Understanding of spatial correspondence does not contribute to representational understanding: Evidence from the Model Room and False Belief tasks

Catherine M. Sayer

Martin J. Doherty[1]

School of Psychology, University of East Anglia, Norwich, NR4 7TJ, UK

Author Note

We wish to thank the children, parents and teachers of Watton Westfield Infant and Nursery School, Chestnut Boutique Nursery School St Giles, Fleggburgh Church of England Voluntary Controlled Primary School, Little Explorers Nursery, East Dereham Day Nursery, Treetots Playgroup and Cheeky Smiles Ltd. We also thank Dr Patricia González Peña for methodological discussions and assistance, Dr Louise Ewing and Professor Andrew Bayliss for invaluable discussion about the manuscript, and Dr Robin N. Campbell for many long discussions about the topics raised in this study. Experiment 1 and the first phase of Experiment 2 were not preregistered. The second phase of Experiment 2 was preregistered as part of a commitment henceforth to preregister all work from our lab. See https://aspredicted.org/PRV_5CD

[1] Corresponding author

Understanding of spatial correspondence does not contribute to representational understanding: Evidence from the Model Room and False Belief tasks

Abstract

We examine the longstanding claim that understanding relational correspondence is a general component of representational understanding (Perner, 1991). Two experiments with 175 preschool children located in [blinded] examined use of a scale model (DeLoache, 1987) comparing performances on a 'Copy' task, measuring abstract spatial arrangement ability, and the False Belief task. Consistent with previous studies, younger children performed well in scale model trials when objects were unique (e.g., one cupboard) but poorly at distinguishing objects using spatial layout (one of three identical chairs). Performance was specifically associated with Copy task but not False Belief performance. Emphasizing the representational relation between model and room was ineffective. We find no evidence for understanding relational correspondence as a general component of representational understanding.

Public significance statement: This study compares performance on a classic hiding game, using a room and its scale model, with preschool children's abstract spatial ability and understanding of mental representation. We found performance on the classic Model Room task only related to spatial abilities, potentially important in later science and mathematics education.

Keywords: Spatial correspondence; Model Room task; False Belief task; Theory of Mind

Here we examine a long-standing claim about humans' general understanding of representation. Types of representation include mental states, language, pictures, and scale models. By 'understanding of representation', we mean the ability to think about the relation between the representational medium and the thing it represents, also known as metarepresentation. For example, theory of mind is the ability to think about the relation between the mental states of others or oneself (belief, desire, emotions, and so on) and the things these represent or are about.

The claim under consideration was first made by Perner (1991, p.78): "Understanding the use of representations to truthfully inform about the represented world requires understanding of correspondence". Representations typically correspond in systematic ways to their referents: relations that exist in the referent (say, an object is inside a wardrobe in a room) are also encoded in the representation of the referent (e.g., a scale model, or picture of, or belief, or statement about the room). This correspondence is necessary for representations to perform their function. Perner's claim goes further, to suggest that to be able to understand representations as such, one must acknowledge this mirroring of relations within representation and referent.

Hoyos et al. (2020) make the similar claim that mental state concepts develop through structural mapping processes (Gentner, 2005). They argue that extracting common structural relations between mental states and reality, or two mental states representing the same reality, is a key process in theory of mind development. Baldwin and Saylor (2005) and Bach (2014) make related proposals.

Common or domain general mechanisms promoting representational understanding might help explain findings of general metarepresentational development around the age of four years. The False Belief task is widely accepted as a diagnostic test of theory of mind. It requires children to predict action based on an agent's mental state even when it does not match the real situation. Success on the False Belief task associates with success on tasks measuring non-mental representation. These include understanding that things can have alternate names (Doherty & Perner, 1998; Perner et al. 2002), that one word (e.g., *bat*) can have two meanings (e.g., Doherty, 2000), and that a novel word is likely to refer to a novel object (Gollek & Doherty, 2016; Karadaki, 2019; Karadaki & Doherty, 2017). In the pictorial domain, children can acknowledge that ambiguous figures have two interpretations (e.g., duck or rabbit; Jastrow, 1900) around this age, and this ability is associated with understanding of false belief (Doherty & Wimmer, 2005), alternative naming, and homonymy (Wimmer & Doherty 2011). False belief understanding is also associated with proficient use of pictorial information when completing jigsaws (Doherty et al., 2021).

These findings suggest common developments in understanding of mental and non-mental representation around the age of four years. Conceptually this has been described as a general understanding of *perspective*, allowing the understanding that representations of the same thing can differ in content, e.g., different beliefs about the same situation, or different ways of referring to the same object (Perner et al., 2002; see Doherty & Perner, 2020, for a review and further theoretical analysis). Here we consider understanding of correspondence as a possible common factor supporting these conceptual developments.

We examine this claim in the context of a classic task involving a scale model of a room (DeLoache, 1987). A room and its scale model, complete with model furniture, are laid out in exact spatial correspondence. Children see an object hidden in one room and search for a similar object "in the same place" in the other room. Typically, children succeed around 36 months.

The Model Room task apparently requires children to understand the correspondence relations between the two spaces. Correspondence has typically not been defined in past research, and here requires it. The dictionary definition of 'correspond' includes distinct senses of the word. One emphasises similarity ('have a close similarity; match or agree almost exactly'). The other emphasises analogy ('be analogous or equivalent in character, form, or function') (Apple Dictionary Version 2.3.0). In the model room task 'correspondence' is used in the analogical sense. To perform the task, one must understand that facts about the model (e.g., a sticker is inside the miniature wardrobe) are informative about the room (that a sticker is also inside the large wardrobe).

This makes it clear that correspondence is relevant to representation. Typically, representations function to provide information about their referents, and the two should therefore correspond. Beliefs should mirror reality, photographs should encode contemporaneous information about the scene, maps and scale models should inform about the spaces they correspond to. However, although correspondence and representation have similarities, understanding correspondence is <u>not</u> equivalent to understanding representation. Perner (1991, p.81) illustrates this point with an analogy: British houses in the same neighbourhood are frequently built to the same plan, so knowledge of one's own house allows you to find the bathroom in a neighbour's. Clearly the houses correspond, but do not represent one another. Perner's claim is that understanding this correspondence is a necessary but not sufficient component of full representational understanding. Nevertheless, understanding correspondence may be a common factor underlying understanding of representation in general.

Two levels of correspondence can be distinguished in the Model Room task. In the task sets of similar objects are laid out in the same relative spatial arrangement. In the original task, each object was unique (one bed, one couch, and so on). Thus, the target could be found by noting which type of furniture it was under in the model, and locating the same type of furniture in the room, without regard for the layout of either space. We refer to this here as item-to-item correspondence. Alternatively, both the identities of the furniture and of the spatial layout could be encoded and used to locate the item: ‘spatial correspondence’.

It was tacitly assumed that the model room task measured spatial correspondence. This is natural: for us adults, the identical layout of the two spaces is highly salient. However, data from Blades and Cooke (1994) suggests that 3-year-old only use item-to-item correspondence to pass the task.

Blades and Cooke (1994) modified the task so that two of the items of furniture were identical. Thus, the spatial arrangement had to be used to identify which of these was the correct hiding place. Three-year-olds, who performed well on the standard task, were at chance. Four-year-olds successfully chose between identical hiding places on about 80% of trials. Blades and Cooke concluded that three-year-olds did not consider the spatial relationships encoded in the model[1].

