Title

Number In Perspective: Why is it Hard for Preschoolers to Attribute False Belief About Numerosity?

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors
Jakab, Zoltan
Kiss, Szabolcs

Publication Date
2024

Peer reviewed

Number In Perspective: Why is it Hard for Preschoolers to Attribute False Belief About Numerosity?

Zoltan Jakab (zoltan.jakab@barczi.elte.hu)
Institute for the Psychology of Special Needs, Eötvös Loránd University of Sciences
1097 Budapest, Ecseri út 3, Hungary

Szabolcs Kiss (kiss.szabolcs@pte.hu)
Institute of Psychology, University of Pécs
7624 Pécs, Ifjúság útja 6. Hungary

Abstract

In this paper we report an investigation of how concepts of integer number combine with those of mindreading. We used tasks that require explicit thought and verbal responses, and examined children between 6-10 years of age. We designed four experiments to look at the intersection of quantification and mindreading in development using two combination tasks: (i) visual perspective taking and number; (ii) false belief and number. In both, children needed to coordinate between simple mathematical operations (counting and addition), and reconstructing an agent's visual or mental perspective. Although all preschoolers were proficient in counting, and the majority of them passed the false-belief task, the false belief and number task proved surprisingly difficult, and was not mastered before age 8. After briefly discussing theories of concept combination, we offer a performance-based explanation of this difficulty.

Keywords: false belief attribution, integer numbers, concept combination, perspective taking

Number and Mindreading at Preschool Age

Most typically developing 5 to 6 year old children pass standard verbal false-belief attribution tasks (Wellman et al., 2001; Perner & Ruffman, 2005; Apperly & Butterfill, 2009; Carruthers, 2016), manage simpler and more complex forms of visual perspective taking (i.e., what someone else can see and how someone represents a jointly viewed object; Flavell et al., 1981; Gopnik & Astington, 1988; Moll & Meltzoff, 2011; Sodian, Thoermer, & Metz, 2007; Samson et al., 2010; Surtees & Apperly, 2012; Qureshi, Apperly, & Samson, 2010), and can also count (Gelman & Gallistel, 1978; Gallistel & Gelman, 1992, 2000; Wynn, 1990, 1992, 1998; LeCorre et al. 2006; LeCorre & Carey, 2007; Carey, 2009, Ch 8). An interesting question is, at what age do children manage to combine their number concepts with those of false belief (mental perspective) and perceptual perspective? The topic of concept combination in cognitive psychology and cognitive science has seen a growing number of studies in recent years (Murphy, 2002, Ch12; Hinzen et al., 2012; Hampton & Winter, 2017). Fodor (1998) famously argued that most existing psychological theories of concepts cannot account for combination (Fodor, 1998, p114; for criticism, see Laurence & Margolis, 1999, 67-8; Murphy (2002, Ch 12). From a developmental perspective, it must be evident that young children can efficiently combine the concepts they possess. For example, a four-year-old child passing the Sally-Ann task evidently combines concepts of physical objects, and those of mental states (e.g., when entertaining the idea that Sally believes that her ball is in the basket). The question is exactly how – by means of what cognitive mechanism – children manage to combine their concepts. Here we discuss two accounts of concept combination, and based on both we make predictions about preschoolers' performance in number-perspective combination tasks.

Theories of Concept Combination

According to Barsalou (2017), concept combination is a process called *multimodal simulation*, often supported by *extensional feedback*. On this view, concepts are represented by distributed multimodal networks in the brain. These networks can reproduce (recall), or creatively produce (i.e., simulate) complex experiential states (or 'neural and bodily states' e.g., Barsalou, 2017, p17) pertaining to interactions with members of the category in question. A straightforward case is BICYCLE, but the same principle is applied to highly abstract concepts like TRUTH (op.cit. p13) where a starting point for many people is how truth is understood in courts of law. This view entails that conceptual content has a situation-specific character, which also affects concept combination. For example, in tokening the concept LAWN, people are likely to simulate green blades of grass whereas in tokening ROLLED-UP LAWN (a complex concept) they are much more likely to activate multimodal representations of dirt and roots (as evidenced by reports of experimental subjects). This illustrates that concept combination tends to "reveal normally occluded features" (p18). Extensional feedback is the idea that the environment can contribute to concept combination. PET FISH may be constructed entirely from the two constituent categories in cultures that do not keep fish as pets; in cultures that do, features not obvious from the constituent concepts are supplied by the environment (i.e., the referent or extension of the complex concept). We think this theory predicts that preschoolers should be good at concept combination including number and perspective, simply because they are good at handling multimodal, episodic information (Piaget, 1962; Harris, 2000, 2022), and this allows them to reconstruct in imagination an ignorant agent's perspective. We will return to this in the discussion.

