# Universiteit Leiden
## The Netherlands

**Self-projection in early childhood: no evidence for a common underpinning of episodic memory, episodic future thinking, theory of mind, and spatial navigation**

Immel, A.-S.; Altgassen, M; Meyer, M.; Endedijk, H.M.; Hunnius, S.

# Self-projection in early childhood: No evidence for a common underpinning of episodic memory, episodic future thinking, theory of mind, and spatial navigation

A-S. Immel [a],[*], M. Altgassen [b], M. Meyer [a], H.M. Endedijk [c], S. Hunnius [a]

[a] *Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, 6525 GD Nijmegen, The Netherlands*
[b] *Department of Psychology, Johannes Gutenberg University Mainz, 55122 Mainz, Germany*
[c] *Educational Science, Leiden University, 2300 RA Leiden, The Netherlands*

## ARTICLE INFO

## ABSTRACT

Buckner and Carroll [*Trends in Cognitive Sciences* (2007), Vol. 11, pp. 49–57] argued that episodic memory (EM), episodic future thinking (EFT), theory of mind (ToM), and spatial navigation all build on the same mental mechanism—self-projection, that is, the ability to disengage from the immediate present and shift perspective to alternative temporal, mental, or spatial situations. Developmental studies indeed show that all four abilities undergo profound developmental changes around 4 years of age, and there are first indications of developmental interrelations between some of the abilities. However, strong evidence for the self-projection account, namely that *all four* abilities are interrelated in their emergence during early childhood, is still lacking. To thoroughly investigate the self-projection hypothesis, we tested 151 4-year-old children on 12 different tasks assessing their EM, EFT, ToM, and spatial navigation skills (3 tasks per ability). Structural equation modeling under maximum likelihood estimation was conducted on a final sample of 144 children to evaluate a model with the 12 tasks as indicators and self-projection as the latent factor. The model showed a very good fit to the data. However, the factor loadings, indicating the strength of association between the latent factor and the indicators, were very low—which speaks against the validity of the measurement model. In summary, the results do not support the assumption of a common latent factor underlying the

* Corresponding author.

*E-mail address:* anna-sophie.immel@posteo.de (A-S. Immel).

various abilities EM, EFT, ToM, and spatial navigation. Implications of our results for the self-projection account and possible related theoretical and methodological challenges are discussed.

## Introduction

What do the various activities of remembering last year's summer holidays, envisioning a picnic upcoming weekend, understanding your friend's perspective, and finding your way around the city center have in common? Whereas one might be inclined to spontaneously answer this question with "very little," recent findings suggest that the related abilities episodic memory (EM), episodic future thinking (EFT), theory of mind (ToM), and spatial navigation might share much more than previously assumed. First, several developmental studies show that all four abilities undergo profound developmental changes around the same time during early childhood, namely around 4 years of age (EM and EFT: Atance, Louw, & Clayton, 2015; Atance & O'Neill, 2005; Prabhakar & Hudson, 2019; Suddendorf, Nielsen, & Von Gehlen, 2011; ToM: Frith & Frith, 2005; Wellman & Liu, 2004; spatial navigation: Nardini, Burgess, Breckenridge, & Atkinson, 2006). Furthermore, individual difference approaches have demonstrated developmental interrelations between some of the skills, for instance, between EM and ToM (Perner, Kloo, & Gornik, 2007), between EM and EFT (Busby & Suddendorf, 2005; Hayne, Gross, McNamee, Fitzgibbon, & Tustin, 2011; Perner et al., 2010), and between ToM and the EFT-related capacity prospective memory (Causey & Bjorklund, 2014; Ford, Driscoll, Shum, & Macaulay, 2012; Kretschmer-Trendowicz, Ellis, & Altgassen, 2016). Finally, a range of neuroimaging studies with adults have shown that a strikingly similar core neural network is activated when being engaged in tasks related to these different abilities (e.g., Addis, Wong, & Schacter, 2007; Andrews-Hanna, Saxe, & Yarkoni, 2014; Spreng & Grady, 2010; Spreng, Mar, & Kim, 2009; Szpunar, Watson, & McDermott, 2007). Together, these findings speak in favor of the self-projection account, an integrative approach proposed by Buckner and Carroll (2007; for similar accounts, see also Hassabis & Maguire, 2007, 2009; Mahr, 2020; Suddendorf & Corballis, 1997, 2007). Buckner and Carroll (2007) argued that EM, EFT, ToM, and spatial navigation, which have traditionally been studied in isolation, all build on the same cognitive mechanism—self-projection, which they define as the ability to disengage from the immediate present and to shift perspective to alternative temporal, mental, or spatial situations. All four abilities suggested to be related to self-projection are crucial milestones in children's cognitive and social-cognitive development. Moreover, a growing body of adult and child research indicates that individuals with autism spectrum disorder (ASD) show deficits in all four abilities (evidence from adult studies—EM: Crane & Goddard, 2008; EM and EFT: Lind, Williams, Bowler, & Peel, 2014; ToM: Spek, Scholte, & Van Berckelaer-Onnes, 2010; evidence from child studies—EFT: Terrett et al., 2013; EM, EFT, ToM, and spatial navigation: Lind, Bowler, & Raber, 2014). The proposal of a common mechanism underlying these diverse abilities therefore not only has important implications for a better understanding of crucial hallmarks of human cognition and behavior but also could, if the account holds, critically inform programs fostering the development of the four abilities in typically and atypically developing children. Despite this importance, to date nearly all the available evidence on the interrelation of the four abilities solely comprises data on associations between a subset of the target constructs. However, strong evidence for the self-projection account, namely that *all four* abilities are interrelated during the time in development when they are beginning to emerge, is still lacking. Therefore, by investigating individual developmental differences in a large sample of 4-year-olds, this study set out to close this gap and to inquire whether the seemingly diverse abilities of EM, EFT, ToM, and spatial navigation arise in a related fashion during early childhood.

In the following, we give a short overview of the four abilities, their proposed link to self-projection, and developmental research investigating the interrelations among the abilities. Each section is concluded with a brief outline of the approach and hypotheses of the current study.