This shift to using the corresponding spatial relations suggests two competing explanations. One is representational: the shift results from understanding the representational relations between the model and the room. This new understanding may alert children to the function of the model (as a representation of the room) and how it serves this function (by being laid out in an identical manner), thus causing them to attend to the spatial correspondence between the two. The developmental trajectory is very similar to other metarepresentational developments.

[1] Troseth et al. challenge this conclusion. They modified the Model Room task by removing the hiding and retrieval elements, simply asking children to find which piece of furniture in the model “looks like” the one on the room. Two-and-a-half-year-old children performed well at this matching task. They therefore claim that even 2-year-olds can detect “low-level object correspondences. However, although the task requires noting the similarity of two objects, absence of the hiding and retrieval elements means it does not require understanding correspondence in the analogical sense. Both Item-to-item and spatial correspondence require this sense. Thus 2-year-olds’ success on the matching task does not indicate anything about later success in model room tasks.

The second possible explanation is that it directly reflects a development in children's sensitivity to spatial relations within a stimulus. Various studies suggest that when assessing similarity young children first focus on similarity of objects, then later shift to also focus on similarity of relations between objects. This occurs with grammatical development (e.g., verb learning; Childers et al., 2016), analogical reasoning (e.g., Richland et al. 2006), and spatial mapping (Yuan et al., 2017). This shift of focus from objects to relations has been claimed to reflect general processes of analogical reasoning developing in the preschool period (Gentner, 2010). The two explanations are not necessarily entirely distinct: general processes of analogical reasoning may influence metarepresentational understanding, potentially consistent with Perner's (1991) claim. Bach (2014) and Hoyos et al. (2020) also suggest a role for analogical reasoning in the development of false belief understanding.

The ability to encode and decode analogical spatial relations has received considerable attention in the context of maps. Maps typically stand in spatial correspondence relationships to their referent[2], and the development of map reading is clearly relevant to the claim that a general understanding of spatial correspondence is developing in the preschool period. Maps differ from scale models in that they are less iconic, two dimensional, and typically only provide an 'overhead' viewing angle. These differences plausibly make determining correspondence between maps and their referents more difficult. Nevertheless, most research does suggest that basic map reading skills emerge in the preschool period, and undergo protracted subsequent development (Blades & Spencer, 1994; Liben & Yekel, 1996).

---

[2] However, not always. For example, the diagrammatic map of the London Underground shows serial positions of stations along lines, with little correspondence to geography; Degani, 2013)

In a classic study, Bluestein and Acredelo (1979) used a room with exclusively non-unique target locations (four identical blue boxes), indicated on an aerial view map by blue squares (with a minimal perspective view of their front sides). When viewing the map inside and aligned with the room, only 55% of 3-year-olds reached the pretest criterion of 2/3 trials correct, compared to 86% of four-year-olds and all 5-year-olds.

Performance on simpler maps also seems somewhat challenging for 3-year-olds. Huttenlocher et al. (1999) used a minimally simple rectangular map with a single dot representing the position of a buried object in a featureless rectangular sand pit. All 4-year-olds and 60% of 3-year-olds were able to use these maps to find the object. The 40% of 3-year-olds who failed the task did not appear to be using the map at all, suggesting conceptual rather than performance deficits. Shusterman et al. (2008) examined 4-year-olds' ability to identify non-unique locations using simple 3-location maps. Performance was mostly good if one location was unique, functioning as a landmark. When all three locations were circles, performance did not exceed chance for linear or isosceles triangular arrays but did for a right-angled triangular array. This again suggests a basic understanding of spatial correspondence by late three or four years old, but complex layouts continue to pose additional difficulties.

One study found some ability to identify non-unique locations in a map task in younger children. Winkler-Rhoades et al. (2013) used simple maps similar to Shusterman et al.'s (2008). They found that 2.5-year-olds could correctly identify the central of 3 non-unique locations if they were in a line, or the apex of an isosceles triangular formation. They were at chance for the two ends of the linear formation, or the two base locations of the isosceles triangle. In fact, their isosceles triangle performance was better than Shusterman et al.'s 4-year-olds, who did not exceed chance even at the apex.

These findings are potentially important in suggesting early signs of very basic understanding of spatial correspondence Winkler-Rhoades et al. (2013) speculate that they find considerably better performance than most other studies because of the abstract nature of their maps. However, Shusterman et al.'s (2008) maps were equally abstract. Salsa et al. (2019), using similarly abstract maps and locations, found 2.5-year-old's performance did not differ from chance (39% compared to a 33% chance baseline) with children not performing well until four years. However, they did not separately analyse central linear or apex triangular performances, the locations at which Winkler-Rhoades et al. found above chance performance.

Taken together, most evidence suggests that the basic ability to use maps to identify locations based on spatial layout develops in the preschool period, roughly around four years. It continues to develop thereafter as children become able to deal with the additional factors of decoding relative distances, angles, and geometrical form from maps, and to use maps when not aligned with or viewed outside of the referent space (Bluestein & Acredelo, 1979; Dillon et al. 2013; Dillon & Spelke 2018. Liben & Yekel, 1996). Development of these skills extends well into adulthood (e.g., Liben et al. 2008). Much of this can be seen as performance issues. Here we are interested in the basic conceptual insight that two spaces can correspond in their spatial layout, and whether this understanding relates to understanding of representation.

We examine this using model rooms. Scale models are more concrete and iconic than maps, and like the referent spaces are three dimensional. The Model Room task is arguably suitable for three additional reasons: 1. Perner made his correspondence claim using this task as the example; 2. DeLoache and colleagues claim the task measures symbolic and representational understanding[3]; 3. the task has also been employed in research examining development of relational reasoning (Hoyos et al., 2020; Loewenstein & Gentner, 2001).

Two existing studies have examined the hypothesis that correspondence understanding underlies general metarepresentational development. Walker and Murachver (2012) and Lillard and Kavanaugh (2014) examined longitudinal relations between Model Room and False Belief task performances. Walker and Murachver found that scale model performance at 36 months and 42 months showed low associations with false belief performance at 42 months (r = .30 and r =.22 respectively, one-tailed ps<.05), and associations between earlier scale model performance and false belief understanding at 48 months were fully mediated by language. Lillard and Kavanaugh found that performance on the Model Room task (combined with a similar hiding task using a photograph instead of the scale model) was not significantly related to later theory of mind performance (r =.26 and r =.24 at 4- and 5-years, two-tailed ps >.05).

Neither study's findings suggest that the two tasks share a common factor. However, both used the standard version of the Model Room task which, as discussed, requires understanding of only item-to-item correspondence. Here we adopt Blades and Cooke's (1994) extension of the task to compare the more complex spatial correspondence understanding.

Present study

We aim to compare three abilities:

1. Object search using scale models, including unique and non-unique hiding places.

2. Theory of mind.

[3] DeLoache and colleagues treat the terms 'symbolic' and 'representational' as synonymous (e.g., Troseth et al. 2007, p. 763), as is common in the literature (e.g., Callaghan, 2020). We will use 'representational' throughout.