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

The theory of mental files (Recanati, 2012, 2016) also has resources to account for combination. In psychology, Perner et al. (2015) present a detailed account of how metarepresentations and ordinary representations combine with one another (see also Huemer et al., 2018). According to Perner et al. (2015), preschoolers can form *regular mental files* (vehicles of first-order mental representation) and *vicarious files* (to meta-represent other agents' mental states), but tend to make certain mistakes in handling them. Children overcome these difficulties when they understand second-order false belief. In philosophy of language, Pryor (2016) outlines an extended version of mental file theory. He argues that a complete theory of mental files will have to include a model of the structure of complex thoughts (propositional attitudes). Such an account plausibly takes the form of semantic networks, or *mental graphs*, in Pryor's terms. One reason for this is that relations between different objects are better represented by connecting files, and eventually setting up a semantic network of some kind, than by being stored in a single file.

We propose the following addition to mental file theory. Suppose that a certain kind of mental file is anchored to contextually relevant sets; we might call them *set files*. Set files contain an explicit numerosity slot that *may* be filled (if the child possesses concepts of integers). Using a set file amounts to an implicit existential quantification (e.g., on looking at a plate, and realizing that there are crackers on the plate). Our motivation for this addition is that in some circumstances we do treat sets of objects as a single entity. Our perception often zooms in on such sets – like *the set of crackers on a plate*. We may also conceive of such sets as a single entity (e.g., *that pile of crackers*). Still it is likely that certain constraints associated with the concept OBJECT are relaxed in such cases; one example is object permanence (Carey, 2009, pp100-102). Arguably, we think, this account predicts that preschoolers perform poorly at combining concepts of perspective with those of number, for they probably experience difficulties with handling regular and vicarious files, although these difficulties are somewhat different from those formulated by Perner et al. (2015; Huemer et al. 2018). We return to this issue after presenting our studies.

However, mental file theory is not the only possible explanation of why preschoolers falter – if indeed they do – at number-perspective combination. Before considering such an intricate theory we need to examine simpler, performance-related (or domain-general) factors that may at least contribute to preschoolers' difficulties. Thus the goal of the present paper is to examine performance factors that might affect tasks designed to study number-perspective combinations.

Study Question

Given the conflicting predictions above, do or do not preschoolers have a difficulty combining concepts of number with those of perceptual and mental perspective?

Study 1

Participants

We examined 110 participants: Group1 (5;5-6;11, mean: 6;0, n=47, 22 females), Group2 (8;1-9;1, mean: 8;5, n=34, 19 females), and Group3 (8;3-9;0, mean: 8;9, n=29, 25 females). Note that Group2 and Group3 had very similar ranges and means, although the mean age of Group3 was significantly higher than that of Group2 ($t=4.75$; $p<0.001$; difference of the means: 3.55 months, CI95=[2.05, 5.04], effect size: $r=0.521$). However, Group3 came from one higher grade-level than Group2. Hereafter Group2-3 will refer to the union of these two grade schooler groups. All participants were native speakers of Hungarian from Budapest.

Materials and Tasks

Four tasks were administered to all participants. To control for false belief attribution we used the deceptive box task; counting ability was checked by the Give-a-Number task (LeCorre et al., 2006; Carey, 2009, Ch8). Concept combination was measured by the following two tasks.

Smurf Task: Visual Perspective and Number Participants saw a certain number of toy objects (a dog, and a cat, or a few mushrooms, depending on trial type) on a table, and a Papa Smurf character on the opposite side of the table who could see either all objects, or only a subset of them because the remaining ones were, visibly to the participant, blocked out from his view (by what we called his "house"). The cover story was that Papa Smurf was returning home from a long trip during which either mushrooms grew in his house and garden, or a dog and a cat sneaked in, and started to play.