*Episodic memory and episodic future thinking*

Tulving (1972, 1985a, 1985b) defined episodic memory as the mental recollection of autobiographical experiences, distinguishing it from semantic memory, which refers to the recall of factual knowledge. He argued that a central aspect of EM is the subjective awareness (also called autonoetic consciousness) of the recollected events (Tulving, 1985b; see also Wheeler, Stuss, & Tulving, 1997). This implies a reliving or reexperiencing of the events as having happened to oneself, leading to an active and vivid remembering in contrast to mere knowledge about their occurrence (Dafni-Merom & Arzy, 2020; Perner, Kloo, & Gornik, 2007; Tulving, 1985b, 1987).

The future-oriented counterpart of EM, episodic future thinking, refers to the construction of mental representations of future events (e.g., Atance & O'Neill, 2001; Schacter, Benoit, & Szpunar, 2017). In parallel with EM but in contrast with other forms of future thinking such as planning and prediction, the "episodic" in EFT underlines that it regards a specific category of mental simulations, namely those comprising autobiographical and thereby self-referential aspects (Szpunar, Spreng, & Schacter, 2014). Furthermore, it has been proposed that EFT critically builds on EM because it is likely that the retrieval and recombination of past experiences allow the creation of future-directed mental simulations and time travel (e.g., Klein, 2016; Schacter & Addis, 2007).

Indeed, mentally traveling back (EM) or forward (EFT) in time, or in other words projecting oneself into a subjective time other than the current one, matches well Buckner and Carroll's (2007) description of self-projection, defined as the mental disengagement of the self from the here and now into an alternative perspective (temporal in the case of EM and EFT). Next to the shift of viewpoint, it is the pronounced self-referential aspect encompassed by both constructs that suits well the self-projection framework: Not any event is thought to be recalled or anticipated when engaging in EM and EFT but specifically those events with a link to autobiographical elements and subjective experiences. Developmental support for the assumption of a common underlying cognitive mechanism of the two abilities EM and EFT comes from research showing that both abilities undergo substantial changes during the preschool years (Atance et al., 2015; Atance & Meltzoff, 2005; Busby & Suddendorf, 2005; Prabhakar & Hudson, 2019) and are closely associated in development (Atance & Sommerville, 2014; Busby & Suddendorf, 2005; Prabhakar & Hudson, 2019; Quon & Atance, 2010; Richmond & Pan, 2013; Ünal & Hohenberger, 2017). Atance and Sommerville (2014), for instance, assessed 3- to 5-year-olds' EFT abilities with a set of item selection tasks in which children were confronted with a problem in one room (e.g., a locked box) and, after a short delay, were presented in a second room with four items (e.g., a comb, tape, a crayon, and a key), one of which could solve the previously encountered problem. After item selection, children's EM abilities were assessed by asking children about their memory of the problem. Results showed that children's proportion of correct item choices were significantly correlated with the proportion of correct responses to the memory question, also after controlling for age. In line with the self-projection hypothesis, the authors concluded that both abilities "likely draw on overlapping although not identical processes" (p. 124).

*Theory of mind*

Another ability suggested to depend on self-projection is theory of mind, which is defined as the ability to attribute mental states like intentions, beliefs, and emotions to others and the ability to infer their actions based on these attributions (Premack & Woodruff, 1978; Wimmer & Perner, 1983). This social-cognitive skill is considered to be crucial for successful interpersonal interactions (e.g., Astington & Jenkins, 1995). It has been argued that in order to understand another agent's mental states, one needs to create a mental simulation of the other person's emotions, desires, or beliefs (Goldman, 2006; Wu, Liu, Hagan, & Mobbs, 2020). In accordance with the self-projection account, it is plausible to assume that this simulation requires a mental detachment from one's own current point of view and a simultaneous projection into an alternative perspective, in this case the one of another agent. Similarly, with regard to ToM, research has confirmed an important developmental shift around 4 years of age (Wellman & Liu, 2004). Furthermore, it has been proposed that ToM and EM are intertwined in development (Perner, 2000), and studies with 3- to 6-year-olds do indeed suggest a developmental interrelation between the two abilities (Perner et al., 2007), although not always

consistently across all age groups (Naito, 2003). In one of these studies, Perner et al. (2007) tested preschoolers on a ToM test battery and on two memory tasks, one of which involved the direct experience of events (actively looking at self-picked cards) and the other of which provided the events indirectly (watching a video about what was depicted on the cards). The latter manipulation was introduced in order to tease apart the specific influence of EM—that is, vividly remembering an event (direct experience memory task) from mere knowledge of an event (indirect information memory task). Results yielded that children with high ToM competence remembered the items from the direct experience memory task better than the ones from the indirect information memory task. This points to an association between ToM and EM and corroborates the assumption of the self-projection account that self-referential processes are essential to the relationship between the abilities.

Similarly, a developmental interrelation between ToM and EFT has been suggested (Ford et al., 2012), but the empirical evidence for this association is mixed. Hanson, Atance, and Paluck and colleagues (2014), in testing 3- to 5-year-olds on several ToM and EFT tasks, found no association between both constructs after controlling for age. However, developmental research on prospective memory (an EFT-related ability) and ToM suggests a link between the two skills (Causey & Bjorklund, 2014; Ford et al., 2012; Kretschmer-Trendowicz et al., 2016).