3. Non-representational spatial relational ability.

For (1) we adopt Blades and Cooke's (1994) extension of the Model Room task to compare: unique correspondences, where one exemplar of a furniture type is in each space; and non-unique correspondences, where three identical pieces of furniture were in each space, distinguished by location. We used two identical model rooms, simultaneously visible, to maximise the clarity of the correspondence relationships and minimise memory requirements. Blades and Cooke also used identical model rooms in one experiment, and a real room and its scale model in a second experiment, finding equivalent performance in the two experiments.

Most Model Room experiments present the room and its model as distinct rooms, typically one belonging to big Snoopy and one to little Snoopy (using appropriate dolls). The reason given for the correspondence is that little Snoopy and big Snoopy like to do the same things (i.e., be in corresponding locations in their rooms). While this cover story may communicate the correspondences involved simply, it does not present them as representational, but instead based on the arbitrary preferences of the rooms' owners. Plausibly this makes it harder to approach the task as researchers intend.

Here, instead of a cover story, we directly described the similarity of the rooms, then presented the task as a hiding game in which the target was "in the same place in the other room". In Experiment 2, we further emphasised the representational nature of the relation between model and room, by using different sized model rooms, so that one is a true scale model of the other, and explicitly stating the function of the smaller room is to show where the object was hidden in the big room.

For (2) we employ the unexpected transfer version of the explicit False Belief task. It is widely accepted as a diagnostic of understanding of mental representation (Doherty, 2009) and has served as the comparator task in the non-mental representation research reviewed above. Children typically pass the task around four years (Wellman et al., 2001).

For (3) we use a task in which children copy a simple spatial array. No representational relations are implied, and objects are not thematically related. This task requires arranging a set of four objects to create a spatial match with four identical objects on a six-square grid. It has been extensively piloted, and performance improves rapidly between two-and-half and four years (Doherty & Anderson, 2003).

We investigate two hypotheses predicting a developmental relation between the adapted Model Room task and False Belief task. The first more general one is that both require an understanding of correspondence, and therefore performance on each will be related. The second hypothesis is more specific. The ability to identify hiding places based on spatial correspondence develops around four years, roughly the age children succeed at the False Belief task and comparable non-mental representation tasks. This synchrony supports the hypothesis that understanding spatial correspondence is an immediate precursor to understanding mental and non-mental representation. We will conclude that neither the general nor specific hypotheses are supported.

Experiment 1

Participants

Participants were 89 children (47 girls): 35 2- to 3-year-olds (M = 3;5, SD = 5.5m, range 2;7 to 3;11), 38 4-year-olds (M = 4;5, SD = 3.5m, range 4;0 to 4;11), and 16 5-year-olds (M = 5;4, SD = 3.0m, range 5;0 to 5;9) recruited from predominantly working-class preschools in [blinded].

Inclusion criteria were informed parental consent, and child assent immediately prior to testing. No child met the exclusion criterion of a teacher- or parent-indicated special-needs diagnosis. The stopping criterion was all available children had been tested.

Previous Model Room studies have modest sample sizes. Studies discussed above range from eight (e.g., Troseth & Deloache, 1998, Experiment 2) to 77 (Lillard & Kavanaugh, 2014), the largest we are aware of. These sizes may reflect practical difficulties maintaining a real domestically furnished room in a laboratory. We instead used two identical model rooms, following Blades and Cooke (1994). This additionally maximises the clarity of the correspondence relations and minimises memory requirements.

Our principal planned analysis was correlation of the main task variables. The G-Power 3.1.9.2 statistical tool (Faul et al., 2007) estimated sample sizes required to achieve a power of 0.80 with two-tailed alpha set to 0.05 as follows: correlation, medium effect size of $r = .30$, $N = 84$; comparison of means using Wilcoxon signed-rank tests, medium effect size of $d = .30$, paired samples $N = 94$, independent samples N = 184. Given the lower power of the Wilcoxon tests, we confirmed all comparisons with t-tests, medium effect size of $d = .50$, paired samples $N = 34$, independent samples $N = 64$. These comparisons yielded equivalent results, so we report only the Wilcoxon test results. The data are not normally distributed, but parametric tests are robust to violation of the assumption of normality (see Rasch & Guiard, 2004, for simulation-based evidence of robustness). Experiments 1 and 2 had 89 and 86 participants, ensuring sufficient power for correlation and parametric comparison of means.

Design

Each child was tested in a quiet, familiar location for around 15 minutes on the Model Room, two False Belief, and Copy tasks. Task order was randomised, observing the constraints that Model Room and Copy Tasks were presented first or third and the false belief tasks presented second and fourth. The first author collected all data in single sessions.

Materials and Procedure

Model Room Task

Two identical model bedrooms (30cm wide x 22cm deep x 14cm high) were presented side by side in the same orientation and layout. Each contained a bed, wardrobe, cupboard, and three identical ('non-unique') chairs arranged as in Figure 1. Children were told "Here we have two rooms, and look, they are the same". They were then shown that each item in one room was the same as the corresponding item in the other, e.g. "This bed [points] is the same as this bed [points to the other bed]". The task was presented as a hiding game, "We are going to play a hiding game with these stickers, and if I hide a sticker in one room, the same sticker will be in the same place in the other room. Every time you find a sticker you can keep it".

Figure 1

Model Room Task Materials for Experiments 1 and 2.

Standard Task

Explicit Task used in Experiment 2

To familiarise children with the materials the researcher placed a sticker in full view on a piece of furniture in one room and asked, “Can you put this sticker in the same place in the other room?”. Feedback was given: if children were correct, they were told “That is right. See how they are both on the *bed*”. Six children (*range* = 37-56m, *M* = 46.2m) responded incorrectly; they were told “See how the sticker is on the *bed* but you put the sticker on the *chair*, shall we put it on the *bed* so they are in the same place”.

In six experimental trials, the sticker was concealed underneath or inside the furniture. Children were shown where the sticker was in one room and told "The sticker is in the same place in the other room". Questions were asked in order:

Retrieval Analogous: Can you find it in the other room?

Retrieval Original: Where's the sticker in this room? [indicating the original room.]

Each location was tested once, order counterbalanced between children; chair trials were never consecutive. Success was correct search on the first attempt, 0 to 3 trials for unique and 0 to 3 trials for non-unique items of furniture.