There were four trials: two objects trials that did not require counting, only object recognition, and two number trials that did require counting. Both the objects trials and the number trials consisted of one consistent-view trial, and one inconsistent-view trial. In the *objects consistent* trial, the participant saw two small plastic animal characters, a cat and a dog, playing in Papa Smurf's garden, thus being visible to both the participant and the smurf. Participants were asked two questions: "Can Papa Smurf see the dog?" (correct: yes), and "Can Papa Smurf see the cat?" (correct: yes). In the *objects inconsistent* trial, the dog entered the house before Papa Smurf appeared in the scene, therefore only the cat remained visible to both the participant and Papa Smurf. Thus participants needed to realize on this trial that even though they could see the dog, Papa Smurf could not. The same two questions were asked as in the objects consistent trial (now the correct answer to the first question is no; to the second it is still yes). On the *numbers consistent* trial, three mushrooms "grew" in the garden by the time Papa Smurf arrived home (therefore Papa Smurf could see all of them); finally, on the *numbers inconsistent* trial, five mushrooms grew in the garden, and two in the house. On the latter two trials participants were asked the question "How many mushrooms can Papa Smurf see?" (correct: "three" in the consistent condition, and ”five” in the inconsistent condition). Regarding question order, participants were subdivided into four groups: consistent first vs. inconsistent first crossed with numbers first vs. objects first. For example, in the *consistent first, objects first* trial order the participant first saw the objects consistent trial; then came the objects inconsistent trial; then the numbers consistent scene, and finally the numbers inconsistent one.

Balls-in-Box Task: False Belief and Number Participants saw four short videos. In each video, first an assistant (whom the experimenter called Orsi) entered a room in which there was a box on a table. Orsi dropped a certain number of balls into the box. Then, in the false belief (FB) condition Orsi left the room. Next, another assistant (named Claire) entered through another door which was on another side of the room suggesting that the two did not meet outside. Claire dropped one more ball in the box, then left. Finally, Orsi returned, and reached into the box. At this point the image froze showing Orsi with her arm reaching into the box. The contents of the box were visibly blocked from her point of view. While participants saw the frozen image they were asked "How many balls Orsi thinks there are in the box?" In the control (true belief, TB) condition, Orsi stayed in the room and watched while Claire came in and dropped her ball into the box.

The box in the video was transparent toward the participants: they were able to see the balls in the box after they were dropped, and keep track of their number. Four trials were used: two numerosities (1+1 and 5+1) crossed with two observation conditions – stay vs. leave. In the 1+1 condition Orsi left one ball in the box, and Claire left another. In the 5+1 condition, Orsi left 5 balls in the box, and Claire left a sixth one. All participants received the 1+1 trials first, followed by the 5+1 ones. The order of the stay and leave conditions was counterbalanced: half the participants received the stay condition first (for both numerosities), and the other half saw the leave condition first.

After the first three trials participants were told by the experimenter that Orsi, after reaching in the box, removed all balls from it, and left the box empty. This was to prevent speculations by our participants that on a given trial, balls from the previous trial(s) were still in the box, or at least Orsi thought they were. The correct response was 1,2,5,6: it took responding "one" on the *1+1 leave* trial, "two" on the *1+1 stay* trial; "five" on the *5+1 leave* trial, and "six" on the *5+1 stay* trial.

Results

29 out of 47 Group1 members passed the deceptive box task, while in Group2-3 60 out of 63 did so. That is, Group2-3 did better ($\chi^2(1)=19.599$; $p<0.001$, OR=0.0825 (odds ratio), CI95=[0.014, 0.315]). Based on the Give-a-Number task, all 110 participants were cardinality knowers.

In the Smurf task, 46 out of 47 Group1 members, and all 63 participants in Group2-3 responded correctly in all three control trials. For this reason only the results of the fourth trial (numbers inconsistent) will be analyzed further. In Group1, 29 out of 47 participants passed (62%), whereas in Group2-3, 53/63 participants did so (84%). This difference is significant ($\chi^2(1)=7.134$; $p<0.01$; OR=0.307, CI95=[0.111, 0.810]). Group2 and Group3 performed equally well: 28/34 participants in Group2 (82%), and 25/29 participants in Group3 (86%) answered correctly. Regarding question order, when the object trials preceded the number trials, performance was somewhat better than when the number trials came first, at the level of the entire sample ($\chi^2(1)=5.62$, $p<0.02$; OR=2.659, CI95=[1.088, 6.730]).