## Spatial navigation

Successful spatial navigation, also called "mental space travel" (e.g., Adornetti et al., 2021) is supported by the use of egocentric and allocentric strategies (e.g., Burgess, 2006; Galati, Pelle, Berthoz, & Committeri, 2010; O'Keefe & Nadel, 1978; Ruggiero, D'Errico, & Iachini, 2016). An egocentric frame of reference is based on person-centered representations of the environment that relate the surroundings to oneself. An allocentric frame of reference, on the other hand, is built on representations of the environment in which the location of objects is expressed with relation to external references like landmarks. The latter reference system allows a flexible retrieval of the environment because it is vantage point independent. Specifically, this allocentric strategy has been proposed to be related to self-projection because it possibly involves a mental detachment from the current spatial location and shift to an alternative (spatial) perspective (Burgess, 2008; McNamara, Rump, & Werner, 2003; Spreng et al., 2009). Classic accounts suggest that the use of egocentric strategies is already present from an early age and foregoes the use of allocentric strategies (e.g., Acredolo, 1978; Piaget & Inhelder, 1967). The employment of an allocentric frame of reference emerges from 3 to 5 years of age (Nardini et al., 2006; van Hoogmoed, 2014), and its use gradually increases and becomes more efficient during middle childhood (Bullens, Iglói, Berthoz, Postma, & Rondi-Reig, 2010; Yang, Merrill, & Wang, 2019). With the exception of a clinical study that investigated possible links between spatial navigation and EFT, EM, and ToM in children with ASD (Lind, Bowler, et al., 2014), there is no developmental literature that included spatial navigation in design when exploring the self-projection hypothesis. In this clinical study (Lind, Bowler, et al., 2014), children with ASD (mean age ∼8 years) and typically developing children (control group) were tested on a computerized spatial navigation task. First, they needed to navigate through a virtual island and encode the positions of objects that were indicated by flags. Then, on subsequent trials, they were asked to navigate through the same environment, but this time they needed to (re)find their way to the objects without any indicators (i.e., the objects were now hidden). Furthermore, children were assessed on their EM, EFT, and ToM abilities by providing short verbal narratives of past and potential future events (EM and EFT, respectively) and by describing the interactions of moving triangles (ToM). Results indicated that in the clinical group, children showed impairments in their spatial navigation, EM, and EFT performance in comparison with the control group, which the authors interpreted as speaking in favor of the self-projection account, although there were no correlations between the different tasks for either group (Lind, Bowler, et al., 2014). Importantly, it also remains an open question whether the detected group differences are specific to self-projection or reflect more general differences between the clinical and control groups. Lind, Bowler, et al.'s (2014) study therefore provides only mixed evidence for Buckner and Carroll's (2007) hypothesis. Still, that spatial navigation skills and skills in EM, EFT, and ToM share a similar developmental trajectory supports the assumption of a common origin of all four abilities.

*Investigations of the self-projection hypothesis including all four abilities*

In addition to the evidence from the developmental literature, there are clinical and neuroimaging studies with adults corroborating the self-projection account and the notion of a shared underlying neural network (e.g., Addis et al., 2007; Addis, Wong, & Schacter, 2008; Hassabis, Kumaran, & Maguire, 2007; Klein, Loftus, & Kihlstrom, 2002; Kurczek et al., 2015; Okuda et al., 2003). However, even when taking this body of research into consideration, to date there are only three studies that have addressed the self-projection hypothesis by investigating the interrelations among *all four* abilities: a quantitative meta-analysis of neuroimaging studies with adults (Spreng et al., 2009) and two clinical studies with adults and the earlier described study on children with ASD (Lind, Bowler, et al., 2014; Lind, Williams, Raber, Peel, & Bowler, 2013). The meta-analysis provides quantitative evidence for a core neural network being shared by EM, EFT, ToM, and spatial navigation (Spreng et al., 2009). Specifically, across all four target abilities, conjunction analyses revealed a high level of corresponding activity in key brain regions related to the neural network that has been suggested to underlie self-projection (Buckner & Carroll, 2007), namely in the lateral prefrontal cortex, medial-temporal lobe, posterior cingulate cortex, and temporo-parietal junction (Spreng et al., 2009). Lind and colleagues' empirical investigation of the self-projection account, however, yielded mixed results. The adult study as well as the developmental study showed that taken across the two studies, individuals with ASD displayed deficits in all the investigated abilities (adults displayed deficits in EM, ToM, and spatial navigation but not EFT; children showed deficits in EM, EFT, and spatial navigation but not ToM) which altogether speaks in favor of the self-projection account (Lind, Bowler, et al., 2014; Lind et al., 2013). Although, as reported above, when investigating the interrelations among the four abilities in primary school children, no associations between the tasks were found (Lind, Bowler, et al., 2014). However, it could be argued that possible interrelations among the four abilities may have gone undetected due to a lack of statistical power given that each group in this study consisted of only 20 participants.

To the best of our knowledge, there is no study that has tested the self-projection account by investigating the interrelations among all four abilities in a large sample at a time when these abilities first emerge, that is, around 4 years of age. However, showing that all four abilities are related on an interindividual level during early childhood would provide strong evidence for the assumption of a common mechanism underlying these diverse cognitive and social-cognitive key capacities of human cognition.

*Goal and approach of the current study*

To examine the self-projection hypothesis in a thorough manner, we set out to test a large sample of 4-year-olds on a range of EM, EFT, ToM, and spatial navigation tasks. Each ability was assessed by three tasks, allowing us to comprehensively capture each construct. To test the self-projection hypothesis, we planned to use a multidimensional latent factor approach with the four abilities as the first-order factor and to use self-projection as the second-order factor. Because the target abilities presumably were related to general cognitive development, we also assessed verbal ability and reasoning skills and used these measures as additional predictors in our model. If there is indeed a shared cognitive mechanism (i.e., self-projection) underlying the diverse abilities EM, EFT, ToM, and spatial navigation, we would expect to identify a common latent factor being associated with all four abilities.

**Method**

*Participants*

The final sample on which the analyses are based consisted of 144 4-year-old children (mean age = 54.6 months, *SD* = 2.3, range = 50–59; 84 female) recruited through and tested at primary schools in The Netherlands. All children were Dutch native speakers. An additional 7 children were tested but excluded from the analyses due to missing data on more than seven tasks (*n* = 5) or to not being native speakers (*n* = 2). We had a complete data set—that is, data for all 12 tasks and the 2 control variables (see Table 1)—for 90 participants. From the remaining 54 children, there were on average 1.83 tasks