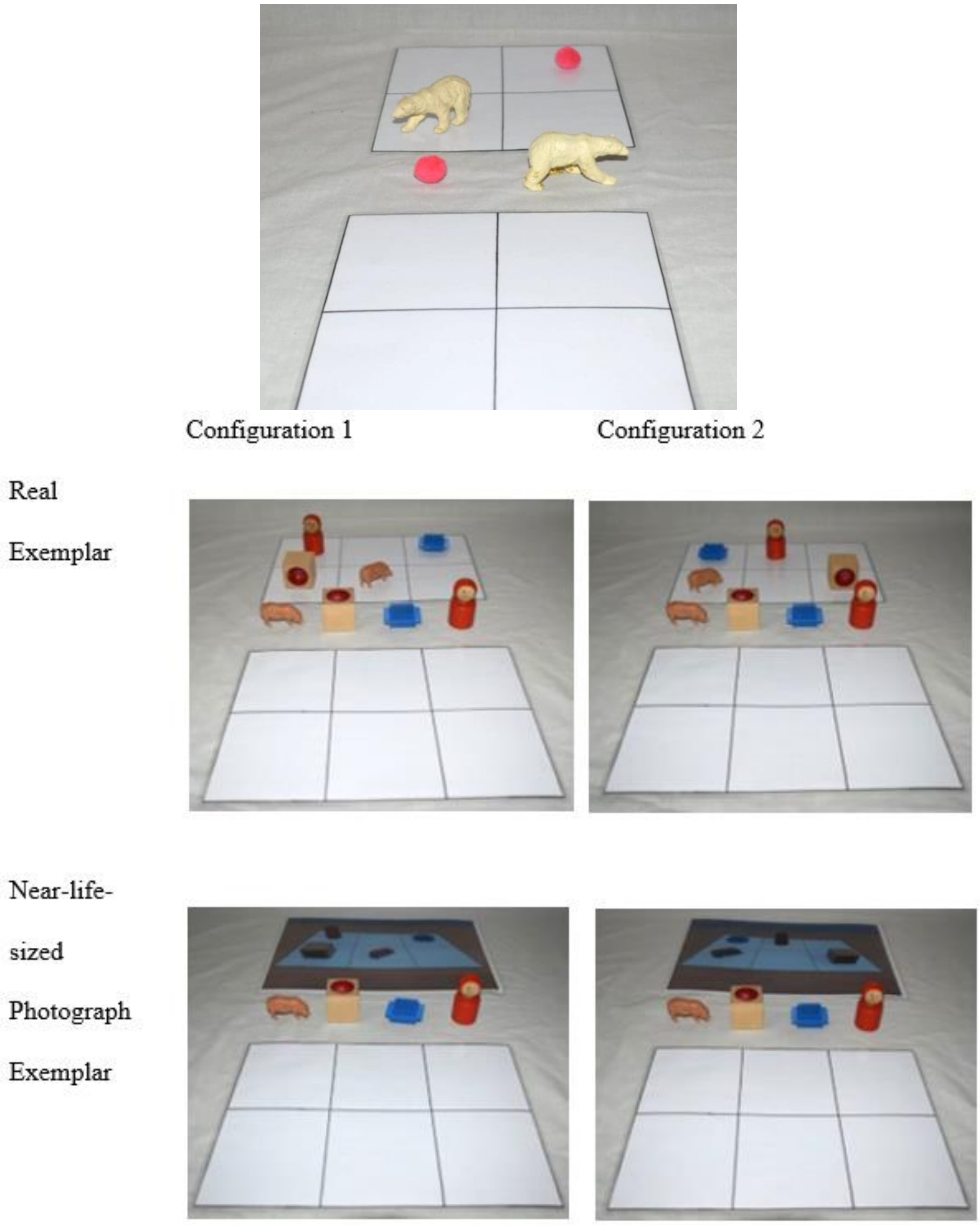
Copy Task

In the warm-up task children were shown two 2x2 grids (18x18cm). One had two toys placed as shown in Figure 2. Children were shown the other grid and two identical toys and asked, "Can you make this [indicates empty grid] look like this [indicates completed grid]?". They were given feedback: if participants placed an object correctly, they were praised, and the two corresponding quadrants were pointed out. If they placed it incorrectly, the researcher pointed out the different locations of the object in the exemplar and test grids and moved the test objects pointing out that they were now "the same".

The experimental task had a 2x3 grid (18x27cm) and four toys distributed irregularly as shown in Figure 2. Between the exemplar grid and the child, an empty grid and four identical toys were placed. The child was asked to make their grid look like the exemplar. Half the children saw the real exemplar grid on the first trial then a near-life-sized exemplar photograph presenting the same point of view; half had the opposite order. Order was crossed with the two grid configurations shown.

Figure 2

Copy Task Materials: Warm-Up, Test Configuration and Exemplar Type



Performance was scored on the number of matching squares between exemplar and the child’s grid; since it was impossible to match only five squares (a single misplaced piece yields two incorrect squares) fully correct performance received five points, yielding an ordinal scale score of zero to five. Task chance baseline is 1.33 (based on a total score for all 720 possible combinations of 958).

False Belief Task

A short story was acted out with two Playpeople dolls (5cm), a marble, an opaque jar (5 x 2.5cm), and a box (3 x 4cm). Sally placed a marble in the box and left. Tony then moved the marble to the jar. Sally returned and children were asked the following questions:

False belief test question: “Where will she look first for her marble?”

Reality question: “Where is the marble really?”

Memory question: “Where did Sally put the marble in the beginning?

Children passed the False Belief task if they answered all three questions correctly.

The second procedure was the same with dolls Lucy and Billy (3.5cm), an oval box (5cm high x 5cm x 9cm), a square box (4cm x 5cm), and a penny.

Figure 3

False Belief Materials

Sally and Tony

Lucy and Billy

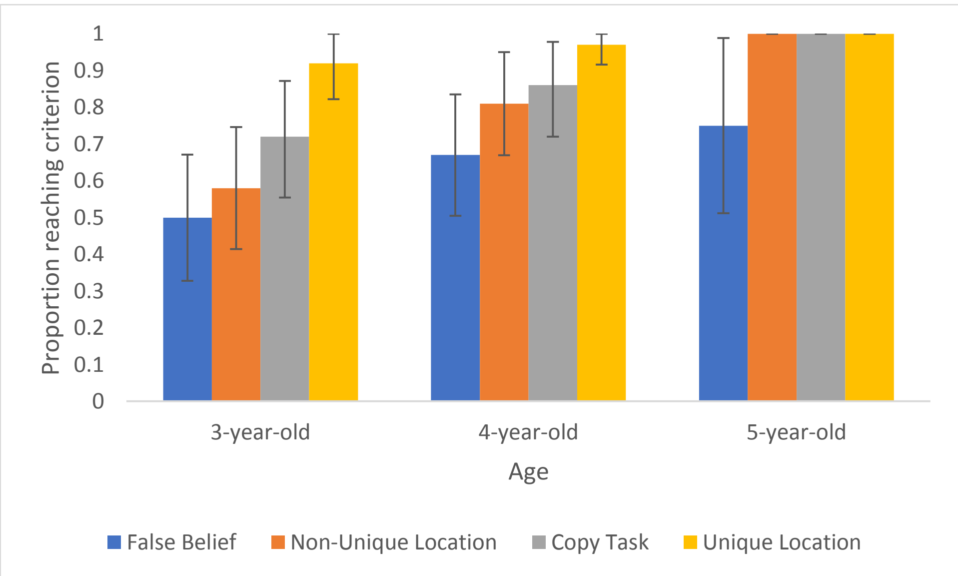
Both experiments were carried out in accordance with ethical standards at [blinded] Research Ethics Committee (2019-0010-001516), “Children's understanding of scale models.”. Experiment 1 was not preregistered.

Results

Preliminary analysis showed no gender differences on any measure in either experiment, so gender was not analysed further. Analyses were conducted using raw scores, except for between-task comparisons (where different chance baselines risk misleading findings). Performances were classified as pass or fail using the lenient criterion of the first integer above chance on their respective scales. For the Model Room task, the passing criterion was 2 out of 3 because, assuming children understand at least item-to-item correspondence, chance performance on Non-unique trials is 1 out of 3. For the Copy task the chance baseline is 1.33, so passing criterion was 2 out of 5 in each trial. For the False Belief task, younger children typically systematically fail rather than guess, so passing criterion was 1 of 2 trials correct. Figures 4 and 5 show pass/fail data; Tables 1 and 6 show raw scores.