On the Balls-in-Box task, Group1 performed poorly: only 5 out of 47 participants answered correctly on all four trials. Of the 42 participants who answered incorrectly, 34 responded “two” in both 1+1 trials, and “six” in both 5+1 trials; it seemed that they simply added up all balls that were dropped into the box, regardless of observation condition. Hereafter we will call such participants *unconditional counters*. Of the 63 Group2-3 members, 38 (60.3%) passed the Balls-in-Box task (including all four trials); of the 25 participants who made some mistake, 17 were unconditional counters. Group2-3 did significantly better ($\chi^2(1)=27.90$; $p<0.001$; OR=0.080; CI95=[0.022, 0.241]). We also found a significant difference between Group2 and Group3 ($\chi^2(1)=8.01$; $p<0.01$; OR=0.212; CI95=[0.055, 0.711]). Of the 34 Group2 members, 15 (44%) answered correctly, and there were twelve unconditional counters (and seven miscellaneous errors). In Group3 (29 participants), 23 passers (79%), five unconditional counters, and one participant exhibiting a miscellaneous error were found.

Question order had an effect in Group 2-3: the *leave* (false belief) condition coming first (for both numerosities) was easier than the *stay* (true belief) conditions arriving first ($\chi^2(1)=7.46$, $p<0.01$; OR = 0.233; CI95=[0.065, 0.761]). Of the 25 Group2-3 members who made some mistake, only seven were in the leave-stay (FB first) condition, and 18 were in the stay-leave (TB first) condition.

Finally, Group1 did much better on the Smurf task than on the Balls-in-Box task: 4 participants passed both tasks, 17 failed both tasks, 25 passed the Smurf task but failed the Balls-in-Box task whereas only 1 passed the Balls-in-Box task while failing the Smurf task (McNemar exact: $p<0.0001$; OR=25; CI95=[4.092,1026.445]). Figure 1 summarizes the results of this study.

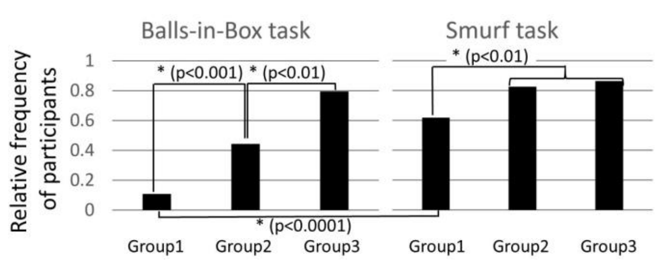

Figure 1: Relative frequencies of correct answers in the two tasks in Study 1.

Interim discussion

One finding of Study 1 is that preschoolers (Group1) performed significantly worse on both tasks than grade schoolers (Group2-3). Still preschoolers did much better on the Smurf task than on the Balls-in-Box task. Altogether the 62% correct ratio on the Smurf task in the preschooler group seems reasonable. More surprising, however, is the low passing rate of the same group on the Balls-in-Box task. This deserves further attention.

For another note, one might object why we used two separate groups with overlapping age ranges. Our rationale is that Group3 had one more year of schooling experience (third grade) than Group2 (second grade). Should this decision seem controversial, the two groups can be regarded as a single one – this is what we did in reporting most of the results above, referring to Group2-3. It might be of some interest still that, as Figure 1 shows, these two (sub)groups did differ on the Balls-in-Box task.

At this point an important question arises: can the low performance of Group1 on the Balls-in-Box task be explained by performance factors? Our preschooler subjects had been skilled counters for their age, due to regular training in Hungarian kindergartens. However, they had virtually no training with false belief tasks. Would a quick familiarization with the latter task increase their success rate on the Balls-in-Box task? This question motivated Study 2.

Study 2

In this study we examined the question whether additional training with the false belief task increases kindergarteners' performance on the Balls-in-Box task. We suspected that the difficulty with the latter task stems from the fact that our kindergartener participants had had a lot of training in counting, but none in false belief attribution.

Participants

35 kindergarteners (all native speakers of Hungarian from middle-class families in Budapest) participated (mean age: 6;5, range: 5;8-7;5; 15 females).

Tasks and Procedure

All participants took the deceptive box test followed by the Give-a-Number task. Subsequently two subgroups were formed, an experimental group (Group4), and a control group (Group5). Both groups received a short training consisting of two tasks before the critical Balls-in-Box task.

