Received: 8 December 2022 | Revised: 3 October 2023 | Accepted: 8 October 2023

DOI: 10.1111/sode.12719

ORIGINAL ARTICLE

WILEY

Linguistic but not minimal group membership modulates spontaneous level-2 perspective interference in 8-year-old children

Lívia Priyanka Elek[1] | Ildikó Király[1] | Réka Pető[1] | Renáta Szücs[1] | Fruzsina Elekes[1,2] | Katalin Oláh[1]

[1]MTA-ELTE Social Minds Research Group, Psychology Institute, ELTE Eötvös Loránd University, Budapest, Hungary

[2]Department of Cognitive Science, Central European University, Budapest, Hungary

Correspondence
Ildikó Király, MTA-Momentum Social Minds Research Group, Psychology Institute, ELTE Eötvös Loránd University, Izabella 46, Budapest - 1064, Hungary.
Email: kiralyi@caesar.elte.hu

Funding information
Momentum Program of the Hungarian Academy of Sciences, Grant/Award Number: LP-2017-17/2017

Abstract

This paper presents evidence that social categorization affects spontaneous level-2 visual perspective taking (L2PT) differently depending on the type of social category in 8-year-old. In Experiment 1 (N = 46), children were paired with same-age peers, who belonged to the same or a different minimal group. In Experiment 2 (N = 42) children participated with an adult confederate, who either shared their cultural group membership or was a member of an out-group (inferred from a linguistic cue, accent). In Experiment 3 (N = 80), children were acting together with an adult confederate who was at the same time a member of the same or a different minimal group and shared their linguistic membership. This allowed us to investigate how these social categories influence each other. Spontaneous L2PT was not affected by the minimal group manipulation. However, accent weakened L2PT when it implied that the task partner belonged to an out-group. When both category cues were present, accent—that could be an indicator of shared knowledge attribution—played a more pronounced role in attenuating L2PT. It is argued that social categories that are indicative of the partner's knowledge states but not ad hoc groups influence spontaneous mentalizing.

KEYWORDS
linguistic cue, social categorization, spontaneous level-2 perspective taking

1 | INTRODUCTION

Theory of Mind (ToM), the ability to represent what others see, know, believe or desire -with one term, the ability to mentalize—has traditionally been investigated independently of other processes of the mind, with the main purpose of identifying the core components of this capacity (Flavell, 1999; Frith & Frith, 2005). However, ToM may not be as isolated from other activities of the mind as it has been depicted in empirical research. Undoubtedly, Theory of Mind functions in parallel with and in the context of other cognitive mechanisms that function to cope with social partners. Hence, the possibility arises that these processes provide information for the mindreading system, influencing the specific mental state attributions being performed. Such an important source of information about the target of mentalizing may be the person's social group membership, and the process identifying it: social categorization (Spaulding, 2018; Westra, 2017).

Recent philosophical accounts have proposed different routes through which information about another person's social category membership may influence mentalizing. Spaulding (2018) claims that information derived from one's social category membership affects mentalizing by filtering information that enters a specific mindreading scenario. She argues that information implicitly associated with each social category will become salient over others in specific contexts. The process may be further influenced by one's priority to mentalize accurately or efficiently, relying on deliberately collected information or heuristics, respectively. Westra (2017) argues that social categorization activates stereotypes. Stereotypes are representations of oversimplified opinions and attitudes in relation to others, shared by the members of a group, the core structures of which are character-traits. These, being stable, unobservable psychological properties, have consistent effects on individuals' behavior, and can thus serve as the basis of behavior prediction. The mechanisms described in these theories promote behavior prediction through generalized expectations based on group membership. They provide *shortcuts* to behavior prediction, by constraining the selection process that determines which of the possible mental states (intentions or beliefs) an agent is most likely to hold. For these mechanisms to unfold the mentalizer needs to have relatively elaborate expectations and attitudes (stereotypes) about *specific* social groups. Indirectly, these theories take a stand for a late emerging interplay between mentalizing and social categorization that only manifests itself towards out-groups that the mentalizer has knowledge about.

Specifying the effects of group membership on mindreading is particularly challenging as both mental states and group distinctions may vary along a number of dimensions. Consequently, not all group distinctions carry the same relevance for a given mindreading task. For example, being thin or overweight should not necessarily affect how a person attributes false beliefs on the location of an object but may alter attributed emotions on topics related to eating or cultural stigma. In the domain of cognitive theory of mind, group distinctions that can be directly linked to an identifiable body of knowledge may be of special relevance as these tasks require participants to compute the knowledge state of the social partner. As proposed by Oláh et al. (2019) social categories can influence mentalizing by providing guidance about the range of knowledge its members can be expected to have. At a developmentally early stage this may only entail tracking *whether* someone shares the knowledge that features the child's in-groups without necessarily specifying *how* the out-groups differ. There is evidence that from early childhood children expect certain social categories to denote the borders of shared knowledge (see, e.g., Kinzler et al., 2012; Oláh et al., 2014; Soley & Spelke, 2016). The fast mapping of knowledge/ignorance states is of great importance to young children who need to acquire culturally valid information specifically from individuals who are in possession of such knowledge. Undoubtedly, not all social category distinctions carry the same weight in this respect. Linguistic cues, for example, may be seen as direct and reliable

markers of cultural knowledge (Kinzler et al., 2012; Pető et al., 2018) whereas, for example, hair color or clothing may only gain such significance through associations.

In this study, our goal is to test the idea that arbitrarily created group distinctions would not similarly interfere with mindreading processes as group distinctions that inherently carry relevant information for the task at hand. Specifically, we will focus on how a perceived mismatch of group affiliations that do or do not map onto relevant knowledge differences, affects a basic social function, spontaneous visual perspective taking.

When reasoning about others' visual perspective, the observer may limit their computation to *which objects* another person can see (level-1 perspective taking, L1PT), a capacity which emerges already in the second year of life (Moll & Tomasello, 2006; Luo & Baillargeon, 2007). Assessing *how* an object *appears* from another person's point of view (level-2 perspective taking, L2PT) is a more demanding computation and the ability to explicitly report on level-2 perspectives emerges developmentally later, between 4 and 5 years of age (Flavell et al., 1981; Gzesh & Surber, 1985; Moll & Meltzoff, 2011). Precursors to L2PT can, however, be traced in implicit task contexts even in the second and third years of life (Luo & Beck, 2010; Moll et al., 2013).

Recently, both skills have been found to occur spontaneously, that is, without explicit motivation to do so, in online tests of perspective taking, both in adults (L1PT: Samson et al., 2010; L2PT: Elekes et al., 2016; Freundlieb et al., 2018; Surtees et al., 2016; Yuan et al., 2017), and in 8–9.5-year-old children (LPT1: Surtees & Apperly, 2012; Surtees et al., 2012; LPT2: Elekes et al., 2017). In the above paradigms, spontaneous PT is assumed to occur when the social partner's conflicting perspective interferes with one's own when one is only supposed to consider their own perspective and ignore the other's. This suggests that the other's perspective has been calculated spontaneously, without instruction to do so. The possibility that such online forms of perspective taking may be affected by social categorization is especially intriguing as it would demonstrate flexibility in a speeded, implicit capacity.

A number of studies have started to explore this question, although, as of yet, research is constrained to level-1 perspective taking and to adult participants. In the most common level-1 task, the dot perspective task (Samson et al., 2010), an avatar is displayed in a virtual room with dots on the walls. The avatar, due to their spatial orientation (perspective) sees the same or different number of dots as the participant. Participants make decisions about the number of dots, either from their own or from the avatar's perspective. Results suggest that a mismatch between percepts influences reaction times when responding both from the avatar's and the participant's own perspective.

Two studies (Simpson & Todd, 2017; Schneider et al., 2018) using different group manipulations (university mascot and personality type or nationality, respectively) have found that group membership influences how the participant's own perspective interferes with responding from someone else's, although results are contradictory with respect to the direction of such effects. Importantly, both studies report that the avatar's conflicting visual perspective interfered with participants' self-perspective-based decisions regardless of the avatar's in-group or out-group status. Other group distinctions can however modulate spontaneous perspective taking as well (Ferguson et al., 2018). When adult participants completed the dot perspective task with a child avatar, no such interference was observed, indicating that the effect of social categorization on mentalizing may depend on the type of category in question.