Figure 4

Experiment 1 Mean Task Success by Age Group.



Note. Error bars represent 95% CI.

Table 1

Experiment 1 Mean (SD) Task Performance by Age Group.

Variable	3-year-olds	4-year-olds	5-year-olds	Total
Unique Location (max. 3)	2.75 (*.60*)	2.78 (*.48*)	3.00 (*.00*)	2.81 (*.50*)
Non-unique Location (max. 3)	1.69 (*.89*)	2.16 (*.93*)	2.88 (*.34*)	2.10 (*.93*)
Copy Task (max. 10)	6.44 (*3.14*)	8.22 (*2.69*)	9.81 (*.54*)	7.79 (*2.91*)
False Belief (max. 2)	.72 (*.81*)	1.16 (*.90*)	1.36 (*.89*)	1.02 (*.90*)

By inspection of Figure 4: performance on Unique trials approached ceiling; performances on the other three tasks show comparable improvement with age. Non-unique trial and Copy task performances were similar in magnitude and higher than performance on the False Belief tasks.

Model Room

Performance was higher when searching Unique than Non-unique locations, Wilcoxon's $T = 17.50$, $p<.01$, $r = .66$. Performances were significantly better than chance: for Unique trials, p = 1/6, t(88) = 50.26, $p<.01$, $d = .50$. For Non-unique trials, the chance baseline is either 1 in 6 if children guess at random, or 1 in 3 if children identify the furniture type and choose from exemplars at random: p = 1/6, t(88) = 19.69, $p<.01$, $d = .93$; p = 1/3, t(88) = 17.97, p<.01, $d = .93$. Most (76%) Non-unique trial errors reflected guesses associated with other non-unique chairs. Eight participants (*range* = 28-60m, $M = 48$m) failed one Retrieval Original question, three failed two (32m, 47m, 48m), and one failed four (48m). After the first trial, only 11% of incorrect searches were to the previous trial location, indicating no perseveration.

Copy Task

Performances on real (*M* = 3.90, *SD* = 1.50) and photograph (*M* = 3.89, *SD* = 1.66) exemplars did not differ, Wilcoxon *T* =414.00, *p* =.83, *r* = .02, and were closely associated (r_s =.64, p<.01), so were combined for analysis.

False Belief

Performances on task versions did not differ (M_1 = .53, M_2 = .49, McNemar, *p* =.65) and were closely related (r_s =.58, p<.01; controlling for age r_{partial} =.52, *p*<.01) so were combined for analysis.

Comparison of Tasks

Performances on all tasks were significantly associated with each other and age, as shown in Table 2, except for the unique trials of the Model Room task – which approached ceiling and showed minimal variance (*M* = 93.7%, *SD* = 0.17%). After controlling for age, False Belief performance was not associated with Non-unique trial nor Copy task performances. The correlation between Non-unique trial and Copy task performances was substantial and significant, and remained so when controlling for age.

Non-unique trials and Copy task performances did not differ, Binomial test, p =.143, and were significantly higher than False Belief performance, Binomial test, p =.036 and p =.001 respectively.

Table 2

Correlations Between Age, Non-unique Trials, Copy Task and False Belief in Experiment 1. Partial Correlations controlling for Age are shown Below the Diagonal.

Variable	**Non-unique Trials**	**Copy Task**	**False Belief**
Age (Months)	.58**	.61**	.29**
Non-unique Trials	-	.63**	.31**
Copy Task	.43**	-	.22*
False Belief	.18	.06	-

Note. Correlations are Spearman's Rho. *p<.05, **p<.01, *2-tailed.*

Regression analysis was used further to examine associations between task performances. Hierarchical binary logistic regression examined whether Non-unique trial and Copy task performances explained the variance in False Belief performances beyond that of age. Hierarchical linear regression examined whether Copy task performance explained the variance in Non-unique trial performance beyond that of age. For both regressions, all models were significant. Variation in false belief performance was best explained by the model with age as the sole predictor. Variation in Non-unique trial performance was significantly predicted by age and Copy task performance, with Copy task performance contributing significantly more to the explained variance than age alone. The regressions confirm that Non-unique trial and Copy task performances are specifically related beyond common effects of age, but false belief variation is explained by age alone.

Table 3

Summary of Hierarchical Binary Logistic Regression Analysis for Experiment 1 Variables Predicting False Belief

Variable	χ^2	$\Delta\chi^2$	-2LL	*B*	*SE*	Wald	OR	95% CI OR	
								LL	UL
Step 1	7.43**	7.43**	110.95						
Age Months				.07	.03	6.61**	1.07	1.02	1.13
Step 2	10.65**	3.22	107.73						
Age Months				.04	.03	1.17	1.04	.97	1.10
Non-unique				.05	.10	.201	1.05	.86	1.27
Copy Task				.45	.32	1.92	1.56	.83	2.93

Note. *$p < .05$, **$p < .01$, ***$p < .001$

Table 4

Summary of Hierarchical Regression Analysis for Experiment 1 Variables Predicting Non-Unique Location Score

Variable	F	R^2	ΔR^2	B	SE	t
Step 1	39.84***	.31	.31***			
Age Months				.56	.01	6.31***
Step 2	33.14***	.44	.12***			
Age Months				.32	.01	3.28***
Copy Task				.42	.03	4.30***

Note. *p < .05, **p < .01, ***p< .001

Discussion

To summarise:

1. We clearly replicate Blades and Cooke's (1994) findings that the Model Room task is substantially harder when hiding locations are identical, distinguished only by spatial position. Performance improves rapidly over the period between three to five years. Most errors are to another non-unique location. This strongly suggests younger children's very good performance on unique trials utilizes item-to-item rather than spatial correspondences.

2. Non-unique trial performance was better than False Belief task performance. They were nevertheless of similar magnitude, and significantly associated. However, contrary to hypothesis, this association did not persist when age was statistically controlled for. The hypothesis that contemporaneous success on the tasks indicates development of an understanding of complex correspondence is not supported.

3. The ability to create spatial correspondences in the relatively abstract Copy task developed over the age range tested. This ability is of similar magnitude to Non-unique trial performance, and the two were strongly associated beyond common associations with age.

Thus, the ability to detect, exploit, or arrange spatial correspondences is used in the Model Room task, and develops contemporaneously but is not specifically associated with false belief understanding. This does not support the claim that understanding of spatial correspondence is a component of representational understanding.

Previous studies frequently found that children search the previously correct location after the first trial (see Kuhlmeier, 2005, for review and investigation). This perseverative search is most common in 2-year-old participants, and may reflect inhibitory difficulties. We found no evidence of perseveration, possibly because most of our participants were older than two years. Errors on Unique trials were rare. Non-unique trials were never consecutive, and errors were usually to another exemplar from the same category of furniture as the target. This suggests that most participants were employing at least item-to-item correspondence in their search. Previous studies including Non-unique trials (Blades & Cooke, 1994; Loewenstein & Gentner, 2001) also found little tendency for perseverative search.