Training In the experimental group, the first task was a video version of the Sally-Ann task in which Sally enters the room, leaves her teddy bear in a shopping bag and leaves whereupon Ann enters and moves the teddy bear to a box, then leaves. Upon Sally's return the participant is asked where Sally will look for her teddy. On correct answers we moved on to the second training task. Following incorrect answers an elaboration began: first we asked two questions ("Do you remember where Sally left the teddy?"; "Did Sally see that Ann moved the teddy to the box?"). When the answer to both questions was correct, we offered the following summary: "Because Sally did not see that Ann moved the teddy to the box, she will look for it where she left it – in the bag." When at least one of the answers was incorrect, we showed the video to the child again, during which we gave the correct answer to both questions, followed by a brief explanation. After the video, the same summary was used to conclude. The second training task was a version of the diverse beliefs task (Bartsch & Wellman, 1989). A picture of Linda and her kitten was placed into a scene depicting a bush and a garage. The instruction was: "Here is Linda. She wants to find her kitten which is hiding either in the bush, or in the garage. Where do you think Linda's kitten is hiding?" If the child guessed that the kitten was hiding in the bush, we said "This is a good idea, but Linda thinks her kitten is hiding in the garage.", and we placed the kitten figure on the garage. Upon a "garage" guess we said that Linda thought the kitten was in the bush, and put the kitten beside the bush. Then we asked "So where will Linda look for her kitten – behind the bush, or in the garage?" If the answer was incorrect, we replied "Look! She will look for her kitten where she thinks it is hiding, that is in the _____" filling in the right term, and simultaneously pointing out the cat figure.

In the control group the first task was a version of the diverse desires task (Repacholi & Gopnik, 1997; Wellman & Liu, 2004), and the second task was an emotion recognition task based on facial expressions (Wellman & Liu, 2004). We considered both these tasks to be much less helpful in understanding false belief attribution than the training tasks of the experimental group.

Balls-in-Box Task Only two conditions of this task were used: *5+1 leave* (FB condition, correct answer: 5), followed by *5+1 stay* (TB condition, correct answer: 6), for every participant. A third, *5-1 leave* problem (correct answer: 5) was also presented to those participants who responded correctly to the first two ones.

Results

22 out of 35 participants (63%) passed the deceptive box task. In the experimental group 10 participants passed, and 7 failed; in the control, 12 passed, and 6 failed ($\chi^2(1) = 0.230$; NS). All participants passed the Give-a-Number task up to 6. Regarding the Balls-in-Box task, in the experimental group, 10 out of 17 participants responded correctly to the first two questions, whereas in the control group, only 3 out of 18 did so. This ratio change was significant (Fisher exact: p=0.015, OR=6.703; CI95=[1.222, 50.360]). However, of the 10 participants who correctly answered the first two questions in the experimental group, only 7 passed the third, 5-1 leave condition. In the control group, two of the three participants who correctly answered the first two questions passed the third one. This suggests that despite the improvement that the training in the experimental group seems to have produced, there still remained some performance unstability on this task. All three questions taken together, 7 out of 17

participants passed in the experimental group, and 2 out of 18 participants did so in the control group. This was only marginally significant (Fisher exact: p=0.0599, OR=5.323, CI95=[0.798, 62.608]); thus we must add that further corroborating evidence is needed to fully justify the effectiveness of false belief training in improving performance on the Balls-in-Box task.

Discussion

In the introduction we formulated two opposite predictions, based on two different theories of concept combination, about preschoolers' success in number-and-perspective tasks. Barsalou's multimodal simulation theory predicts that preschoolers should be good at concept combination because they can simulate other agents' viewpoint in imagination (Harris, 2000, 2022). In combining false belief with number in the Balls-in-Box task, preschoolers could in principle resort to multimodal simulations to imaginarily reconstruct Orsi's view, and thus understand what happens in the FB condition (she left the room; the box was blocked out of her view; hence she did not witness further manipulations of the balls in the box). In this framework the abstract inference from Orsi's absence to her ignorance is handled by keeping track of Orsi's visual perspective (five balls dropped; no more seen; hence she should think "five balls in" upon her return).

Mental file theory, with the addition of set files, has resources to explain, even predict, the failure of preschoolers in the balls-in-box task. According to Perner et al., (2015), preschoolers can form regular and vicarious files, but tend to make certain mistakes in handling them. This general principle may have a special application to the balls-in-box task. Namely, participants might naturally unite their own regular files of the two sets of balls (dropped by the two agents respectively) as something like [*Set: balls* (individuating information); *in the box* (predicative information)], or simply deploy a single regular set file from the outset to refer to all the balls. But then, dividing this set into two subsets when hearing the question, for purposes of vicarious file deployment, may be a pretty difficult task for young mental files administrators. Unlike in the Perner at al. case, however, the core difficulty is not handling embedded perspectives (i.e., that a conceptual perspective on an object is not available from an agent's situational perspective), but rather, to attach a vicarious file to a regular set file based on a division of the set referred to by the regular file (so that the vicarious file extends only to a subset of the reference set). This might create a substantial difficulty for preschoolers. Thus the mental files account arguably predicts the difficulties kindergarteners have with the Balls-in-Box task. A similar problem arises in the Smurf task, although in that task, separation may be easier to maintain at the level of regular files (on the *numbers inconsistent* trial): mushrooms in the house are presented as spatially separated from those in the garden.[1] This may have helped children with vicarious file deployment. Figure 2 summarizes the proposed idea of set files.