**Table 1**
Task descriptions.

| | Task | Description |
|---|---|---|
| EM | Treasure Box (Kastner-Koller & Deimann, 2002) | Children needed to learn the position of six items hidden in a treasure box and retrieve them. |
| | Yesterday (Busby & Suddendorf, 2005) | Children were asked to name three specific activities they had done the day before at school. |
| | Cartoon Recognition (Nigro, Brandimonte, Cicogna, & Cosenza, 2014) | Children needed to judge whether scenes had been part of a cartoon they had watched. |
| EFT | Picture Book (Atance & Meltzoff, 2005) | In preparation for an imaginary trip, children needed to choose one item they wanted to take with them and explain their choice. |
| | Tomorrow (Busby & Suddendorf, 2005) | Children were asked to name three specific activities they planned or expected to do the next day at school. |
| | Item Selection (Atance & Sommerville, 2014) | Being in the experimental room, children were confronted with a locked treasure box; after a distraction task on the corridor, they needed to choose one of four items (key, sharpener, color pencil, or comb) that they wanted to take with them upon going back to the experimental room. |
| ToM | False Belief Location (three instances: "Max," "Sally," "Heidi") (e.g., Wimmer & Perner, 1983) | After having observed the displacement of an object in a play, children needed to indicate where the protagonist, who had not observed the object transfer, would search for it. |
| | False Belief Content (three instances: "Milk Package," "Chocolate Sprinkles," "Crayon Box") (e.g., Perner, Leekam, & Wimmer, 1987) | After having seen the unexpected content of a box, children needed to indicate what another person or puppet, who had not seen its content, would think was in the box. |
| | Animated Shape (Abell, Happé, & Frith, 2000) | Children needed to describe and interpret the movements of triangle shapes. |
| SpNa | Board & Cup (Nardini, Burgess, Breckenridge, & Atkinson, 2006) | Children needed to encode an item's position on a large board and retrieve it after having moved around the board. |
| | Map (Shusterman, Lee, & Spelke, 2008) | On a map depicting three circles, children were shown the position where a little toy frog wanted to sit; they then were asked to turn around and put the frog in one of the buckets placed in front of them, mirroring on the floor the spatial arrangement of the circles seen on the map. |
| | Turning Table (Lambrey, Doeller, Berthoz, & Burgess, 2012; Wang & Spelke, 2002) | Children needed to indicate which item had changed position when presenting the scenery from a different perspective. |
| CV | Matrix Reasoning (Wechsler, 2002) | Children needed to complete a matrix of three items with a logically suitable fourth item. |
| | Vocabulary (Wechsler, 2002) | Children needed to explain the meanings of words read out by the experimenter. |

*Note.* Different instances of the False Belief Location and False Belief Content tasks were tested on each of the 3 testing days (see Appendix A). EM, episodic memory; EFT, episodic future thinking; ToM, theory of mind; SpNa, spatial navigation; CV, control variable.

missing ($SD = 1.31$). These children were also included in the analysis. Written informed consent from the parents and the heads of schools was acquired prior to testing. The study was approved by the local ethics committee. As a thank you, each participating class received a gift chosen by the teacher (e.g., a game or book for the class). In addition, at the end of Testing Day 1 participating children could keep the little bouncing ball that they found as a "treasure" in one of the tasks, and at the end of Testing Day 2 children could choose a coloring postcard.

*Procedure*

Children were tested individually by a female experimenter in a quiet room at their primary school. Every child was tested on 12 different tasks investigating their EM, EFT, ToM, and spatial navigation

skills. Furthermore, children were tested on reasoning and verbal ability with two subtests (Matrix Reasoning and Vocabulary) of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-III-NL; Wechsler, 2002). Testing was spread over three testing sessions taking place on separate days and lasting 35 to 45 min each. The three testing sessions were completed within a maximum period of 2 weeks. We carefully piloted all tasks prior to testing to ensure that they suited our age group and that setup, material, and administration motivated children to engage in the tasks. All tasks had been used in previous studies, and whenever exact measures and detailed illustrations of the stimulus material from the original study were available, we set up the task as closely as possible in accordance with these descriptions and the original protocol. Even though this was the case for the majority of the tasks, piloting revealed that some of the paradigms needed to be slightly adapted to our age range or the specific testing context (e.g., if original material was not available in Dutch). Table 1 contains a short description of each paradigm and its source. A more detailed description of the tasks and our adaptations can be found in Appendix A.

The order of tasks was determined based on considerations of children's motivation and attention span and was piloted prior to testing. We took into account, for instance, the length of the task, the level of activity (tasks administered on a table vs. those executed standing and walking), verbal demands (predominantly verbal tasks vs. those not solely based on verbal interaction), and whether screen time was involved. Piloting showed that children stayed motivated and engaged when task characteristics were varied during the testing session. Therefore, we tried to have, for instance, short and long tasks as well as tasks administered on the table and those that required whole body movements in alternation. In this way, we could reduce fatigue due to long periods of sitting or focusing on only one kind of task. An additional factor that we needed to take into consideration was of a practical nature. Some tasks, such as the Board & Cup task, required a longer setup and therefore were placed at the beginning of the testing session. The same order of tasks was used for all children and was as follows: Testing Day 1: Board & Cup, False Belief Location "Max," Tomorrow, Item Selection, False Belief Content "Milk Package"; Testing Day 2: False Belief Content "Chocolate Sprinkles," Picture Book, Cartoon Recognition, Matrix Reasoning, Map, False Belief Location "Sally," Yesterday; Testing Day 3: False Belief Content "Crayon Box," Treasure Box, Animated Shape, Vocabulary, False Belief Location "Heidi," Turning Table. However, the order of the Yesterday and Tomorrow tasks needed to be reversed for some children if the respective testing day (Monday or Friday) did not allow asking what had happened the day before at school or what would happen the day after at school, respectively.

Testing was conducted by 15 female experimenters. Completing in total about 450 testing sessions (3 sessions per child) within a reasonable time span was possible only by testing at different locations in parallel. To that end, a larger group of experimenters was indispensable. Experimenters were bachelor's and master's students. All experimenters were trained on all tasks, so that a child could be tested by the same experimenter on all 3 testing days. To ensure that the administration of tasks was equivalent for each participant, detailed protocols were developed and experimenters needed to complete extensive training prior to testing. First, experimenters were trained with video material showing the exact administration of the tasks. Subsequently, under the supervision of the main experimenter, several practical training sessions were conducted. Experimenters were allowed to test children only if they mastered the administration of all tasks flawlessly. Furthermore, testing sessions were audio-recorded. To ensure comparability of the testing procedure in terms of wording and motivational cues, three of the recorded sessions of each experimenter were evaluated by the main experimenter. This evaluation yielded that all experimenters had administered the tasks competently and correctly and in accordance with the protocols.