Experiment 2

A potential limitation of Experiment 1 is that the relation between the two spaces is not strongly emphasised. Failure to realise that one space is *supposed* to correspond to the other could lead children to neglect the relational information within each space. This would produce the observed difference, since Unique trials do not require this information whereas Non-unique trials do. We address this in two ways designed to clarify the representational nature of the relation.

Scale models are typically smaller than their referents, and absence of this characteristic may make children less likely to infer a representational relation between them.

For half of children the referent room was made twice the size of the model, and the purpose of the small room consistently and emphatically described as to show "where things are in this big room".

Participants

Participants were 86 children (50 female): 47 3-year-olds (M = 3;6, SD = 3.3m, range 3;0 to 3;11), and 39 4-year-olds (M = 4;4, SD = 3.8m, range 4;0 to 5;1) from predominantly working-class preschools in [blinded]. None took part in Experiment 1. Inclusion and exclusion criteria were as before. Data collection was interrupted on the 16th March 2020 when the UK government ordered non-essential contact to cease due to the Covid-19 pandemic. It was resumed on 7th March 2022 after restrictions had been lifted. The first phase included 44 participants; the second phase included 42 participants. The second phase of this experiment was preregistered to confirm the intention to finish the study as planned. See https://aspredicted.org/PRV_5CD. First phase data were analysed in 2020 before it became clear we would be able complete the study. Although we could present each phase as separate identical experiments with equivalent findings, we present it here as a single full-power study.

Design

Each child was tested in a quiet, familiar location for around 15 minutes in quasi-random order on the Model Room, Copy, and False Belief tasks. Participants were assigned to the standard or explicit representation version of the Model Room task using matched pairs. The first author collected all data in single sessions.

Materials and Procedure

The Copy task was identical to Experiment 1, using real exemplar grids for both trials. The False Belief task was the Lucy and Billy task used in Experiment 1.

Model Room Task

The standard task was identical to Experiment 1. For the explicit representation task, one model bedroom was twice the size of the other (60cm wide x 44cm deep x 28cm high, see Figure 1). Instructions emphasised the informative relation between the two rooms: “Here we have two rooms, and this small room shows us where things are in this big room”. Children were shown that each item in the small room showed where the corresponding item was in the big room, e.g., “The bed in the small room [points] shows us where the bed is in the big room [points] [...] We are going to play a hiding game with these stickers. The small room will show us where the sticker is in the big room. Every time you find a sticker you can keep it”. In six experimental trials, children were shown where the sticker was in the small room and told “This shows us where the sticker is in the big room.” The test questions were:

Retrieval Analogous: Can you find it in the big room?

Retrieval Original: Where’s the sticker in this room? [indicating the small room]

In all other respects procedure and scoring matched the standard task. Four children (*range* = 38-47m, M = 42.5) failed the familiarization trial and were given feedback. One child (42m) failed one retrieval trial.

Results

Model Room

Performance on the standard and explicit representation conditions did not differ (Mann-Whitney, Unique trials U = 834.5, p =.12, r = .17; Non-unique trials U = 809.5, p =.30, r = .11; Table 5). The two conditions were therefore combined for subsequent analysis.

Table 5

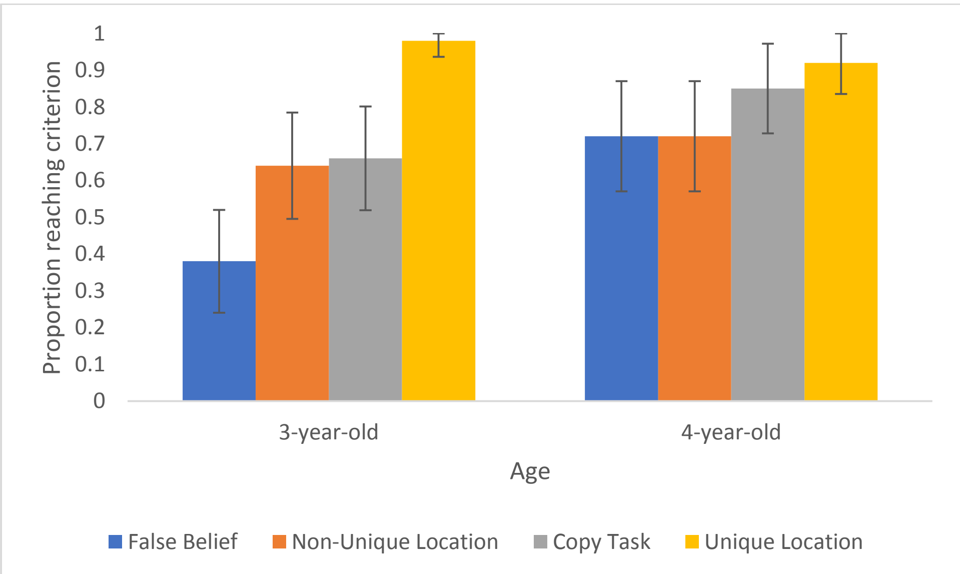
Mean (SD) for Trial Type and Condition by Age

Trial Type	Condition	3-year-olds	4-year-olds	Total
Unique	Standard	2.91 (*.41*)	2.60 (.75)	2.77 (*.61*)
	Explicit	2.96 (.20)	2.95 (*.23*)	2.95 (*.21*)
Non-unique	Standard	2.09 (*.95*)	1.95 (*1.15*)	2.02 (*1.04*)
	Explicit	1.79 (*.97*)	1.95 (*.78*)	1.86 (.89)

Performance was higher on Unique than Non-unique trials for both conditions (respectively: Wilcoxon, $T = 496$, $p<.01$, $r = .76$; $T = 253$, $p<.01$, $r = .65$). Performance was significantly better than chance across conditions and trial types (One-sample t-tests, all *ps* <.01, $d = .21 – 1.04$). Errors on Non-unique trials were at other non-unique locations 98% of the time. Incorrect searches were to the previous trial location only 1% of the time, indicating no perseveration. No errors were made on the Retrieval Original memory question.

Figure 5

Experiment 2 Mean Task Success by Age Group.



Note. Error Bars represent 95% CI.

Comparison of Tasks

Table 6 shows mean performances on all tasks. Unique trial performance approached ceiling and showed minimal variance (*M* = 95.3%, *SD* = 0.15%). Table 7 shows correlations between remaining variables. After controlling for age, again only associations between the Copy Task and Model Room Trials remain significant.

Table 6

Experiment 2 Mean (SD) Task Performance by Age Group.

Variable	3-year-olds	4-year-olds	Total
Unique Location (max. 3)	2.94 (*.32*)	2.77 (*.58*)	2.86 (*.46*)
Non-unique Location (max. 3)	1.94 (*.97*)	1.95 (*.97*)	1.94 (*.96*)
Copy Task (max. 10)	6.81 (*2.75*)	8.36 (*2.41*)	7.51 (*2.70*)
False Belief (max. 1)	.38 (*.49*)	.72 (*.46*)	.53 (*.50*)

Table 7

Correlations Between Age, Non-unique Trials of the Model Task, Copy Task, and False Belief in Experiment 2. Partial Correlations Controlling for Age are Shown Below the Diagonal.