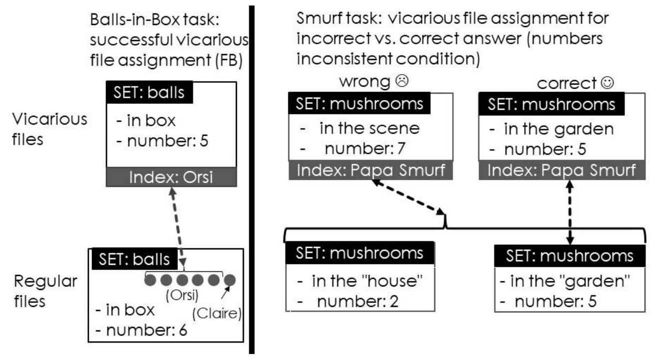

Figure 2: Summary of the idea of set files. The wrong answer to the Balls-in-Box task (*5+1 leave* condition) can be modeled by the horizontal curly brace embracing all six balls in the regular file.

Results of Study 1 showed that, whereas preschoolers achieved a reasonable level of success with combining visual perspective taking with number (in the Smurf task), they were surprisingly unsuccessful at combining false belief with number (in the Balls-in-Box task). Following this we kept on using the latter task exactly because it yielded unexpected results, and we wanted to understand why that happened. Moreover, before drawing hasty conclusions about the correctness/falsity of theories of concept combination, we aimed to peel off any potential performance limitations this task may produce, then, once the ground is cleared, retest it against different theories of concept combination. In this paper we report results of the former stage; the latter one is in progress.

This is our strategy, and we did indeed find a number of indications in our data that domain-general factors play an important role in solving the Balls-in-Box task. Here is a list of such hints.

1. Schooling seems to make a difference (Study 1, Groups 2 versus 3), despite the fact that false belief problems are not specifically addressed or practised in any way in Hungarian primary education.
2. Control (TB) trials coming first in the Balls-in-Box task tended to distract participants' attention from the crucial aspect of the subsequent FB trials (Study 1).
3. Even a little practice with standard FB tasks tended to enhance success on the Balls-in-Box task (Study 2).
4. The Smurf task, which targeted the combination of perceptual perspective with number did not prove unexpectedly difficult for preschoolers, still a minimal practice (i.e., object trials coming before number trials) had a small but significant effect on success (Study 1).
5. Earlier pilot studies (not reported here) indicated that the Balls-in-Box task was extremely difficult even for grade

[1] Unlike in the Balls-in-Box task where balls dropped by the two agents constituted a single pile; see Figure 3.

schoolers (age range from 7;9 to 10;5) when the box was not transparent toward the participant (Figure 3). On this version grade schoolers did not perform any better than preschoolers; in both age groups success rate for the four questions was around 10 per cent.

Figure 3: Transparent and non-transparent box versions of the Balls-in-Box task. Study 1 and Study 2 reported here used the transparent box version (left).

Performance Factors in Our Tasks

We suggest that the difference between the difficulty of the two tasks (Smurf and Balls-in-Box) can, for the most part, be explained by performance factors (or domain-general mechanisms). We identified four such factors that figure in the two tasks; all four of them seem to interestingly differentiate the two tasks. Here is the list.

The first factor is *task complexity*. In the Smurf task, perspective taking can be understood in terms of gaze-tracking: the Smurf's line of sight needs to be assessed to decide whether it is blocked by the wall or not. In the Balls-in-Box task, however, gaze tracking all by itself is insufficient to reach the correct response. In false belief attribution one needs to take into account the attributee's recent personal past as well (e.g., that Max did not see the removal of the chocolate from its original location; Wimmer & Perner, 1983). Similarly, the Balls-in Box task requires taking into account events that happened in the recent past (i.e., having witnessed the initial placement of balls in the box). The Smurf task makes no such demand.