Note that there is a relevant representational difference between L1PT and L2PT which makes it likely that group manipulations would affect them differently. Specifically, L1PT can be achieved by tracing a person's line of sight without further considerations about the agent's features or different perspectives even (Aichhorn et al., 2006). In contrast, representing *how* something appears from another person's point of view (L2PT) necessitates handling alternative representations of reality, sometimes in the form of alternative object *identities*. L2PT requires the appreciation of perspectives (Arora et al., 2015). Importantly, how an agent represents his surroundings depends not only on the properties of the physical world, but also on those of the agent (Surtees et al., 2012). If a single object has different appearances from different viewpoints, the way it is perceived by someone depends on the viewpoint from which the object is perceived in combination with the perceiver's competence to recognize those appearances. Take, for example, an object that appears yellow from one side but green from the other—this object will only give rise to different experiences depending on the viewpoint of those who have intact color vision. While humans' default expectation is

probably that others' color vision is intact, evidence about the contrary should directly influence the mental content attributed to that person (L2PT), but not whether he sees the object at all (L1PT).

Take another example where a piece of paper is lying on the table with a character (6) drawn in the middle. The character's meaning (the magnitude it symbolizes in Western cultures: six or nine) depends on the viewpoint from which we perceive it. However, it will only convey different meaning from different perspectives to those who have knowledge about the existence of both symbols (6,9). For someone who does not have that knowledge, the character will *look* different from the opposing perspectives, it will not, however, *mean* different things. Attributing the representation "nine" to the partner when I see "six" relies on the implicit assumption that the two of us have the same knowledge of the symbol. We argue that due to its truly perspectival, aspectual nature, L2PT should be especially sensitive to cues that imply that the interactional partner's knowledge differs substantially from one's own and therefore he or she may represent the world differently.

We propose that mindreading (mentalization) uses the information derived from social category memberships to adjust to the specifics of the interactional partner and this flexibility manifests itself in visual perspective taking from childhood. We claim that, in the absence of evidence on the contrary (as in most experimental contexts using avatars or live co-actors), humans expect their interactional partner to belong to their own social group and so to have a shared knowledge base, a core common ground (Kecskés & Zhang, 2009). However, cues that violate the default assumption of shared knowledge bases are likely to create *uncertainty* about how the other might understand the scene, as even the simplest visual stimuli may mean different things depending on the perceiver's background knowledge. We propose that such uncertainty should result in slower, possibly more deliberate mentalizing, preventing L2PT to unfold in a quick, spontaneous manner.

Note, that in this theoretical framework, the lack of spontaneous L2PT is not caused by perceiving a difference in group membership per se, rather obtaining evidence that the partner does not share the same broad knowledge as the child themselves. Two predictions follow from this. First, any cue to discrepancies in knowledge states should lead to a decrease in spontaneous L2PT. Second, those cues that delimit groups *without* relevant differences in knowledge states should *not* diminish spontaneous perspective taking.

To confirm that different category cues affect spontaneous L2PT selectively, we manipulated group membership in two ways. In Experiment 1, we employed a minimal group membership manipulation. We expected participants to understand the arbitrariness of this distinction and consequently, that it would not interfere with how participants attribute interpretations of number symbols to the partner. Since the meaning of number symbols are rooted in cultural knowledge—which was unaffected by the manipulation—, we hypothesized that spontaneous level-2 perspective interference would emerge in all conditions. In Experiment 2, we manipulated the accent (foreign, native) of the person with whom children completed the task, thereby informing them about a cultural discrepancy that would be relevant in processing stimuli with culturally determined meaning. We expected this to eliminate level-2 perspective interference, due to the increased uncertainty of how the partner represents stimuli. In Experiment 3, we manipulated both the minimal group membership and the accent of the partner, and tested whether shared knowledge attribution induced by accent play a primary role in modulating L2PT. We tested these questions in 8-year-old children, the youngest age group in which online L2PT has already been demonstrated (Elekes et al., 2016).

2 | EXPERIMENT 1

To capture whether spontaneous level-2 perspective taking is influenced by social cues, we chose the number verification paradigm of Surtees et al. (2012), that has already been adapted to live settings—participant dyads (Elekes et al., 2016; Surtees et al., 2016). Participants in this setting sit across each other facing a monitor that lays on a table in between them. While numbers appear on the screen one by one, they hear the name of a number too, which either corresponds to the one they see, or differs from it. Their task is to decide whether what they see and hear is the same number. The numbers are either asymmetrical, perspective dependent (6,9) or symmetrical, non-perspective

dependent (8,0). In individual trials, one of the participants is sitting with their back towards the monitor, while in joint trials, participants attend the screen simultaneously. An increase in reaction times and a decrease in hit rates to jointly attended asymmetrical trials shows that the others' level-2 perspective has been spontaneously computed.

Although both adults and children have been found to demonstrate spontaneous L2PT in this task (Elekes et al., 2016, 2017), children in the previous experiment were tested in pairs with an adult confederate (Elekes et al., 2017). We therefore conducted a pilot experiment with child-child dyads without group manipulation to confirm that spontaneous perspective taking in children is not constrained to adult partners (all relevant data is included in the Supplemental Materials). Findings of the pilot experiment showed robust L2PT interference, similarly to that reported by Elekes et al. (2017), which enabled us to test the effect of minimal group membership on L2PT in this experimental context.

Children from the age of 5 are already sensitive to minimal group membership manipulations (Dunham et al., 2011), influencing their forced choice decisions in multiple domains of prosocial behavior and social preferences (Plötner et al., 2015). Despite children's ability to register such differences, we predicted that the minimal group manipulation would not influence L2PT, as minimal group membership does not imply relevant differences in core knowledge states.

2.1 | Method

2.1.1 | Participants

Data was collected in urban, public elementary schools. Forty-eight second and third grade White, middle-class children participated in the study in pairs (29 females, 19 males, M_{age}= 8.35 years, SD = .57). Participating pairs were always classmates, and so they were familiar with each other. Children were randomly assigned to one of the 2 conditions: same minimal group membership (24 children) and different minimal group membership (24 children). Each participant had to score at least 50% of trials correctly within each stimulus-category (= 14 correct answers) in order to provide enough data points for analyses. Two children were excluded for not fitting the criteria and only one additional child performed near the criteria (this child's performance was overall above 80% but approached 50% in one stimulus-category). The ratio of correct answers within each category and within each child remained above 70% and was on average 91.89%. The final sample consisted of 46 children, evenly distributed in the two experimental conditions.

Children's parents received an informed consent form, which was handed out and collected by the teachers. Only children who were willing to participate and whose parents signed the consent took part in the experiment. The experiment was approved by the ethical committee of the university.

2.1.2 | Materials

The visual stimuli consisted of two symmetrical (0,8) and two asymmetrical numbers (6,9), which were drawn in Matlab R2013a. Visual stimuli were similar to those in Elekes et al. (2016, 2017) except that numbers were black with different patterned surfaces (lined or dotted). The audio stimuli were presented in a neutral, female voice. The length of the words was as follows: zero ("nulla," 760 ms), eight ("nyolc," 760 ms), six ("hat," 600 ms), nine ("kilenc," 760 ms).

2.1.3 | Procedure

Children in the first experiment were classmates who participated in randomly assigned pairs. The testing room was a currently unused, quiet classroom in the school. After they had been seated across each other with a 13-inch flat

monitor lying on the table in between them, children received the instruction based on the condition: identical minimal group or different minimal group.

To indicate their group membership, children were told that they were going to play a game in which they would be collecting points for the group they had been assigned to by chance. The experimenter explained that their class was divided into two groups which are competing to earn the most points during the game. After explaining the number verification task, children were told that each group would be earning points based on the group member's speed and accuracy in the task. Participating pairs were either told that they had been assigned to the same group (minimal in-groups), or that they were representatives of the two different groups (minimal out-groups). Children were additionally told that they would both have a chance to practice, and that this phase of the experiment is not about collecting points.