# Results

Models were fitted using the package Lavaan (Rosseel, 2012) in R (R Development Core Team, 2016). All other analyses were conducted with SPSS Version 25.0 (IBM Corp., Armonk, NY, USA).

Prior to model fitting, variables were checked for patterns of missingness and normality. According to Little's MCAR (missing completely at random) test, the missing data of all 14 variables (i.e., the 12 experimental tasks and the 2 control tasks) did not show evidence for structure in the missing data

and thus appear random, $\chi^2(161) = 163.17$, $p = .44$. The Shapiro–Wilk test for normality indicated that only the data of the Vocabulary task were normally distributed (see Appendix B, Table B1, for more details). The False Belief Content and False Belief Location tasks showed a U-shaped distribution pointing toward dichotomization and accordingly were recoded into binary dummy variables by combining the original scores 0 and 1 into the category "failed" (score of 0) and combining the original scores 2 and 3 into the category "passed" (score of 1). All other variables were log-transformed. However, also after log transformation, normality was not reached. Therefore, we decided to use the original non-transformed data of these tasks and to account for the non-normal distribution during model fitting with robust test statistics. Hence, model fit of all models was assessed using maximum likelihood estimation with robust (Huber–White) standard errors and a scaled test statistic that is (asymptotically) equal to the Yuan–Bentler test statistics (MLR in Lavaan).

*Descriptive statistics and correlations*

Descriptive statistics of all 12 experimental tasks and the 2 control tasks are provided in Table 2. First, relations between all 12 tasks were analyzed using Pearson correlations. Results are shown in Table 3. Partial correlations (in parentheses) depict the relationship between variables after controlling for reasoning skills (Matrix Reasoning) and verbal ability (Vocabulary).

Correlational relationships between the three tasks that were supposed to operationalize the same ability (i.e., EM, EFT, ToM, or spatial navigation) were weak, and only very few of them reached

**Table 2**
Ranges, means, and standard deviations of all 12 experimental tasks and the 2 control variables.

| hmf | Task | N | Range (possible Min and Max) | M | SD |
|-----|------|---|------------------------------|---|-----|
| EM | Treasure Box | 140 | 2–23 (2–26) | 11.3 | 4.3 |
| | Yesterday | 133 | 0–3 (0–3) | 1.6 | 1.0 |
| | Cartoon Recognition | 136 | 5–22 (0–22) | 16.8 | 3.1 |
| | Picture Book | 143 | 0–9 (0–9) | 4.1 | 2.7 |
| EFT | Tomorrow | 140 | 0–3 (0–3) | 1.8 | 0.9 |
| | Item Selection | 135 | 0–2 (0–2) | 1.0 | 0.8 |
| ToM | False Belief Location | 134 | 0–1 (0–1) | 0.4 | 0.5 |
| | False Belief Content | 123 | 0–1 (0–1) | 0.5 | 0.5 |
| | Animated Shape | 136 | 0–4 (0–6) | 0.8 | 1.0 |
| SpNa | Board & Cup | 140 | −17.8–100 (below 0–100) | 55.6 | 28.0 |
| | Map | 130 | 0.5–6 (0–6) | 3.5 | 1.4 |
| | Turning Table | 138 | 0–4 (0–4) | 1.8 | 1.2 |
| CV | Matrix Reasoning | 141 | 0–20 (0–29) | 10.3 | 4.5 |
| | Vocabulary | 139 | 2–31 (0–45) | 15.5 | 6.0 |

*Note.* For all tasks except the False Belief Location and False Belief Content tasks, range and mean values are based on the raw scores of the tasks. For the two False Belief tasks, values are based on the dichotomized scores. For dichotomizing the variables, we combined the original scores 0 and 1 into the category "failed" (score of 0) and combined the original scores 2 and 3 into the category "passed" (score of 1). Range values refer to the observed range. Possible Min and Max values (parentheses) refer to the possible minimum and maximum scores of the tasks.

**Table 3**

Pearson correlations between all 12 experimental tasks.

| | | | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| EM | 1 | Treasure Box | .106 (.090) | .166 (.128) | .136 (.071) | .082 (.093) | -.138 (-.135) | .120 (.080) | .211* (.177) | .125 (.065) | .163 (.158) | .190* (.177*) | .078 (.020) |
| | 2 | Yesterday | - | .019 (.001) | -.027 (-.047) | .400* (.403*) | -.090 (-.091) | -.061 (-.076) | -.073 (-.092) | .038 (-.016) | .011 (-.009) | .090 (.087) | -.004 (-.019) |
| | 3 | Cartoon Recognition | | - | .196* (.164) | .001 (.004) | .019 (.016) | .076 (.040) | .106 (.063) | .103 (.040) | .025 (.025) | .030 (.023) | .003 (-.032) |
| EFT | 4 | Picture Book | | | - | .071 (.081) | .087 (.098) | .116 (.080) | .221* (.191*) | .107 (.052) | -.052 (-.064) | -.007 (-.026) | .108 (.058) |
| | 5 | Tomorrow | | | | - | -.090 (-.092) | .064 (.068) | .066 (.071) | -.013 (-.009) | .022 (.023) | .166 (.169) | .113 (.122) |
| | 6 | Item Selection | | | | | - | .068 (.070) | -.193* (-.200*) | -.001 (-.004) | -.065 (-.062) | -.007 (-.004) | -.121 (-.117) |
| ToM | 7 | False Belief Location | | | | | | - | .225* (.198*) | .178* (.137) | .010 (.006) | -.012 (-.022) | .082 (.050) |
| | 8 | False Belief Content | | | | | | | - | .059 (.001) | .058 (.057) | .072 (.065) | .196* (.170) |
| | 9 | Animated Shape | | | | | | | | - | -.068 (-.075) | -.049 (.064) | .024 (-.026) |
| SpNa | 10 | Board & Cup | | | | | | | | | - | .081 (.078) | -.014 (-.023) |
| | 11 | Map | | | | | | | | | | - | .264* (.255*) |
| | 12 | Turning Table | | | | | | | | | | | - |