Variable	**Non-unique Trials**	**Copy Task**	**False Belief**
Age (Months)	.078	.39**	.34**
Non-unique Trials	-	.30**	.02
Copy Task	.30**	-	.25*
False Belief	-.01	.14	-

Note. Correlations are Spearman's Rho. *p<.05, **p<.01, *2-tailed.*

To compare performance, success on Non-unique trials and the Copy task was classified as before. Non-unique trials and Copy task performances were very similar (67% and 74% success respectively). Copy task performance was significantly greater than False Belief performance (53% success), Binomial test, $p = .003$. Non-unique trial and False Belief performances did not significantly differ, Binomial test, $p = .08$.

As before, a hierarchical binary logistic regression shows the model with age as the only predictor best explained the variance in False Belief performance.

Table 8

Summary of Hierarchical Binary Logistic Regression Analysis for Experiment 2 Variables Predicting False Belief

Variable	χ^2	$\Delta\chi^2$	-2LL	*B*	*SE*	Wald	OR	95% CI OR	
								LL	UL
Step 1	11.51**	11.51**	107.29						
Age Months				.13	.04	9.73**	1.14	1.051.02	1.24
Step 2	13.47**	1.96	105.33						
Age Months				.11	.04	6.67**	1.12	1.03	1.22
Non-unique				.13	.10	1.92	1.143	.95	1.38
Copy Task				-.12	.26	.20	.89	.53	1.50

Note. **p* < .05, ***p* < .01, ****p*< .001

Hierarchical linear regression shows that both age and Copy task performance are significant predictors of non-unique performance, where Copy task contributes significantly more to the explained variance than age alone.

Table 9

Summary of Hierarchical Regression Analysis for Experiment 1 Variables Predicting Non-Unique Location Score

Variable	F	R^2	ΔR^2	B	SE	t
Step 1	.28	.003	.003			
Age Months				.06	.02	.53
Step 2	3.34*	.07	.07**			
Age Months				-.04	.02	-.38
Copy Task				.29	.04	2.52**

Note. *p < .05, **p < .01, ***p< .001

General Discussion

In both experiments, Unique trial performance on the Model Room task approached ceiling as expected. It was substantially higher than Non-unique trial performance. Performances on Non-unique trials of the Model Room task, the Copy task, and the False Belief task were of similar magnitude. Nevertheless, False Belief performance was lower than the other two, which did not differ. After accounting for general developmental change by controlling for age, Non-unique trial and Copy task performances were substantially and significantly associated. Neither was associated with False Belief performance.

In Experiment 2 the representational relation between model and room was emphasised by using a model smaller than the referent space, as is typical for scale model studies, and explicitly presenting the model's purpose as to show the hiding location in the referent space. This manipulation made no difference to observed outcomes.

We began with the hypothesis that the ability to use spatial information in the Model Room task indicates children now understand the model as a representation of the room. This was not a logical necessity, but was plausible because of the apparently contemporaneous success at other tasks measuring metarepresentational understanding. The results do not support this hypothesis. One should be careful making strong conclusions from null results. However, our findings are only null in respect of the False Belief task. The non-representational spatial Copy task was strongly associated with performance on Non-unique model room trials.

Taken together, these findings provide no evidence for a general understanding of correspondence arising alongside metarepresentational understanding. Nor do the results suggest the understanding of correspondence to be a precursor to metarepresentational understanding. Success rates on the False Belief task and Non-unique model trials were of similar magnitude. The modest differences could reflect task demands rather than differences in target competences, and such a close precursor would be expected to associate strongly.

These findings are consistent with previous studies using the Model Room task. As Lillard and Kavanaugh (2014) note, all prior studies administered the task in isolation, except for Walker and Murachver (2012), so little is known about its relation to other measures of associated processes. These two studies used Unique trials, which measure understanding of item-to-item correspondence. They found little relation to theory of mind development. The present study extends this finding to Non-unique trials, which measure spatial correspondence. None of the studies gives reason to believe that understanding spatial correspondence is related to metarepresentational development.

Implications for the Model Room task

The close relation between Identical trial and Copy task performances supports the alternative hypothesis that Model Room performance reflects spatial abilities. Based on prior literature and present results, these abilities can be divided thus: very young children can recognise the similarity of objects; children from the age of about three years can understand that two similar objects in the two spaces correspond; children from the age of about four years can understand that the two spaces themselves, in terms of their similar spatial layout, correspond. In other words, the developmental sequence is object similarity, item-to-item correspondence, and spatial correspondence.

From present data we cannot speculate about why the stages in the sequence emerge at the ages observed. Map reading abilities, reviewed above, may reflect a similar sequence. However, whether this sequence is common to scale models and maps and what determines this sequence are matters for further empirical study.

Conclusion

We examined a longstanding claim that understanding relational correspondence is a general component of representational understanding (Bach, 2014; Hoyos et al., 2020; Lillard & Kavanaugh, 2014; Perner, 1991; Walker & Murachver, 2012). We employed the well-known Model Room task (DeLoache, 1987), typically considered a test of representational understanding. Success on trials that involved matching targets associated with unique and non-unique items of furniture ‘across rooms’ was compared with performance on the False Belief task, a widely used measure of representational understanding, and the Copy task (Doherty & Anderson, 2003), a test of the ability to establish purely spatial correspondences. In total we had 175 participants between 2- and 5-years, the largest existing sample among studies employing the Model Room task. The findings were that the ability to exploit spatial correspondence in the task developed over the age range and performance was comparable to

Copy and False Belief task performances. After accounting for common effects of age, spatial correspondence ability associated substantially and significantly with the Copy task, and neither associated with the False Belief task. We conclude that the Model Room task measures spatial abilities, but find no evidence for representational abilities, which in turn does not support the claim that understanding of correspondence is a common factor in the development of metarepresentational ability.

References

Apple Dictionary Version 2.3.0

Bach, T. (2014). A Unified Account of General Learning Mechanisms and Theory- of- Mind Development. *Mind & Language*, *29(3)*, 351-381. https://doi.org/10.1111/mila.12055

Baldwin, D. A. and Saylor, M.M. (2005). Language promotes structural alignment in the acquisition of mentalistic concept. In: J. W. Astington and J. A. Baird (eds.) *Why Language Matters for Theory of Mind*. New York: Oxford University Press.

Blades, M., & Cooke, Z. (1994). Young children's ability to understand a model as a spatial representation. *The Journal of Genetic Psychology*, *155*(2), 201-218. https://doi.org/10.1080/00221325.1994.9914772

Bluestein, N., & Acredolo, L. (1979). Developmental changes in map-reading skills. *Child Development*, *50*, 691-697. https://doi.org/10.2307/1128934

Callaghan, T. (2020). The origins and development of a symbolic mind: the case of pictorial symbols. *Interchange*, *51*(1), 53-64. https://doi.org/10.1007/s10780-020-09396-z

Childers, J. B., Parrish, R., Olson, C. V., Burch, C., Fung, G., & McIntyre, K. P. (2016). Early verb learning: How do children learn how to compare events? Journal of Cognition and Development, 17(1), 41–66. https://doi.org/10.1080/15248372.2015.1042580

Degani, A. (2013) A tale of two maps: Analysis of the london Underground "diagram". *Ergonomics in Design*, *21*, 7-16.