Children's task was to decide whether the number they see and the number they hear is the same or different. Throughout the experiment participants had to respond based on *their own* perspective, and the alternative perspective was not mentioned to them. Participants indicated their answers with separate keyboards. The buttons to press were "a" and "s" on one keyboard and "k" and "l" on the other. Key mappings were counterbalanced between subjects. Other buttons were removed from the keyboard, making it easier to press the targeted buttons. Children were first told which buttons to press when getting instructions and an additional second time during their practice phase. They were also reminded of which group they were collecting points for before starting the joint phase of the experiment.

After the task had been explained, participant "A" had a chance to practice through 16 trials, and a following 112 trials to play individually, while the other participant was sitting with her back to the monitor. This was followed by 16 practice trials for participant "B," after which participants played together for the upcoming 112 trials. Finally, participant "B" had a chance to play alone, the same way as participant "A" did. For the program to automatically switch between trials one or both children had to provide answers, depending on the phase of the experiment.

Each test phase included a 30-s-long rest-phase at half of the trials occurring automatically, but the switch between phases was directed by the experimenter, which allowed further rests if children needed them. Children completed the experiment in about 25 min including the rest phases.

Half of the trials depicted symmetric numbers and the other half consisted of asymmetric numbers. Visual stimuli were corresponding with the audio stimuli on half of the trials and non-corresponding in the other half. Correspondence was introduced following the experimental tradition to present a valid task through varying the response types. This led to 4 categories of stimuli with an equal number of trials: symmetric—corresponding (0 – "zero", 8 – "eight"); asymmetric-corresponding (6 – "six", 9 – "nine"); symmetric—non-corresponding (0 – "eight", 8 – "zero"); asymmetric—non-corresponding (6 – "nine", 9 – "six"). Participants had 28 trials for each trial type in both the individual and joint phases of the experiment.

Children heard the names of numbers through the loudspeakers of the laptop. The script that was used for stimulus presentation and response recording was written in PsychoPy 1.81. Trials started with the appearance of a fixation cross at the center of the screen, which was followed by the audio stimulus 500 ms later. The visual stimulus was presented 300 ms after the onset of the audio stimulus and remained on screen till the script received a response from one or both participants depending on condition.

2.1.4 | Data analyses and design

Before analyzing the reaction time data, outliers were removed. Outlier data points were identified individually for each participant. Data points differing by two standard deviations from the mean RT of the given participant were defined as outliers. Datasets submitted to statistical analyses are available online for all experiments: https://osf.io/bfs8p/?view_only=f2c93f08eb5a4e06afe571568c507415

Reaction time data and hit rates were included in separate analyses. Both datatypes were analyzed with repeated measures ANOVA in the statistics software JASP 0.11.1.0. A 2 × 2 × 2 × 2 (Jointness [individual, joint] × Symmetry

TABLE 1 Mean reaction times and hit rates with standard deviations in parentheses as a function of jointness, symmetry, correspondence, and condition in Experiment 1: Minimal groups.

Jointness	Symmetry	Correspondence	Condition	RT (sec) Mean (SD)	Hit rate (%) Mean (SD)
Individual	Symmetric	Corresponding	In-group	1.449 (.36)	93.17 (8.94)
			Out-group	1.438 (.39)	92.7 (8.09)
		Non-corresponding	In-group	1.507 (.35)	96.43 (4.57)
			Out-group	1.537 (.39)	93.94 (6.41)
	Asymmetric	Corresponding	In-group	1.568 (.39)	94.1 (5.56)
			Out-group	1.582 (.50)	95.19 (4.39)
		Non-corresponding	In-group	1.543 (.31)	94.72 (5.98)
			Out-group	1.622 (.53)	92.55 (7.83)
Joint	Symmetric	Corresponding	In-group	1.507 (.24)	88.67 (11.88)
			Out-group	1.488 (.50)	93.32 (8.5)
		Non-corresponding	In-group	1.612 (.29)	91.93 (6.02)
			Out-group	1.613 (.59)	91.93 (7.93)
	Asymmetric	Corresponding	In-group	1.747 (.45)	87.89 (7.12)
			Out-group	1.704 (.63)	92.39 (7.25)
		Non-corresponding	In-group	1.819 (.44)	87.89 (8.73)
			Out-group	1.766 (.69)	89.13 (9.36)

[symmetric, asymmetric] × Correspondence [corresponding, non-corresponding] × Condition [minimal in-group, minimal-out-group]) factorial design was used, in which jointness, symmetry, and correspondence were included as within-subject factors and condition provided the between-subject factor.

Our main focus was on the jointness × symmetry and on the jointness × symmetry × condition interactions. An interaction between jointness and symmetry was expected on both RT-s and hit rates. Since only perspective dependency, namely asymmetry would influence behavior, and specifically in the joint asymmetric trials, we expected increased reaction times and decreased hit rates in joint compared to individual participation for asymmetric but not symmetric numbers. This pattern of data would indicate that the other's perspective had been represented and causes interference when conflicting with the self-perspective. We did not expect this effect to vary with participants' minimal group membership, consequently, we did not predict a three-way interaction between jointness, symmetry and condition. We present the above introduced focal analyses, all the further analyses are described in the Supplemental Materials.

2.2 | Results

For descriptive statistical data on reaction times and hit rate, see Table 1.

2.2.1 | Reaction times

A significant interaction between jointness and symmetry indicated spontaneous perspective taking, since participants' responses were slower in the joint phase when responding to asymmetric (perspective dependent) numbers, *F*

(1, 44) = 11.991, $p = .001$, $\eta_p^2 = .214$. The lack of a three-way interaction between jointness, symmetry and condition was in line with our hypothesis, showing no modulatory effect of minimal group membership on spontaneous perspective interference, $F(1, 44) = 1.504$, $p = .227$, $\eta_p^2 = .033$.

2.2.2 | Hit rates

The interaction between jointness and symmetry reached significance on hit rates as well, $F(1, 44) = 4.812$, $p = .034$, $\eta_p^2 = .099$. The decrease in hit rates between symmetric and asymmetric numbers was larger in the joint phase than when children participated alone. We did not find a three-way interaction between jointness, symmetry and condition, $F(1, 44) = .037$, $p = .848$, $\eta_p^2 = .001$, further supporting the notion that calculating another person's level-2 perspective is not modulated by their minimal group membership.

Additionally, the main effect of jointness reached significance, $F(1, 44) = 19.605$, $p < .0001$, $\eta_p^2 = .308$, hit rates being lower in the joint phase. This was modulated by an interaction between jointness and condition, $F(1, 44) = 4.648$, $p = .037$, $\eta_p^2 = .096$, so that accuracy decreased more in the joint phase compared to the individual phase in the in-group condition. Comparison of the results of Experiment 1 and the Pilot Experiment contribute to the claim that minimal group manipulation does not impact L2PT. This analysis and Bayesian analyses that support this pattern of results are reported in the Supplemental material.

3 | EXPERIMENT 2

In Experiment 1, 8-year-old children demonstrated spontaneous level-2 PT regardless of the minimal group membership of their task partner. According to our account, only those social categories should alter spontaneous perspective taking that suggest a difference in knowledge bases. Minimal group membership conveys no such information and consequently was found not to modulate spontaneous perspective taking. In the second experiment, group membership was induced by the task partner's foreign or native accent. Some suggest that language is a primary social category cue (Kinzler, 2010), others argue that language is just one of the cues that indicate the borders of knowledge-systems, that is, social groups featured by shared knowledge and conventions (Oláh et al., 2014; Oláh & Király, 2019). It is consensual, however, that children can distinguish familiar from unfamiliar accents early in life (Butler et al., 2011) and accent influences children's expectations of other individuals from preschool age. Specifically, 4–5-year-old children demonstrate selective trust in information conveyed by native accented speakers (Kinzler et al., 2011) and weigh native accent more than race when forming social preferences (Kinzler et al., 2009). The significance of accent in social categorization allowed us to use this robust cue of cultural group membership and simultaneously, provided a solution to stimulus presentation: children knew that the partner spoke and understood their native language (and so comprehended the auditory stimuli used in the experiment), but could still be perceived as out-group.