*Note.* Partial correlations controlling for reasoning skills (Matrix Reasoning) and verbal ability (Vocabulary) are shown in parentheses. Correlations between the three tasks testing episodic memory (EM), episodic future thinking (EFT), theory of mind (ToM), and spatial navigation (SpNa) are highlighted in gray.
*p <.05.

statistical significance (see Table 3). For EM and EFT, there were no significant correlations between any of the three tasks testing the respective ability (i.e., neither between the Treasure Box, Yesterday, and Cartoon Recognition tasks nor between the Picture Book, Tomorrow, and Item Selection tasks. For ToM, there was a significant correlation between the two False Belief tasks as well as between the False Belief Location and Animated Shape tasks. However, after controlling for reasoning skills and verbal ability, only the former but not the latter remained significant. There was no significant correlation between the False Belief Content and Animated Shape tasks. For spatial navigation, there was a significant correlation between the Map and Turning Table tasks, also after controlling for reasoning skills and verbal ability, but there were no correlations between the Board & Cup task and any of the other two spatial navigation tasks.

*Structural equation modeling*

We had planned to test our hypothesis using a multidimensional latent variable model with the four abilities EM, EFT, ToM, and spatial navigation as first-order factors, reasoning skills and verbal ability as predictors of the first-order latent factors, and self-projection as a second-order factor (see Fig. 1). Reasoning skills and verbal ability were added as predictors of the second-order latent factor. We also modeled the correlations between the residual variances of the Yesterday and Tomorrow tasks and between the two False Belief tasks, respectively, because in both cases the two tasks were structurally and conceptually very similar, thereby potentially violating the assumption of uncorrelated error variance. However, when fitting the model, the program could not find a solution because the model did not converge. This was due to the fact that, given the weak correlations, the model was not empirically identified anymore. To reach empirical identification, sufficiently strong correlations between the three indicators of each of the first-order factors (i.e., between the three tasks testing the respective abilities EM, EFT, ToM, and spatial navigation) would have been necessary. Because this was not the case, we adapted the model. Our new approach to investigate the self-projection hypothesis entailed testing whether the 12 tasks that were representing the four abilities EM, EFT, ToM, and spatial navigation were taken individually, associated with a common latent factor. Hence, we fitted a model with one latent factor, representing self-projection, and the 12 experimental tasks as indicators (see Fig. 2). To account for reasoning skills and verbal ability, these measures were added as predictors of the latent factor. We also modeled the correlations between the residual variances of the Yesterday and Tomorrow tasks and between the two False Belief tasks, respectively, in order to account for the structural similarity in these tasks. The model showed a very good fit to the data, $\chi^2(74) = 67.02$, $p = .71$, Bentler's comparative fit index (CFI) = 1.00, root mean square error of approximation and its interval (RMSEA) =.000 [.000,.037], scaling correction factor Yuan–Bentler correction =.968. However, factor loadings indicating the strength of association between the latent factor and the indicators were very low (see Fig. 2). None of the factor loadings exceeded the commonly accepted .50 threshold value. Factor loadings at and above this value are considered to represent loadings with practical significance (Hair, Black, Babin, & Anderson, 2014). Despite the good model fit, the low factor loadings speak against the validity of the measurement model. In summary, based on the low factor loadings, the results do not support the assumption of a common latent factor underlying the 12 experimental tasks.
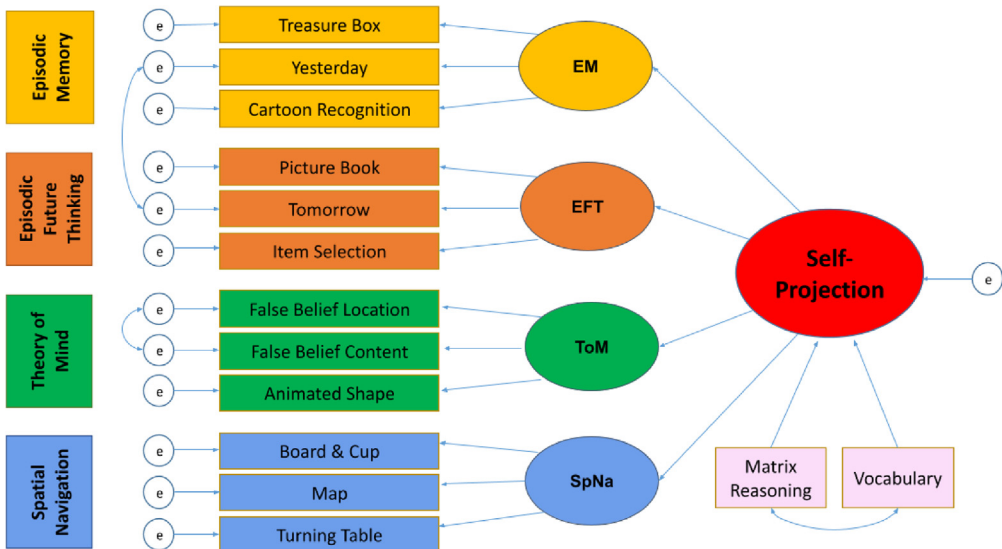


