rather than playing in perfect unison. Future research could further explore how rhythmic activities with more variability and dynamics affect social cognitive processes in adults and children.

In conclusion, this study contributes to the growing literature on children's social cognitive development and the factors that contribute to this process. Children aged 5- to 6-year-olds spontaneously use perspective information in language comprehension even when it is not necessary for completing the task, suggesting that perspective-taking influences moment-by-moment processing. This study is the first to show that coordination in rhythmic activities can influence children's performance in subsequent linguistic tasks. Moreover, perspective-taking was facilitated by prior asynchronous and antiphase synchronous rhythmic interactions, but not by synchronous coordination. This is likely because moving out of synchrony highlights self-other distinction and trains inhibitory control. The finding challenges the common "similar is better" view, suggesting that maintaining self-other boundary and individuality may be beneficial for some aspects of social interactions.

References

Abney, D. H., Paxton, A., Dale, R., & Kello, C. T. (2015). Movement dynamics reflect a functional role for weak coupling and role structure in dyadic problem solving. *Cognitive Processing*, *16*(4), 325-332.

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-Like states? *Psychological Review*, 116(4), 953–970.

Apperly, I. A., Carroll, D. J., Samson, D., Humphreys, G. W., Qureshi, A., & Moffitt, G. (2010). Why are there limits on theory of mind use? Evidence from adults' ability to follow instructions from an ignorant speaker. *Quarterly Journal of Experimental Psychology*, 63(6), 1201–1217. Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*, 67(1), 1–48.

Baimel, A., Birch, S. A., & Norenzayan, A. (2018). Coordinating bodies and minds: Behavioral synchrony fosters mentalizing. *Journal of Experimental Social Psychology, 74*, 281-290.

Bernstein, D. M., Atance, C., Meltzoff, A. N., & Loftus, G. R. (2007). Hindsight bias and developing theories of mind. *Child Development*, *78*(4), 1374-1394.

Branigan, H.P., Pickering, M.J., & Cleland, A.A. (2000). Syntactic coordination in dialogue. *Cognition*, *75,* B13–B25.

Brown-Schmidt, S. (2009). The role of executive function in perspective taking during online language comprehension. *Psychonomic Bulletin & Review*, *16*(5), 893-900.

Brown-Schmidt, S., Gunlogson, C., & Tanenhaus, M. K. (2008). Addressees distinguish shared from private information when interpreting questions during interactive conversation. *Cognition*, *107*(3), 1122-1134.

Chartrand, T.L., & Bargh, J.A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*, 893–910.

Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014). Interpersonal synchrony increases prosocial behavior in infants. *Developmental Science*, *17*(6), 1003-1011.

Clark, H. H. (1996). *Using Language*. Cambridge University Press.

Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, 9(4), 377–395.

Coolin, A., Erdfelder, E., Bernstein, D. M., Thornton, A. E., & Thornton, W. L. (2015). Explaining individual differences in cognitive processes underlying hindsight bias. *Psychonomic Bulletin & Review*, *22*(2), 328-348.

Fan, S. P., Liberman, Z., Keysar, B., & Kinzler, K. D. (2015). The Exposure Advantage: Early Exposure to a Multilingual Environment Promotes Effective Communication. *Psychological Science*, 26(7), 1090–1097.

Finkel, E. J., Campbell, W. K., Brunell, A. B., Dalton, A. N., Scarbeck, S. J., & Chartrand, T. L. (2006). High-maintenance interaction: Inefficient social coordination impairs self- regulation. *Journal of Personality and Social Psychology*, *91*(3), 456.

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual-perception – Further evidence for the level 1-level 2 distinction. *Developmental Psychology*, *17*, 99–103.

Flavell, J. H., Flavell, E. R., Green, F. L., & Moses, L. J. (1990). Young children's understanding of fact beliefs versus value beliefs. *Child Development*, 61(4), 915–928.

Friedman, O., & Leslie, A. M. (2005). Processing demands in belief-desire reasoning: Inhibition or general difficulty?. *Developmental Science*, *8*(3), 218-225.

Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., & Tylén, K. (2012). Coming to terms: Quantifying the benefits of linguistic coordination. *Psychological science*, *23*(8), 931-939.

Garnham, W. A., & Perner, J. (2001). Actions really do speak louder than words - But only implicitly: Young children's understanding of false belief in action. *British Journal of Developmental Psychology*, 19(3), 413–432.

Garnham, W. A., & Ruffman, T. (2001). Doesn't see, doesn't know: is anticipatory looking really related to understanding of belief? *Developmental Science*, 1, 94–100.

Ghrear, S., Baimel, A., Haddock, T., & Birch, S. A. (2021). Are the classic false belief tasks cursed? Young children are just as likely as older children to pass a false belief task when they are not required to overcome the curse of knowledge. *PloS one*, *16*(2), e0244141.