3.1 | Method

3.1.1 | Participants

Data was collected in urban, public elementary schools. Forty-two, second and third grade White, middle-class children participated (23 females, 19 males, M_{age} = 8.9 years, SD = .633). Children were randomly assigned to the conditions. Half of the children participated with a cultural in-group adult confederate and the other half with a cultural out-group adult confederate ($n = 21$ in each group). Exclusion criteria were identical to that in experiment 1. No

child was excluded in this experiment. Informed consents were collected the same way as described in Experiment 1. The experiment was approved by the ethical committee of the university.

3.1.2 | Materials

Materials were identical to that used in Experiment 1.

3.1.3 | Procedure

Children in this study took part in the task with a female adult confederate in her twenties, who was introduced to children as a student. When arriving to the room, children met the confederate, and were simply told that they were going to play a game together. The confederate was either a cultural in-group or an out-group, in the latter case, she spoke the children's native language with a foreign accent. The child and the confederate were asked to shortly introduce themselves to each other, by telling their name, age, where they are from and how far they live from the school. The confederate told children that she attended language courses in the child's native language. The introduction allowed children to realize that while the confederate understands their language, she has a strong accent, indicating that she was a member of another linguistic group. The language of instructions or communication of any sort was conveyed in the children's native language. After briefly introducing the participating partners to each other, the experimenter explained the goal of the game, and participants were asked to be as quick and accurate as they can be. Children were always the first to participate, starting with a 16-trial practice followed by individual participation and the confederate's practice trials. After this they participated jointly, and finally, the confederate completed her individual trials. In all other respects the procedure was identical to that in Experiment 1.

3.1.4 | Data analysis and design

The method for removing outlier data points in the reaction time data was identical to that in Experiment 1. Like in Experiment 1, reaction time data as well as hit rates were analyzed with repeated measures ANOVA in a 2 × 2 × 2 × 2 (Jointness [individual, joint] × Symmetry [symmetric, asymmetric] × Correspondence [corresponding × non-corresponding] × Condition [cultural in-group, cultural out-group]) factorial design. In this experiment, in addition to a significant jointness × symmetry interaction we also expected to find a three-way interaction between jointness, symmetry and condition. We predicted that the partner's cultural group membership would influence spontaneous level-2 perspective taking, so that spontaneous perspective taking would be limited to the in-group condition. We introduce in the main text the focal analyses, all the further analyses are described in the Supplemental Materials.

3.2 | Results

For descriptive statistical data on reaction times and hit rate, see Table 2.

3.2.1 | Reaction times

The jointness x symmetry interaction reached significance, F (1, 40) = 5.377, p = .026, η_p^2 = .118, showing that decisions about asymmetric numbers take longer when visual stimuli are attended jointly, compared to individual

TABLE 2 Mean reaction times and hit rates with standard deviations in parentheses as a function of jointness, symmetry, correspondence, and condition in Experiment 2: Linguistic groups.

Jointness	Symmetry	Correspondence	Condition	RT (sec) Mean (SD)	Hit rate (%) Mean (SD)
Individual	Symmetric	Corresponding	In-group	1.394 (.36)	96.6 (3.48)
			Out-group	1.346 (.34)	96.26 (6.73)
		Non-corresponding	In-group	1.525 (.36)	97.28 (2.97)
			Out-group	1.404 (.3)	96.09 (3.55)
	Asymmetric	Corresponding	In-group	1.408 (.39)	96.09 (5.41)
			Out-group	1.347 (.36)	95.92 (5.45)
		Non-corresponding	In-group	1.556 (.33)	96.6 (4.15)
			Out-group	1.451 (.33)	96.09 (4.91)
Joint	Symmetric	Corresponding	In-group	1.427 (.5)	97.45 (2.8)
			Out-group	1.265 (.32)	95.24 (5.91)
		Non-corresponding	In-group	1.498 (.4)	96.77 (3.9)
			Out-group	1.330 (.28)	95.75 (4.46)
	Asymmetric	Corresponding	In-group	1.520 (.56)	93.03 (6.63)
			Out-group	1.269 (.32)	96.43 (3.39)
		Non-corresponding	In-group	1.717 (.72)	92.52 (5.64)
			Out-group	1.409 (.37)	94.39 (6.64)

performance or jointly attended symmetric numbers. Most importantly, a marginally significant interaction between jointness, symmetry and condition appeared, $F(1, 40) = 3.220$, $p = .080$, $\eta_p^2 = .074$.

A jointness × condition interaction, $F(1, 40) = 3.982$, $p = .053$, $\eta_p^2 = .091$, shows that while jointly attending stimuli with an in-group member caused children to slow down, the opposite can be observed when they participated with an out-group member.

Condition-wise analysis

Our main focus in this experiment was how the between-subjects factor, condition, would influence spontaneous level-2 perspective taking. Although the interaction between jointness, symmetry and condition (reflecting this relationship) was only marginally significant, condition emerged in meaningful two-way interactions as well, specifically with jointness (see also Supplemental Materials). These interactions provided a statistical basis to perform separate analyses on the two conditions. A 2 × 2 × 2 repeated measures ANOVA was thus conducted in both conditions separately.

In the cultural in-group condition, we found a significant two-way interaction between jointness and symmetry, replicating previous findings that perspective taking occurs spontaneously for members of the cultural in-group, $F(1, 20) = 4.713$, $p = .042$, $\eta_p^2 = .191$.

In contrast, there was no interaction between jointness and symmetry in the out-group condition, $F(1, 20) = .670$, $p = .423$, $\eta_p^2 = .032$. A Bayesian 2 (jointness) × 2 (symmetry) × 2 (correspondence) repeated measures ANOVA was conducted in the out-group condition to provide confirmatory evidence for the null. The best model included only the main effects of jointness and correspondence, $BF_{10} = 2{,}145{,}000$. The BF_{10} for the jointness × symmetry interaction was .262, providing evidence against the interaction. These findings confirm our hypothesis that perspective calculation does not occur spontaneously when the partner is a cultural out-group member.

3.2.2 | Hit rates

In line with findings on the reaction time data, the jointness × symmetry interaction reached significance on hit rates as well, indicating spontaneous level-2 PT, $F(1, 40) = 5.049$, $p = .030$, $\eta_p^2 = .112$. As expected, the three-way interaction between jointness, symmetry, and condition was found significant, $F(1, 40) = 5.530$, $p = .024$, $\eta_p^2 = .121$, which allowed us to break down our analyses by condition.

Condition-wise analysis

The analysis of the in-group condition revealed the expected interaction between jointness and symmetry, $F(1, 20) = 9.613$, $p = .006$, $\eta_p^2 = .325$. On the other hand, in the out-group condition there was no jointness x symmetry interaction, $F(1, 20) = .006$, $p = .939$, $\eta_p^2 = .0$. Bayesian analysis showed that the best model was the null model, supporting the conclusion that perspective interference was limited to the in-group condition.

4 | EXPERIMENT 3

In Experiment 1, 8-year-olds spontaneous L2PT was not affected by their partner's minimal group membership, however according to the results of Experiment 2, the confederate's cultural background (native speaker vs. partner speaking with accent) did influence it, namely children demonstrated spontaneous L2PT only if their partner previously spoke their native language without accent. The goal of Experiment 3 was twofold: first, to shed light on how the previously investigated two categories (minimal and cultural group membership) influence each other. In particular, this experiment could provide further evidence in support of the claim that cues of background knowledge modulate L2PT, while minimal group membership does not. Second, while in Experiment 1 the partner was a classmate, in Experiment 2 she was an adult whom children have never met before. Experiment 3 was also created to screen out the possible effect of this difference as well.

4.1 | Method

4.1.1 | Participants

Data was collected in urban, public elementary schools, like it was in Experiment 1 and 2. 82 s and third grade children (43 male and 39 female, $M_{age} = 8.92$, $SD = .66$) participated, but one child was excluded from the final sample due to providing less than 50% correct answers and an additional child decided to finish the study after a few trials. Children were randomly assigned to one of the following four conditions: confederate was a 1, minimal in-group and linguistic in-group ($n = 20$, $M_{age} = 8.94$, $SD = .7$), or 2, minimal in-group and linguistic out-group ($n = 20$, $M_{age} = 8.82$, $SD = .61$), or 3, minimal out-group and linguistic in-group ($n = 21$, $M_{age} = 8.89$, $SD = .66$) or 4, minimal out-group and linguistic out-group ($n = 19$, $M_{age} = 8.94$, $SD = .61$) person. Informed consents were collected the same way as described in Experiment 1. The experiment was approved by the ethical committee of the university.