DeLoache, J. S. (1987). Rapid change in the symbolic functioning of very young children. *Science*, *238*(4833), 1556-1557. https://doi.org/10.1126/science.2446392

Dillon, M. R., & Spelke, E. S. (2018). From map reading to geometric intuitions. Developmental Psychology, 54(7), 1304–1316. https://doi.org/10.1037/dev0000509

Dillon, M. R., Huang, Y., & Spelke, E. S. (2013). Core foundations of abstract geometry. Proceedings of the National Academy of Sciences, 110(35), 14191-14195. https://doi.org/10.1073/pnas.1312640110

Doherty, M. J. (2000). Children's understanding of homonymy: Metalinguistic awareness and false belief. *Journal of Child Language*, *27(2)*, 367-392. https://doi.org/10.1017/S0305000900004153

Doherty, M.J. (2009). *Theory of Mind*. Psychology Press.

Doherty, M.J., & Anderson, J. (2003). Children's ability to copy spatial arrays. Poster presented at the BPS Developmental Section Annual Conference, University of Coventry, September 2003.

Doherty, M., & Perner, J. (1998). Metalinguistic awareness and theory of mind: Just two words for the same thing? *Cognitive Development*, *13(3)*, 279-305. https://doi.org/10.1016/S0885-2014(98)90012-0

Doherty, M. J., & Perner, J. (2020). Mental files: Developmental integration of dual naming and theory of mind. *Developmental Review*, *56*, 100909. https://doi.org/10.1016/j.dr.2020.100909

Doherty, M. J., & Wimmer, M. C. (2005). Children's understanding of ambiguous figures: Which cognitive developments are necessary to experience reversal? *Cognitive Development*, *20(3)*, 407-421. https://doi.org/10.1016/j.cogdev.2005.05.003

Doherty, M. J., Wimmer, M. C., Gollek, C., Stone, C., & Robinson, E. J. (2021). Piecing together the puzzle of pictorial representation: How jigsaw puzzles index metacognitive development. *Child Development*, *92*(1), 205-221. https://doi.org/10.1111/cdev.13391

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G-Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191. https:// doi.org/10.3758/BF03193146

Gentner, D. (2005). The development of relational category knowledge. In:

Gershkoff-Stowe, L., & Rakison, D. H. (Eds.). (2005). *Building object categories in developmental time* (pp. 263-294). Psychology Press.

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science*, *34(5)*, 752-775. https://doi.org/10.1111/j.1551-6709.2010.01114.x

Gollek, C., & Doherty, M. J. (2016). Metacognitive developments in word learning: Mutual exclusivity and theory of mind. *Journal of Experimental Child Psychology*, *148*, 51-69. https://doi.org/10.1016/j.jecp.2016.03.007

Hoyos, C., Horton, W. S., Simms, N. K., & Gentner, D. (2020). Analogical comparison promotes theory- of- mind development. *Cognitive Science, 44(9)*, e12891. https://doi.org/10.1111/cogs.12891

Huttenlocher, J., Newcombe, N., & Vasilyeva, M. (1999). Spatial scaling in young children. Psychological Science, 10(5), 393-398. https://doi.org/10.1111/1467-9280.00175

Jastrow, J. (1900). *Fact and fable in psychology*. Boston: Houghton Mifflin.

Karadaki, T. (2019). Children's referent selection and theory of mind. Unpublished doctoral dissertation, School of Psychology, University of East Anglia.

Karadaki, T., & Doherty, M. J. (2017, May). Preschool metacognitive developments allow learning and use of co-referential words. In Poster presented at the IX. Dubrovnik Conference on Cognitive Science, Dubrovnik, Croatia.

Kuhlmeier, V. (2005). Symbolic insight and inhibitory control: Two problems facing young children on symbolic retrieval tasks. *Journal of Cognition and Development*, *6(3)*, 365-380. https://doi.org/10.1207/s15327647jcd0603_3

Liben, L. S., Myers, L. J., & Kastens, K. A. (2008, September). Locating oneself on a map in relation to person qualities and map characteristics. In *International conference on spatial cognition* (pp. 171-187). Springer, Berlin, Heidelberg.

Liben, L. S., & Yekel, C. A. (1996). Preschoolers' understanding of plan and oblique maps: The role of geometric and representational correspondence. Child Development, 67(6), 2780–2796. https://doi.org/10.2307/1131752

Lillard, A. S., & Kavanaugh, R. D. (2014). The contribution of symbolic skills to the development of an explicit theory of mind. *Child Development*, *85*(4), 1535-1551. https://doi.org/10.1111/cdev.12227

Loewenstein, J., & Gentner, D. (2001). Spatial mapping in preschoolers: Close comparisons facilitate far mappings. *Journal of Cognition and Development*, *2(2),* 189-219. https://doi.org/10.1207/S15327647JCD0202_4

Perner, J. (1991). *Understanding the Representational Mind*. The MIT Press.

Perner, J., Stummer, S., Sprung, M., & Doherty, M. (2002). Theory of mind finds its Piagetian perspective: Why alternative naming comes with understanding belief. *Cognitive Development*, *17*(3-4), 1451-1472. https://doi.org/10.1016/S0885-2014(02)00127-2

Rasch, D., & Guiard, V. (2004). The robustness of parametric statistical methods. *Psychology Science*, *46*, 175-208.

Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology*, *94(3)*, 249-273. https://doi.org/10.1016/j.jecp.2006.02.002

Salsa, A., Gariboldi, M. B., Vivaldi, R., & Rodríguez, J. (2019). Geometric maps as tools for different purposes in early childhood. *Journal of Experimental Child Psychology*, *186*, 33-44. https://doi.org/10.1016/j.jecp.2019.05.004

Shusterman, A., Ah Lee, S., & Spelke, E. S. (2008). Young children's spontaneous use of geometry in maps. Developmental science, 11(2), F1-F7. https://doi.org/10.1111/j.1467-7687.2007.00670.x

Troseth, G. L., Bloom Pickard, M. E., & DeLoache, J. S. (2007). Young children's use of scale models: Testing an alternative to representational insight. *Developmental Science*, *10(6)*, 763-769. https://doi.org/10.1111/j.1467-7687.2007.00625.x

Troseth, G. L., & DeLoache, J. S. (1998). The medium can obscure the message: Young children's understanding of video. *Child Development*, *69*(4), 950-965. https://doi.org/10.2307/1132355

Yuan, L., Uttal, D., & Gentner, D. (2017). Analogical processes in children's understanding of spatial representations. *Developmental Psychology*, *53(6)*, 1098–1114. https://doi.org/10.1037/dev0000302

Walker, R. F., & Murachver, T. (2012). Representation and theory of mind development. *Developmental Psychology*, *48*, 509–520. https://doi.org/10.1037/a0025663

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. https://doi.org/10.1111/1467-8624.00304

Wimmer, M. C., Doherty, M. J., & Collins, W. A. (2011). The development of ambiguous figure perception. *Monographs of the Society for Research in Child Development*, i-130. https://doi.org.uk/10.1111/j.1540-5834.2011.00589.x

Winkler- Rhoades, N., Carey, S. C., & Spelke, E. S. (2013). Two- year- old children interpret abstract, purely geometric maps. Developmental science, 16(3), 365-376. https://doi.org/10.1111/desc.12038