Giles, H., Coupland, N., & Coupland, J. (1992). Accommodation theory: Communication, context and consequences. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of accommodation*. Cambridge, England: Cambridge University Press.

Hanna, J. E., & Tanenhaus, M. K. (2004). Pragmatic effects on reference resolution in a collaborative task: Evidence from eye movements. *Cognitive Science*, *28*, 105-115.

Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory & Language*, *49*, 43-61.

Heller, D., Grodner, D., & Tanenhaus, M. K. (2008). The role of perspective in identifying domains of reference. *Cognition*, 108(3), 831–836.

Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition*, *27*(6), 949-960.

Jin, X., Li, P., He, J., & Shen, M. (2018). How you act matters: The impact of coordination on 4-year-old children's reasoning about diverse desires. *Journal of Experimental Child Psychology*, *176*, 13-25.

Kampis, D, Duplessy, H.L., & Southgate, V. (2021). Training self-other distinction facilitates perspective taking in young children. *PsyArXiv*.

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, *11*(1), 32-38.

Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior, 31*(5), 354–364.

Lagattuta, H. K., Sayfan, L., & Harvey, C. (2014). Beliefs about thought probability: Evidence for persistent errors in mindreading and links to executive control. *Child Development*, *85*(2), 659-674.

Li, P., Jin, X., Liao, Y., Li, Y., Shen, M., & He, J. (2019). Cooperation turns preschoolers into flexible perspective takers. *Cognitive Development*, *52*, 100823.

Magnuson, J. S., McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2003). Lexical effects on compensation for coarticulation: The ghost of Christmash past. *Cognitive Science*, *27*(2), 285–298.

Moll, H., & Tomasello, M. (2007). Cooperation and human cognition: The Vygotskian intelligence hypothesis. *Philosophical Transactions of the Royal Society of London: Biological Sciences*, *362*, 639–648.

Nadig, A. S., & Sedivy, J. C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science*, 13(4), 329–336.

Nilsen, E. S., & Graham, S. A. (2009). The relations between children's communicative perspective-taking and executive functioning. *Cognitive Psychology*, 58(2), 220–249.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308(5719), 255–258.

R Core Team (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rabinowitch, T. C., & Meltzoff, A. N. (2017). Synchronized movement experience enhances peer cooperation in preschool children. *Journal of Experimental Child Psychology*, *160*, 21-32.

Rauchbauer, B., Dunbar, R. I., & Lamm, C. (2020). Being mimicked affects inhibitory mechanisms of imitation. *Acta Psychologica*, *209*, 103132.

Reddish, P., Bulbulia, J., & Fischer, R. (2014). Does synchrony promote generalized prosociality? *Religion, Brain and Behavior*, *4*(1), 3-19.

Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, *29*(6), 1045–1060.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination: Common ground and the coupling of eye movements during dialogue. *Psychological Science*, *18*(5), 407–413.

Riley, M. A., Richardson, M., Shockley, K., & Ramenzoni, V. C. (2011). Interpersonal synergies. *Frontiers in psychology*, *2*, 38.

Santiesteban, I., White, S., Cook, J., Gilbert, S. J., Heyes, C., & Bird, G. (2012). Training social cognition: from imitation to theory of mind. *Cognition*, *122*(2), 228-235.

Savitsky, K., Keysar, B., Epley, N., Carter, T., & Swanson, A. (2011). The closeness-communication bias: Increased egocentrism among friends versus strangers. *Journal of Experimental Social Psychology*, 47(1), 269–273.

Shockley, K., Richardson, D., & Dale, R. (2009). Conversation and Coordinative Structures. *Topics In Cognitive Science*, *1*(2), 305-319.

Shockley, K., Santana, M.V., & Fowler, C.A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 326–332.

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, 18(7), 587–592.

Surtees, A. D., & Apperly, I. A. (2012). Egocentrism and automatic perspective taking in children and adults. *Child Development*, *83*(2), 452–460.

Surtees, A., Apperly, I., & Samson, D. (2016). I've got your number: Spontaneous perspective-taking in an interactive task. *Cognition*, *150*, 43–52.

Wallot, S., Mitkidis, P., McGraw, J. J., & Roepstorff, A. (2016). Beyond synchrony: joint action in a complex production task reveals beneficial effects of decreased interpersonal synchrony. *PloS one*, *11*(12), e0168306.

Wan, Y., & Zhu, L. (2021). Effects of rhythmic turn-taking coordination on five-year-old children's prosocial behaviors. *Developmental psychology*, *57*(11), 1787–1795.

Wellman, H. M., & Bartsch, K. (1988). Young children's reasoning about beliefs. *Cognition*, *30*(3), 239–277.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*(3), 655-684.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103–128.