4.1.2 | Materials

Materials were identical to that used in Experiments 1 and 2.

4.1.3 | Procedure

The procedure was similar to Experiment 2. Children were escorted by the experimenter to a silent room in their school where they met the confederate who was a female adult and were told that they were going to play a game together. Before starting the game, the child and the confederate were told to introduce themselves to each other in a few sentences, the same way as in Experiment 2 (see Section 3.1.3). The confederate spoke in children's native language with (linguistic out-group) or without accent (linguistic in-group). After the short introduction the experimenter explained the rules of the game and also mentioned further details regarding the minimal group membership. The child and the confederate were either told that they were collecting points for the same group (minimal in-group) or for two different groups (minimal out-group). Further events of procedure (children were asked to be quick and accurate, children were always the first to participate…etc.) reflected that of Experiment 2.

4.1.4 | Data analysis and design

The data cleaning process was the same as it was in Experiments 1 and 2. Like in the previous experiments, a repeated measures ANOVA—with Jointness (joint-individual), Symmetry (symmetric- asymmetric), Correspondence (corresponding—non-corresponding) and Condition 1 (Linguistic in-group vs. out-group), and Condition 2 (minimal ingroup vs. outgroup) as factors was conducted on all measures. The main focus was on (1) the main effects of each condition, since we were interested in the difference between the effect of the two different types of grouping conditions (minimal and linguistic /cultural group membership), whether they affect the performance of participants overall or not. In addition, we concentrated on the (2) symmetry × jointness interaction, with special focus on the three-way interactions between jointness × symmetry × each of the two conditions. We were specifically interested in whether they modulate spontaneous L2PT—an interference in the form of slower reaction times, and/or higher error rates—and predicted that the jointness and symmetry interaction, the above interference effect, would be limited to linguistic (cultural) group membership. We present here the above introduced focal analyses, all the further analyses are described in the Supplemental Materials.

4.2 | Results

For the descriptive statistical data, see Table 3.

4.2.1 | Reaction times

The jointness × symmetry interaction was found to be significant, $F\,(1, 76) = 7.419$, $p = .008$, $\eta_p{}^2 = .002$. The jointness x symmetry × language condition × minimal group condition did not reach significance, $F\,(3, 76) = .126$, $p = .724$, $\eta_p{}^2 = 4.490\text{e-}4$ and none of the group conditions interacted separately with jointness × symmetry. The main between subject effect of language condition was also found to be significant, $F\,(1, 76) = 5.962$, $p = .017$, $\eta_p{}^2 = .052$, but not the main effect of minimal group, $F\,(1, 76) = .334$, $p = .565$. Children were in general slower in the presence of a linguistic in-group confederate in contrast to a confederate speaking with an accent. An additional two-way, symmetry × correspondence interaction approached significance, $F\,(1, 76) = 27.508$, $p < .001$, $\eta_p{}^2 = .005$.

TABLE 3 Mean reaction times and hit rates with standard deviations in parentheses as a function of Jointness, Symmetry, and Condition in Experiment 3:Contrasting linguistic groups with minimal groups.

Jointness	Symmetry	Language	Minimal group Condition	RT (sec) Mean (SD)	Hit rate (%) Mean (SD)
Individual	Asymmetric	Out-group	In-group	1.154 (.24)	96.311 (4.05)
			Out-group	1.113 (.31)	93.580 (5.77)
		In-group	In-group	1.256 (.23)	94.672 (8.50)
			Out-group	1.288 (.39)	94.895 (3.45)
	Symmetric	Out-group	In-group	1.109 (.22)	95.343 (4.41)
			Out-group	1.083 (.29)	92.752 (5.56)
		In-group	In-group	1.243 (.21)	93.864 (6.96)
			Out-group	1.247 (.28)	96.101 (3.15)
Joint	Asymmetric	out-group	In-group	1.141 (0.31)	91.012 (5.77)
			Out-group	1.057 (.29)	93.707 (5.11)
		In-group	In-group	1.269 (.24)	91.452 (7.68)
			Out-group	1.221 (.30)	93.612 (4.90)
	Symmetric	Out-group	In-group	1.049 (.25)	91.847 (7.83)
			Out-group	1.018 (.28)	93.170 (5.12)
		In-group	In-group	1.179 (.22)	93.834 (6.21)
			Out-group	1.115 (.23)	91.853 (8.01)

Condition-wise analysis

The main question was to explore how the between subject factors would influence L2PT. Since we were interested in the overall effect of the different group conditions (and the main effect of language condition provided statistical basis for this as well), we conducted a 2 × 2 × 2 repeated measures ANOVA on the two language conditions separately.

In the linguistic in-group condition, the main effect of jointness, $F(1, 39) = 4.186$. $p = .048$, $\eta_p^2 = .012$, symmetry F (1, 39) = 10.722, $p = .002$, $\eta_p^2 = .012$, and correspondence $F(1, 39) = 64.649$, $p < .001$, $\eta_p^2 = .056$, reached significance. Crucially, the jointness × symmetry interaction was significant, ($F(1, 39) = 5.720$, $p = .022$, $\eta_p^2 = .004$), participants were slower in responding to the asymmetric numbers specifically in the joint situation. The 3-way interaction of jointness × symmetry × minimal group did not reach significance, $F(1, 39) = .039$, $p = .844$. In addition, also the two-way interaction of symmetry × correspondence was found to be significant: $F(1, 39) = 4.676$, $p = .037$, $\eta_p^2 = .001$.

In the linguistic out-group condition, neither the jointness and symmetry, nor the 3-way interaction of jointness × symmetry × minimal group reached significance ($F(1, 39) = .840$, $p = .365$).

As the results indicate, the expected and most important jointness × symmetry interaction was significant only in the linguistic or cultural in-group condition.

4.2.2 | Hit rates

The interaction of jointness × symmetry × minimal group, $F(1, 76) = 6.123$, $p = .016$, $\eta_p^2 = .004$ reached significance. The between subject factor main effects of linguistic, $F(1, 76) = .094$, $p = .760$, and minimal group, $F(1, 76) = .025$, $p = .874$, were not significant. The jointness × symmetry × minimal group interaction was significant which allows us to analyze the data in minimal in- and out-group separately.

TABLE 4 The emergence of the jointness × symmetry interaction in the different conditions of Experiment 3.

	Reaction times	Hit rates
Linguistic Ingroup		
Cooperative	✓	✓+
Competitive	X	✓
Linguistic Outgroup		
Cooperative	X	X
Competitive	X	X

Condition-wise analysis

We split the sample by minimal group and conducted a 2 × 2 × 2 repeated measures ANOVA separately.

In the minimal in-group condition, the jointness × symmetry interaction reached significance, F (1, 38) = 4.260, $p = .046$, $\eta_p^2 = .006$.

In the minimal out-group condition there was no interaction that approached significance.

4.2.3 | Analyses of each group

Since the analyses of reaction times and hit rates yielded ambiguous results, in order to better understand the effects of the group manipulations, we decided to analyze the potential interference effect (jointness × symmetry interaction) of perspective taking separately in each group by conducting a repeated measures ANOVA with Jointness (individual—joint), Symmetry (symmetric—asymmetric), and Correspondence (corresponding—non corresponding) as factors on each measures. Since we were in search for the two-way jointness × symmetry interaction, we report only these analyses in the following section.

Group tested with same language speaker-cooperative partner

Reaction Times: In this group condition, the interaction of jointness × symmetry reached significance F (1, 19) = 6.055, $p = .024$, $\eta_p^2 = .015$. ***Hit Rates***: The analyses revealed that the interaction of jointness × symmetry F (1, 19) = 3.985, $p = .060$, $\eta_p^2 = .024$ was almost significant, F (1, 19) = 3.985, $p = .060$, $\eta_p^2 = .024$.