**Fig. 1.** The originally intended multidimensional first- and second-order latent factor model. EM, episodic memory; EFT, episodic future thinking; ToM, theory of mind; SpNA, spatial navigation.
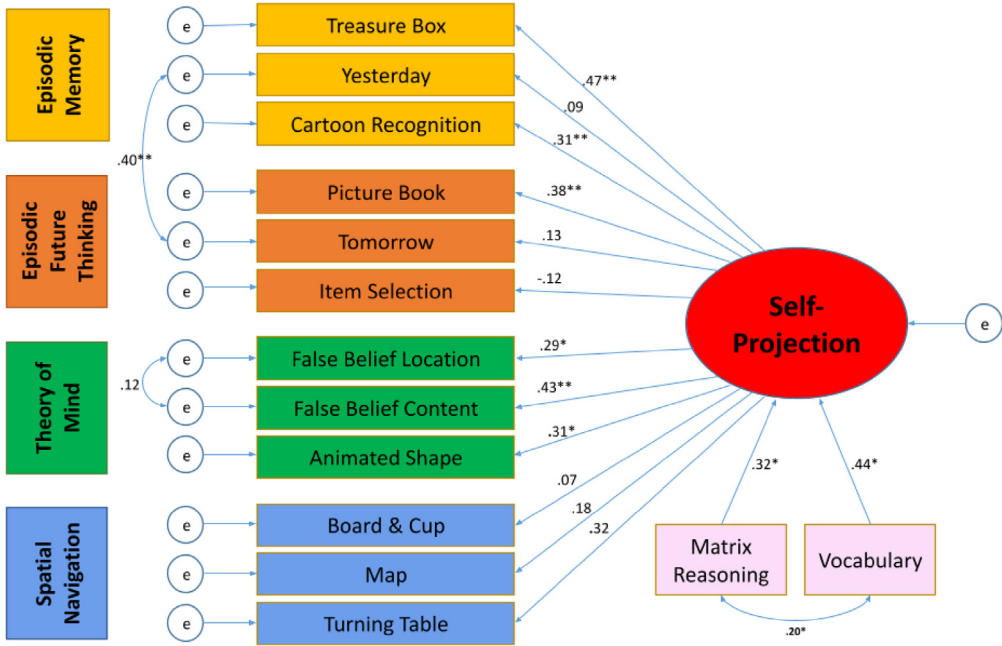
**Fig. 2.** The estimated unidimensional latent factor model with standardized factor loadings. *$p$ <.05, **$p$ <.01.

## Discussion

The self-projection account proposes that the diverse cognitive and social-cognitive abilities EM, EFT, ToM, and spatial navigation share a common cognitive underpinning, namely the ability to disengage from the current state and to project oneself mentally into an alternative temporal, social, or spatial situation (Buckner & Carroll, 2007). We set out to rigorously investigate this account by testing a large cohort of typically developing 4-year-old children on a variety of EM, EFT, ToM, and spatial navigation tasks. In addition, children were tested on reasoning skills and verbal ability to account for general cognitive abilities.

Contrary to our expectation, correlations within the set of three tasks gauging each of the four target abilities (i.e., EM, EFT, ToM, and spatial navigation) were very low. Due to this, the initially intended multidimensional latent variable model did not converge. We alternatively fitted a unidimensional model for which we considered each task as an individual indicator of the respective ability and tested whether a common latent factor, accounting for the interrelations between the tasks, could be identified. Model fit indices suggest that the model fitted the data well. However, factor loadings, indicating the relationship between the assumed latent factor and the indicators, were very low. None of the loadings reached the .50 threshold value, that is, a minimal threshold for practical significance of the association between latent variable and indicator (Hair et al., 2014). Without a sufficiently strong association between the latent variable and indicator, there is no empirical basis to conclude what the latent factor actually represents. Therefore, despite the good model fit, based on the low factor loadings, the validity of the model cannot be justified. In summary, our data do not support the assumption of a common latent factor underlying the diverse abilities EM, EFT, ToM, and spatial navigation.

*Consequences for the self-projection account and its investigation in early childhood development*

The fact that no latent factor could be identified challenges the idea of a common cognitive mechanism underlying these abilities—at least during early childhood development. However, it should be kept in mind that factor loadings are finally based on the correlations between the indicators, that is,

in our case between the administered tasks. The low factor loadings thereby can also be seen as a reflection of what the correlation matrix already indicated, namely that the relations between the tasks were very low. Our results can be interpreted in two ways: either as challenging the current theoretical framework of the self-projection account or as challenging the measures that are used in early childhood research to investigate the self-projection-related abilities EM, EFT, ToM, and spatial navigation.

### Challenges to the self-projection account

#### Lack of positive relationships between different tasks of the same ability

The finding that many of the tasks designed to measure the same ability did not correlate with each other was puzzling, especially given the fact that all tasks had been used in previous studies to assess exactly these respective abilities. However, with regard to EFT, our study is not the first to report this phenomenon. Hanson and colleagues (2014) tested 3-, 4-, and 5-year-olds on a range of EFT tasks, including the Picture Book and Tomorrow tasks that were also used in the current study, and found that, after controlling for age and language skills, none of the tasks were significantly correlated. Our study replicates their results and consolidates that this finding was not limited to their specific study design but rather seems to represent a more robust phenomenon that also persists in a large sample ($N$ = 144) with a very narrow age range (i.e., only 4-year-olds). Furthermore, our results indicate that this puzzling phenomenon not only applies to EFT but also holds for EM and to some extent also for ToM and spatial navigation.

#### Are the investigated abilities unified concepts?

What could be the reason that performances on tasks that are supposed to measure the same concept do not correlate with each other? One possible explanation could be that the investigated abilities do not represent homogeneous concepts but rather consist of different aspects or subcomponents and that hence the respective tasks do not tap into the same subcomponents of the corresponding abilities. However, subcomponents of an ability also could still develop simultaneously. Therefore, the lack of correlations may imply that those subcomponents do not develop at the same time or at least not at the same pace. This is a crucial assumption that should be taken into consideration when investigating these abilities during early childhood development. With regard to the self-projection account, this would demand an adaptation of its current theoretical framework. In particular, it should be specified which of the subcomponents of the target abilities are supposed to be related to the suggested underlying mechanism self-projection.

### Challenges of assessing cognitive capacities during early childhood

#### The task impurity problem

Our results could also be explained by the different demands of the tasks used. Perhaps the selected tasks suffer from a "task impurity problem" (for similar discussions in the realm of executive functions, see Burgess, 1997; Denckla, 1994; van der Sluis, Jong, & van der Leij, 2007). Similar to our results, a study investigating executive functions with a broad test battery found no correlations between different tasks (Miyake et al., 2000). Van der Sluis and colleagues (2007) argued that this may be due to the fact that the different paradigms are "multi-cognitive in nature," which means that they also involve additional cognitive abilities, such as working memory and verbal ability, that are not directly related to the tested concept itself. This makes it difficult to assess whether the measured performance is due to demands specific to the target ability or rather due to additional other cognitive demands.