The second factor is *prepotent response inhibition*. Unconditional counting needs to be inhibited to arrive at the correct response on the critical trials. This became an important factor in our tasks because all of our subjects were good at counting and simple addition tasks, likely due to regular kindergarten training.[2] Hence adding all the balls or mushrooms together was a habitual response for them when encountering two sets with small numerosity. This aspect is present in both tasks, although, as we noted above, spatial separation of the two subsets of objects in the Smurf task mayS have been helpful for participants.

The third factor is *different levels of practice and automatization*. Contrary to counting, our participants had little familiarity with false belief attribution. In the Smurf task, children again can deploy their counting powers, and visual perspective taking is arguably also familiar to them by ages 5 to 6. Starting early on, children gain substantial practice in visual perspective taking, which may come from games like peekaboo, then, later, hide-and-seek (e.g., Peskin & Ardino, 2003). However, it is also arguable that by age 5, children become familiar with deceiving (that is, inducing false belief in) others (Chandler et al., 1989; Sodian et al., 1991). Still, actively inducing false belief in someone versus understanding that, in a given circumstance, someone will come to have a false belief are somewhat different abilities. Even though, at ages 5 to 6 deception and false belief attribution are likely subserved by similar metarepresentational mechanisms (Jakubowska & Białecka-Pikul, 2019), we suspect that deception may be more familiar to children from everyday life than false belief attribution "in the abstract", involving neither the deceptive agent role nor the suspicion that they themselves might be victims of deception. Here we see an analogy with abstract versus realistic versions of Wason's selection task (see Evans, 2012 for a review).

Finally, the fourth domain-general factor is *difference between interpreting an event sequence vs. a static scene*. Not only do false belief tasks take a certain aspect of the attributee's recent past as a premise of attribution, but to actually take the other person's past into account, a series of unfolding events needs to be held in working memory.[3] On the contrary, the Smurf's perspective can be assessed from a single static scene.

To summarize, the Smurf task invokes perspective taking which is less complex (first factor), less demanding (fourth factor), and possibly more familiar (third factor) to preschoolers than false belief attribution. In addition, Study 2 supplies some evidence that a little practice with false belief attribution improves kindergarteners' performance on the Balls-in-Box task. Prepotent response inhibition (second factor) seems to be a major obstacle in the Balls-in-Box task (recall the high frequency of unconditional counters in the experiments), but not so in the Smurf task. We think this is due to an interaction with the other three differentiating factors which made the Balls-in-Box task more difficult, as well as to the spatially separated presentation of the two subsets in the Smurf task.

---

[2] Of the 42 Group1 members who did not respond correctly to all four questions in the Balls-in-Box task, only four made errors *in adding up all the balls* in the box. Of the 18 members of Group1 who erred on the Smurf task (*numbers inconsistent* trial), again only four children committed an error in adding up the mushrooms in the house and those in the garden.

[3] This is definitely true of the Balls-in-Box task, even though some versions of the Sally-Ann task (for example, ones that present the scenario in the form of a comic strip so that each stage is simultaneously present and can be revisited during solving the task) apparently require much less working memory involvement. Still, taking into account the attributee's recent past is logically necessary to reach the correct response in both versions, regardless of whether the WM demand is sustained.

Acknowledgments

The authors wish to thank Ágnes Kovács, Ernő Téglás, and Josef Perner for their helpful comments and suggestions at earlier stages of this research. We are grateful to Hajnalka Hajnal, director of the Kindergarten of the Hungarian Academy of Sciences, for her help in recruiting participants. We are also indebted to Orsolya Bárdos and Andrea Sándor for their help with data collection.

References

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.

Barsalou, L., W. (2017). Cognitively Plausible Theories of Concept Combination. In J. A. Hampton & Y. Winter (Eds.), Compositionality and Concepts in Linguistics and Psychology. Language, Cognition, and Mind 3, Springer.

Bartsch, K., & Wellman, H. (1989). Young Children's Attribution of Action to Beliefs and Desires. *Child Development*, Vol. 60, No. 4, 946-964.

Carey, S. (2009). *The Origin of Concepts*, Oxford University Press

Carruthers, P. (2016). Two Systems for Mindreading? *Rev.Phil.Psych*. 7, 141–162.

Chandler, M., Fritz, A. S., & Hala, S. (1989). Small-Scale Deceit: Deception as a Marker of Two-, Three-, and Four-Year-olds' Early Theories of Mind. *Child Development*, *60*, 1263-1277.