Group tested with same language speaker-competitive partner

Reaction Times: In this group condition no interaction turned out to be significant. ***Hit Rates***: As a result of the analysis, the interaction of jointness × symmetry reached significance F (1, 20) = 4.985, $p = .037$, $\eta_p^2 = .014$.

Group tested with other language speaker-cooperative partner

Reaction Times: In this group condition, the interaction of jointness × symmetry did not reach significance. ***Hit Rates***: In this analysis, the interaction of jointness × symmetry did not reach significance.

Group tested with other language speaker-competitive partner

Reaction Times: In this group condition, the interaction of jointness × symmetry did not reach significance. ***Hit Rates***: In this group condition, the interaction of jointness × symmetry did not reach significance.

This set of results (summarized in Table 4) draw attention to that linguistic group membership is more likely to alter the perspective taking interference than minimal group membership, since in the linguistic ingroup conditions the effect of jointness and symmetry interaction was found to be significant on either or both the reaction time and hit rate measures, while this pattern was lacking as a function of minimal group manipulation.

5 | DISCUSSION

In a set of three experiments, we aimed to investigate how social categorization influences spontaneous level-2 perspective taking (LPT2), that is, computing how an interactional partner represents certain visually presented information. The representations that people form (even of such simple visual stimuli as digits) are never objective reflections of the environment, rather they are interpretations of it. Presuming that the process of L2PT forms predictions about a *specific person's* perspective, we hypothesized that it would be sensitive to those characteristics of the individual which may influence the mental states they form. Social categories are relevant in this respect as they carry information about the knowledge states of their members. We predicted that perspective taking can unfold spontaneously only when the assumption of shared knowledge is intact, that is, children have reason to assume that given the same viewpoint, and the partner would represent the object in the same way as they themselves do. When information is available through social categorization to indicate a difference in cultural background, how the partner interprets the visual scene with culturally determined content becomes uncertain, which uncertainty disrupts the spontaneity of PT. Importantly, we expected that a difference in social categories would only modulate spontaneous PT if the category in question can be assumed to represent a meaningful distinction between groups in terms of background knowledge rather than ad hoc established minimal groups.

In the first study, we assigned children to minimal groups and motivated our 8-year-old participants to collect points for the group they have randomly been assigned to. We predicted that minimal group membership would not influence the spontaneous computation of each other's level-2 perspective. In line with our predictions, we found L2PT to occur without explicit prompting (i.e., task instructions) and irrespective of which minimal group the task partner belonged to. Our findings are in line with results based on spontaneous level-1 perspective taking (Simpson & Todd, 2017) which also show no relation between someone's minimal group membership and the spontaneous representation of their perspective.

It is important to highlight that our participants were classmates. The fact that children shared membership of a salient group, their class, may have diminished the relevance of their minimal group membership. Although this possibility may not be ruled out, it is worth noting that in the specific testing context, great emphasis was put on making children's minimal groups as salient and important as possible. Furthermore, in the seminal study of Tajfel et al. (1971), participants were also classmates and minimal group effects were robustly demonstrated.

In the second study children participated with a young adult confederate, whose cultural group membership was created by linguistic cues. In the in-group condition, children participated with a native speaker, whereas the outgroup confederate spoke children's native language with a foreign accent. Such subtle linguistic cues have been shown to alter preschoolers' social preferences (Kinzler et al., 2009) and beliefs about the applicability of the speaker's knowledge to their own behavior (Kinzler et al., 2011). Language often marks the boundaries of long-standing social groups with group-specific knowledge and habits, and from 5 years of age, children expect the members of linguistic groups to share knowledge of cultural information (Soley & Aldan, 2018). In line with these findings and also with our predictions, language based social categorization altered spontaneous L2PT. Consistent with previous works, spontaneous level-2 perspective interference emerged when children participated with a linguistic (cultural) in-group person, whereas, such an interference effect was not detectable when participating with a foreign accented speaker. This pattern was observable in both hit rates and reaction times. The findings speak for the possibility that L2PT in children is sensitive to the partner's linguistic group membership, and it seems to be reliant on children's expectation of having a common ground (shared representational space, Oláh et al., 2019) with their interacting partners.

The motivation of Experiment 3 was to manipulate both category formation induced by linguistic cues (accented vs. native language use) and by minimal cues (cooperation vs. competition) within the same study, in a contrastive design. This design allowed us to find out more about (a) the robustness of the effect and (b) the relationship of these category cues on visual perspective taking. First of all, the results of Experiment 3 revealed that with a young adult confederate—irrespective of the manipulation of the group membership of the partner—children responded slower

when the partner held a competing perspective. This result replicates that 8-year-old children show spontaneous L2PT (Elekes et al., 2017). Importantly, the result that children were in general faster when acting together with a linguistic out-group partner, in contrast to a linguistic in-group partner allows us to consider that children might pay more attention in general to social partners who are more familiar to them in some respect. Indeed, this familiarity itself might stem from the likelihood that children have better access to their knowledge. The question remained whether this supposed access to the knowledge base alters perspective taking specifically. Follow-up analyses of the third experiment revealed that only when paired with linguistic in-group partners in a cooperative setting did children manifest the interference effect of L2PT—children were slower in their response when they perceived the contrast in the knowledge dependent perspective of themselves and their partner. These results, overall, are in line with the results of the previous studies.

However, children made fewer errors when paired with a minimal out-group competitive partner, in contrast to a cooperative one. This partial result is in line with previous research that has demonstrated that the minimal group manipulation exerts its effect even if implemented *within* real-life groups of children (Hartstone & Augoustinos, 1995). In itself it is interesting that in contrast to the results of Experiment 1 and Experiment 2, in Experiment 3 the two measures, reaction time and hit rates revealed a slightly different pattern of results. This alteration might be caused by that characteristic of Experiment 3 that we manipulated accent and cooperation/competition at the same time, and the motivation to succeed with a newly encountered adult confederate (and not with a well-known peer partner) could have led children to concentrate on the accuracy of their responses resulting in an overall slowing down when making decisions together with a partner (reflected in an overall effect of jointness on reaction times specifically).

The question, whether and to what extent social categorization influences mindreading processes is relatively new in cognitive science, and our findings contribute to this field with novel insights. First, we present evidence that social categorization can alter level-2 perspective taking. Prior research has focused on whether group membership modulates performance in a level-1 context, where participants judge visual access to objects rather than their appearance (Schneider et al., 2018; Simpson & Todd, 2017). These studies report that the avatar's group membership influences the ease with which participants can free their decisions of their *own privileged access* to stimuli when performing explicit judgments about the avatar's perspective. They do not, however, find that group membership would modulate how the partner's perspective interferes with responding from one's own. We hypothesize that this discrepancy is caused by the structural difference between level-1 and level-2 tasks. While tracking visual access requires only a line-of-sight decision without necessarily appreciating different modes of presentation, representing object identity *as seen by someone else* calls for just that (Aichhorn et al., 2006). Consequently, L2PT demonstrates more sensitivity to the features of the agent's mind whose perspective is the target of mindreading.

It has been suggested that L2PT is not automatic but rather a spontaneous process (Elekes et al., 2016), which means that although overt goals are not necessary for it to happen, it is dependent on the perceiver's covert, more general goals (see also Carruthers, 2017). For example, Elekes et al. (2016) have found that L2PT did not take place if the social partner's task was irrelevant to the participant's (they were instructed to pay attention to a different feature of the numbers). Thus, the group membership of the social partner may influence the context and through it, the covert goals of the participants. As a result, participants—not necessarily consciously—first might judge the social partner's perspective to be either not relevant or possible to compute, and start to disregard the other person's representation of the scene.

Our results also demonstrate that it is not simply the partner's in-group or out-group status that the mind reading system registers (and adjusts its functioning to) when receiving input from social categorization processes. The variation in L2PT was found to be a function of linguistic category cues, that is, accent. This suggests that children from an early age associate different relevance with minimal and linguistic cues, potentially interpreting the latter as an indicator of groups that are characterized by shared knowledge. Nonetheless, the results that children made fewer errors when acting together with a competitive young adult confederate and more errors with a cooperative one raises the possibility that our minimal group manipulation—the purpose of collaboration- also made children pay attention to the perspective of their partner.

This result could also be interpreted in terms of children's tendency to monitor what the partner can perceive when it is relevant –, since for collaboration, sharing common ground and common knowledge is important (Vesper et al., 2017).

However, linguistic and minimal groups differ not solely in the inferences they allow regarding shared knowledge. An alternative explanation to the minimal group—linguistic group difference could be that the type of social category influenced participants' implicit social motivation to understand their partner. It could be argued that the more meaningful the group-difference, the less likely participants are to want to engage in interaction, therefore, the less relevant the partner's mental states become, causing a lack of spontaneous PT in the linguistic out-group condition. However, minimal groups have been demonstrated to create in-group favoritism, out-group derogation, and more perceived similarity within groups (Dunham et al., 2011; Tajfel et al., 1971; Turner et al., 1979)—that is, minimal groups can also influence the willingness for and nature of social interactions. We have no reason to assume that our group manipulations differed in this respect. Rather, we argue that the relevant difference between group manipulations was whether it impacted the assumption of shared knowledge.

Since language-based and minimal groups differ in relevance—as we have argued -, the question arises whether the two distinctions also differ in terms of memorability. Specifically, is it easier to remember linguistic than minimal group membership? We cannot rule out this possibility; however, the fact that similar manipulations of minimal group membership have produced measurable differences in behavior in previous studies (see above) suggests that children would remember this distinction in our study as well. In our experiments, children received the same amount of information that could be used to create language based and minimal category representations and we avoided using any salient visual cues that could provide further guidance. Nonetheless, the fact that linguistic groups may be seen as more relevant could inherently affect mnemonics related to the group manipulation.

On a related note, one might ask why the age difference between participating children and the adult confederate did not eliminate spontaneous level-2 perspective taking in Experiment 2 in the linguistic in-group condition as well. The confederate was substantially older than participating children, which might have created a child—adult group difference. However, being an older member of the child's linguistic (cultural) ingroup, the adult confederate likely did not elicit assumptions of difference in knowledge and consequent epistemic uncertainty—leaving L2PT intact. It is also possible that being already school-aged children, our participants did not assume that the age difference signified substantial difference in knowledge states. The same numerical age difference may, however, yield an effect in younger children who are not yet experts in their culture. For instance, Seehagen et al. (2018) have shown that 4-year-old children are more likely to succeed in attributing false-beliefs to fellow children than to adults, possibly due to a tendency to overattribute privileged knowledge to adults.

Our findings can be embedded in a broader literature showing that preschoolers differentiate between in-group and out-group agents in natural speech, using more mentalistic terms when describing ingroup individuals. McLoughlin and Over (2017) report that 5–6-year-old children are more likely to spontaneously attribute mental states to their own group members along dimensions of gender and nationality. Furthermore, children's explicit mindreading performance is more accurate for linguistic (cultural) in-group members compared to out-groups (Gönültaş et al., 2020). Similar results are reported from studies with adults, whose threshold to attribute minds to others varies with group membership (Hackel et al., 2014; Mcclung & Reicher, 2018), and their mental state inferences are slower and less accurate for cross-cultural targets than for within-culture ones (Perez-Zapata et al., 2016).

To summarize, our findings argue for the context sensitivity of spontaneous level-2 perspective taking in children. The interactional partner's social group membership was found to affect mindreading. This sort of sensitivity appears to be more complex than hypothesized before, as results diverge based on the specific manipulation used to create in-group and out-group membership. On this basis, we suggest that the mind reading system is sensitive to the knowledge states of others (inferred from group membership), and functions differently based on contextual factors that are informative in this sense. Altogether, our results suggest that there is a nuanced interplay between social categorization and mentalization.

ACKNOWLEDGMENTS

This study was supported by the Momentum Program of the Hungarian Academy of Sciences to I. Király (LP-2017-17/2017). We thank children participating in this study and Rebeka Zsoldos for her help with data collection.

CONFLICT OF INTEREST STATEMENT

The authors are in full consensus on the content of the manuscript and the ordering of authors. All co-authors contributed to the present manuscript substantially, and will be informed about any decisions regarding the manuscript by the corresponding author. There are no conflicts of interests.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in OSF at: https://osf.io/bfs8p/?view_only=f2c93f08eb5a4e06afe571568c507415

ETHICS APPROVAL

All the experiments presented in the paper were approved by the ethical committee of the University (Ethical Committee of the Eötvös Loránd University, Budapest, approval No.: 2018/126-2).

ORCID

Ildikó Király https://orcid.org/0000-0002-4857-1764
Katalin Oláh https://orcid.org/0000-0001-9570-3558

REFERENCES

Aichhorn, M., Perner, J., Kronbichler, M., Staffen, W., & Ladurner, G. (2006). Do visual perspective tasks need theory of mind? *NeuroImage*, *30*, 1059–1068. https://doi.org/10.1016/j.neuroimage.2005.10.026

Arora, A., Weiss, B., Schurz, M., Aichhorn, M., Wieshofer, R. C., & Perner, J. (2015). Left inferior-parietal lobe activity in perspective tasks: Identity statements. *Frontiers in Human Neuroscience*, *9*, 360. https://doi.org/10.3389/fnhum.2015.00360

Butler, J., Floccia, C., Goslin, J., & Panneton, R. (2011). Infants' discrimination of familiar and unfamiliar accents in speech. *Infancy*, *16*(4), 392–417. https://doi.org/10.1111/j.1532-7078.2010.00050.x

Carruthers, P. (2017). Mindreading in adults: Evaluating two-systems views. *Synthese*, *194*(3), 673–688. https://doi.org/10.1007/s11229-015-0792-3

Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of "minimal" group affiliations in children. *Child Development*, *82*(3), 793–811. https://doi.org/10.1111/j.1467-8624.2011.01577.x

Elekes, F., Varga, M., & Király, I. (2017). Level-2 perspectives computed quickly and spontaneously: Evidence from eight-to 9.5-year-old children. *British Journal of Developmental Psychology*, *35*(4), 609–622. https://doi.org/10.1111/bjdp.12201

Elekes, F., Varga, M., & Király, I. (2016). Evidence for spontaneous level-2 perspective taking in adults. *Consciousness and Cognition*, *41*, 93–103. https://doi.org/10.1016/j.concog.2016.02.010

Ferguson, H. J., Brunsdon, V. E., & Bradford, E. E. (2018). Age of avatar modulates the altercentric bias in a visual perspective-taking task: ERP and behavioral evidence. *Cognitive, Affective, & Behavioral Neuroscience*, *18*(6), 1298–1319. https://doi.org/10.3758/s13415-018-0641-1

Flavell, J. H. (1999). Cognitive development: children's knowledge about the mind. *Annual Review of Psychology*, *50*, 21–45. https://doi.org/10.1146/annurev.psych.50.1.21

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*(1), 99. https://doi.org/10.1037/0012-1649.17.1.99

Freundlieb, M., Kovács, Á. M., & Sebanz, N. (2018). Reading your mind while you are reading—Evidence for spontaneous visuospatial perspective taking during a semantic categorization task. *Psychological Science*, *29*(4), 614–622. https://doi.org/10.1177/0956797617740973

Frith, C., & Frith, U. (2005). Theory of mind. *Current Biology*, *15*(17), R644–R645. https://doi.org/10.1016/j.cub.2005.08.041

Gönültaş, S., Selçuk, B., Slaughter, V., Hunter, J. A., & Ruffman, T. (2020). The capricious nature of theory of mind: Does mental state understanding depend on the characteristics of the target? *Child Development*, *91*, e280–e298. https://doi.org/10.1111/cdev.13223

Gzesh, S. M., & Surber, C. F. (1985). Visual perspective-taking skills in children. *Child Development*, *56*, 1204–1213. https://doi.org/10.2307/1130235

Hackel, L. M., Looser, C. E., & Van Bavel, J. J. (2014). Group membership alters the threshold for mind perception: The role of social identity, collective identification, and intergroup threat. *Journal of Experimental Social Psychology*, *52*, 15–23. https://doi.org/10.1016/j.jesp.2013.12.001

Hartstone, M., & Augoustinos, M. (1995). The minimal group paradigm: Categorization into two versus three groups. *European Journal of Social Psychology*, *25*(2), 179–193. https://doi.org/10.1002/ejsp.2420250205