Evans, J. St. B. T. (2012). Dual-Process Theories of Deductive Reasoning: Facts and Fallacies. In: Holyoak, Keith, J., Morrison, Robert, G. (Eds.): *The Oxford Handbook of Thinking and Reasoning*, Oxford University Press, 115-133.

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.

Fodor, J. A. (1998). *Concepts – Where Cognitive Science Went Wrong*, Clarendon Press, Oxford

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, 44, 43–74.

Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: from reals to integers. *Trends in Cognitive Sciences*, 4, 59-65.

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

Gopnik, A., & Astington, J., W. (1988). Children's understanding of representational change, and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, 59, 26-37.

Hampton, J., A., & Winter, Y. (Eds.). (2017). *Compositionality and Concepts in Linguistics and Psychology*. Language, Cognition, and Mind 3, Springer.

Harris, P., L. (2000). *The Work of the Imagination*. Blackwell

Harris, P., L. (2022). *Children's Imagination*. Cambridge University Press

Hinzen, W., Machery, E., Werning, M. (Eds.). (2012). *The Oxford Handbook of Compositionality*, Oxford University Press

Huemer, M., Perner, J., & Leahy, B. (2018). Mental files theory of mind: When do children consider agents aqcuainted with different object identities? *Cognition, 171*, 122-129.

Jakubowska, J., & Białecka-Pikul, M. (2019). A new model of the development of deception: Disentangling the role of false-belief understanding in deceptive ability. *Social Development*, *29*(1), 21-40.

Laurence, S., & Margolis, E. (1999). Concepts and Cognitive Science, in S. Laurence and E. Margolis (Eds.), *Concepts*, Cambridge Mass. The MIT Press

LeCorre, & M., Brannon, E. M., Van de Walle, G. A., & Carey, S. (2006). Re-visiting the competence/performance debate in the acquisition of the counting principles. *Cognitive Psychology*, 52(2), 130–169.

LeCorre, M., & Carey, S. (2007). One, two, three, four, nothing more: How numerals are mapped onto core knowledge of number in the acquisition of the counting principles. *Cognition*, 105, 395–438.

Moll, H., & Meltzoff, A., N. (2011). Perspective-Taking and its Foundation in Joint Attention, In Johannes Roessler, Hemdat Lerman, Naomi Eilan (Eds.) *Perception, Causation, and Objectivity*, Oxford University Press, Oxford (pp.286-304)

Murphy, G. (2002). *The Big Book of Concepts*. MIT Press

Perner, J., Huemer, & M., Leahy, B. (2015). Mental files and belief: A cognitive theory of how children represent belief and its intensionality, *Cognition* 145 (2015) 77–88.

Perner, J., & Ruffman, T. (2005). Infants' Insight into the Mind: How Deep? *Science*, *308*, 214-216.

Peskin, J., & Ardino. V. (2003). Representing the Mental World in Children's Social Behavior: Playing Hide-and-Seek and Keeping a Secret, *Review of Social Development* 12(4): 496-512.

Piaget, J. (1962). *Play, Dreams, and Imitation*. Routledge and Kegan Paul.

Pryor, J. (2016). Mental Graphs. *Review of Philosophy and Psychology*, *7*(2), 309-341.

Qureshi, A. W., Apperly, I. A., & Samson, D. (2010). Executive function is necessary for perspective selection, not Level-1 visual perspective calculation: Evidence from a dual-task study of adults. *Cognition*, *117*(2), 230-236.

Recanati, F. (2012). *Mental Files*. Oxford University Press

Recanati, F. (2016) *Mental Files in Flux*, Oxford University Press

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology*, *33*(1), 12–21.

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.

Sodian, B., Harris, C. T., Harris P. L., & Perner, J. (1991). Early Deception and the Child's Theory of Mind: False Trails and Genuine Markers. *Child Development*, *62*, 468-483.

Sodian, B., Thoermer, C., & Metz, U. (2007). Now I see it but you don't: 14-month-olds can represent another person's visual perspective. *Developmental Science*, *10*(2), 199-204.

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

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.

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, *75*(2), 523-541.

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

Wynn, K. (1990). Children's understanding of counting. *Cognition*, *36*, 155–193.

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, *24*, 220–251.

Wynn, K. (1998). Psychological foundations of number: numerical competence in human infants. *Trends in Cognitive Sciences*, *2*, 296–303.