Kecskés, I., & Zhang, F. (2009). Activating, seeking, and creating common ground: A socio-cognitive approach. *Pragmatics & Cognition*, *17*(2), 331–355. https://doi.org/10.1075/p&c.17.2.06kec

Kinzler, K. D. (2010). Priorities in social categories. *European Journal of Social Psychology*, *40*, 581–592. https://doi.org/10.1002/ejsp.739

Kinzler, K. D., Corriveau, K. H., & Harris, P. L. (2011). Children's selective trust in native-accented speakers. *Developmental Science*, *14*(1), 106–111. https://doi.org/10.1111/j.1467-7687.2010.00965.x

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2012). 'Native'objects and collaborators: Infants' object choices and acts of giving reflect favor for native over foreign speakers. *Journal of Cognition and Development*, *13*(1), 67–81. https://doi.org/10.1080/15248372.2011.567200

Kinzler, K. D., Shutts, K., DeJesus, J., & Spelke, E. S. (2009). Accent trumps race in guiding children's social preferences. *Social Cognition*, *27*(4), 623–634. https://doi.org/10.1521/soco.2009.27.4.623

Luo, Y., & Baillargeon, R. (2007). Do 12.5-month-old infants consider what objects others can see when interpreting their actions? *Cognition*, *105*(3), 489–512. https://doi.org/10.1016/j.cognition.2006.10.007

Luo, Y., & Beck, W. (2010). Do you see what I see? Infants' reasoning about others' incomplete perceptions. *Developmental Science*, *13*(1), 134–142. https://doi.org/10.1111/j.1467-7687.2009.00863.x

McClung, J. S., & Reicher, S. D. (2018). Representing other minds: Mental state reference is moderated by group membership. *Journal of Experimental Social Psychology*, *76*, 385–392. https://doi.org/10.1016/j.jesp.2017.11.013

McLoughlin, N., & Over, H. (2017). Young children are more likely to spontaneously attribute mental states to members of their own group. *Psychological Science*, *28*(10), 1503–1509. https://doi.org/10.1177/0956797617710724

Moll, H., & Meltzoff, A. N. (2011). How does it look? Level 2 perspective-taking at 36 months of age. *Child Development*, *82*(2), 661–673. https://doi.org/10.1111/j.1467-8624.2010.01571.x

Moll, H., Meltzoff, A. N., Merzsch, K., & Tomasello, M. (2013). Taking versus confronting visual perspectives in preschool children. *Developmental Psychology*, *49*(4), 646. https://doi.org/10.1037/a0028633

Moll, H., & Tomasello, M. (2006). Level 1 perspective-taking at 24 months of age. *British Journal of Developmental Psychology*, *24*(3), 603–613. https://doi.org/10.1348/026151005X55370

Oláh, K., Elekes, F., Bródy, G., & Király, I. (2014). Social category formation is induced by cues of sharing knowledge in young children. *PloS one*, *9*(7), e101680. https://doi.org/10.1371/journal.pone.0101680

Oláh, K., Elekes, F., & Király, I. (2019). Creating a shared representational space. PsyArXiv. https://psyarxiv.com/g2z74/

Oláh, K., & Király, I. (2019). Young children selectively imitate models conforming to social norms. *Frontiers in Psychology*, *10*, 1399. https://doi.org/10.3389/fpsyg.2019.01399

Perez-Zapata, D., Slaughter, V., & Henry, J. D. (2016). Cultural effects on mindreading. *Cognition*, *146*, 410–414. https://doi.org/10.1016/j.cognition.2015.10.018

Pető, R., Elekes, F., Oláh, K., & Király, I. (2018). Learning how to use a tool: Mutually exclusive tool–function mappings are selectively acquired from linguistic in-group models. *Journal of Experimental Child Psychology*, *171*, 99–112. https://doi.org/10.1016/j.jecp.2018.02.007

Plötner, M., Over, H., Carpenter, M., & Tomasello, M. (2015). The effects of collaboration and minimal-group membership on children's prosocial behavior, liking, affiliation, and trust. *Journal of Experimental Child Psychology*, *139*, 161–173. https://doi.org/10.1016/j.jecp.2015.05.008

Powell, L. J., & Spelke, E. S. (2013). Preverbal infants expect members of social groups to act alike. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(41), E3965–E3972. https://doi.org/10.1073/pnas.1304326110

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(5), 1255–1266. https://doi.org/10.1037/a0018729

Schneider, D., Grigutsch, A., Schurz, M., Zäske, R., & Schweinberger, S. R. (2018). Group membership and the effects on visual perspective taking. Psyarchive. 10.31234/osf.io/wnrk6

Seehagen, S., Dreier, L., & Zmyj, N. (2018). Overrated adults: 4-year-olds' false belief understanding is influenced by the believer's age. *Journal of Experimental Child Psychology*, *167*, 328–335. https://doi.org/10.1016/j.jecp.2017.11.007

Simpson, A. J., & Todd, A. R. (2017). Intergroup visual perspective-taking: Shared group membership impairs self-perspective inhibition but may facilitate perspective calculation.*Cognition*, *166*, 371–381. https://doi.org/10.1016/j.cognition.2017.06.003

Soley, G., & Aldan, P. (2018). Children and adults selectively attribute shared cultural knowledge to speakers of the same language. *Child Development*, *91*, https://doi.org/10.1111/cdev.13161

Soley, G., & Spelke, E. S. (2016). Shared cultural knowledge: Effects of music on young children's social preferences. *Cognition*, *148*, 106–116. https://doi.org/10.1016/j.cognition.2015.09.017

Spaulding, S. (2018). Do you see what I see? How social differences influence mindreading. *Synthese*, *195*(9), 4009–4030. https://doi.org/10.1007/s11229-017-1404-1

Surtees, A. D., & Apperly, I. A. (2012). Egocentrism and automatic perspective taking in children and adults. *Child Development*, *83*(2), 452–460. https://doi.org/10.1111/j.1467-8624.2011.01730.x

Surtees, A., Apperly, I., & Samson, D. (2016). I've got your number: Spontaneous perspective-taking in an interactive task. *Cognition*, *150*, 43–52. https://doi.org/10.1016/j.cognition.2016.01.014

Surtees, A. D., Butterfill, S. A., & Apperly, I. A. (2012). Direct and indirect measures of level-2 perspective-taking in children and adults. *British Journal of Developmental Psychology*, *30*(1), 75–86. https://doi.org/10.1111/j.2044-835X.2011.02063.x

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, *1*(2), 149–178. https://doi.org/10.1002/ejsp.2420010202

Todd, A. R., Hanko, K., Galinsky, A. D., & Mussweiler, T. (2011). When focusing on differences leads to similar perspectives. *Psychological Science*, *22*(1), 134–141. https://doi.org/10.1177/0956797610392929

Turner, J. C., Brown, R. J., & Tajfel, H. (1979). Social comparison and group interest in ingroup favouritism. *European Journal of Social Psychology*, *9*(2), 187–204. https://doi.org/10.1002/ejsp.2420090207

Vesper, C., Abramova, E., Bütepage, J., Ciardo, F., Crossey, B., Effenberg, A., Hristova, D., Karlinsky, A., McEllin, L., Nijssen, S. R. R, Schmitz, L., & Wahn, B. (2017). Joint action: Mental representations, shared information and general mechanisms for coordinating with others. *Frontiers in Psychology*, *07*, 2039. https://doi.org/10.3389/fpsyg.2016.02039

Westra, E. (2017). Stereotypes, theory of mind, and the action–prediction hierarchy. *Synthese*, *196*(7), 2821–2846. https://doi.org/10.1007/s11229-017-1575-9

Yuan, X., Wang, N., Geng, H., & Zhang, S. (2017). Mentalizing another's visual world—A novel exploration via motion aftereffect. *Frontiers in Psychology*, *8*, 1535. https://doi.org/10.3389/fpsyg.2017.01535

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Elek, L. P., Király, I., Pető, R., Szücs, R., Elekes, F., & Oláh, K. (2023). Linguistic but not minimal group membership modulates spontaneous level-2 perspective interference in 8-year-old children. *Social Development*, 1–20. https://doi.org/10.1111/sode.12719