In our study design, we tried to address this issue by controlling for children's reasoning and verbal skills. Still, additional cognitive abilities not captured by these two control variables—such as sustained attention, being able to speak freely, and flexibly switching between tasks—might have played a role here. This points to the importance of carefully choosing and thoroughly examining the administered tasks when further investigating the self-projection account. Ideally, of course, new paradigms, which do not involve such confounds, should be developed.

*Relationships across abilities*

Overall, correlation coefficients between tasks across abilities also were very low. These results converge with the preliminary findings from Lind and colleagues' (2014) clinical study of the relationships across abilities in elementary school children with ASD and a typically developing control group. In our study, an exception was a moderately high effect ($r = .40$) between the Yesterday and Tomorrow tasks that also held after controlling for verbal ability and reasoning skills. This replicates the findings from Busby and Suddendorf (2005), the original study from which the tasks had been adapted, and those from a follow-up study (Suddendorf, 2010). The correlation seems to speak in favor of a relationship between the corresponding abilities EM and EFT. However, taking into consideration the low correlations between the other EM and EFT tasks of our study, one might speculate that, instead of showing a genuine developmental link between the two involved abilities, it simply reflects the high similarity of the two tasks with regard to their structure and additional demands.

Taken together, the results of our study, in which we used a more exhaustive test battery to evaluate relationships across abilities than has been done before, point toward a more cautious interpretation of previous findings. Future investigations of the relationships across the abilities EM, EFT, ToM, and spatial navigation with more comprehensive test batteries are needed to advance our understanding on potential relationships between these skills or the lack thereof.

*Possible limitation of the current approach: Can self-projection be detected in 4-year-olds?*

Previous research has shown that at 4 years of age, the abilities EM, EFT, ToM, and spatial navigation are beginning to emerge (e.g., Atance et al., 2015; Nardini et al., 2006; Prabhakar & Hudson, 2019; Wellman & Liu, 2004). This co-emergence is considered as an indicator of the development of the neural network of the suggested underlying common mechanism (i.e., self-projection) and has been put forward as an important argument in favor of the self-projection account; the emergence of all four abilities around 4 years of age is considered to be "evidence of a common origin" (Buckner & Carroll, 2007, p. 49). By testing children at an age when the neural underpinnings are beginning to develop, we aimed to investigate the self-projection account in a thorough and unique manner making use of meaningful interindividual differences at the point of emergence of these abilities. If these abilities share a common underlying mechanism, emerging proficiency in one ability should go along with emerging proficiency in the three other self-projection-related abilities; if children have not yet developed one of the abilities, this should be reflected by low performance in the other three abilities. However, one could ask whether the underlying neural structures that are assumed to underpin self-projection are sufficiently developed in 4-year-olds, so that interrelations among the abilities can be detected if present. Buckner and Carroll (2007) suggested that self-projection is subserved by neural structures that are very similar to the brain's default mode network (DMN), an assembly of brain regions comprising the frontal, lateral, and medial parietal regions as well as medial temporal lobe structures, which has been linked to introspective and self-referential mental states (Buckner & Carroll, 2007; Raichle et al., 2001). Neuroimaging studies with children, especially those of preschool age, are scarce. However, the DMN has been investigated in 7- to 9-year-olds (Fair et al., 2008; Supekar et al., 2010). These studies show that at this age structural connectivity between the nodes of the DMN is only sparsely developed. Importantly, however, functional connectivity between some of the nodes is already in place and even partially reaches adult-like levels (Supekar et al., 2010). Hence, maturation and functional connectivity of the brain regions related to the DMN might be sufficient for the abilities EM, EFT, ToM, and spatial navigation to emerge; the brain network does not need to be fully structurally interconnected to detect potential associations among the abilities supposed to be subserved by the network. Furthermore, in an electroencephalogram (EEG) study with 4-year-olds, Sabbagh, Bowman, Evraire, and Ito (2009) found that children's ToM performance was positively related to the estimated current density of their dorsal medial prefrontal cortex and right temporal parietal junction (crucial frontal and parietal hubs of the DMN), suggesting that the maturation of these key regions is already sufficiently present at this early age (Sabbagh et al., 2009). Based on this body of research, we conclude that if there really is a relationship between the self-projection-related abilities, this link should already be detectable at 4 years of age.

*Alternative theoretical frameworks*

The results of the current study do not support the self-projection account. But do our findings perhaps speak in favor of alternative theoretical frameworks? The most prominent alternative framework to Buckner and Carrolls' (2007) self-projection account is the "scene construction theory" by Hassabis and Maguire (2007, 2009). However, the two frameworks are very similar insofar as they both suggest that EM, EFT, and spatial navigation share one underlying cognitive mechanism (self-projection and scene construction, respectively). They only differ with regard to a fourth ability included in their accounts, which is ToM for self-projection and imagination for the scene construction framework. Other theoretical approaches concentrating on "episodic simulation" (Schacter & Addis, 2007; Schacter, Addis, & Buckner, 2008) and "mental time travel" (Suddendorf & Corballis, 1997, 2007; Tulving, 1983, 2002) have a strong focus on the time component (i.e., on projecting oneself into the past or future) but include additional abilities like spatial navigation (Schacter et al., 2008; Suddendorf & Corballis, 2007), ToM (Schacter et al., 2008), and prospective memory (Schacter et al., 2008) only tangentially. Because the focus of our study was the investigation of the self-projection account and not testing it against alternative accounts, we only included paradigms aiming to measure abilities related to self-projection. This renders impossible a direct comparison of Buckner and Carroll's (2007) self-projection account with, for instance, Hassabis and Maguire's (2007) scene construction theory. However, because we found no strong evidence of interrelations among the three abilities that are included by both theoretical frameworks (i.e., EM, EFT, and spatial navigation), we can derive that if our findings are robust, the scene construction theory also does not seem to be corroborated from a developmental perspective. However, clearly, additional studies are needed to shed further light on these preliminary conclusions.

*New avenues for investigating the self-projection account*

Extensive support for the self-projection account comes from adult studies and more specifically from neuroimaging studies (e.g., Addis et al., 2007; Gallagher & Frith, 2003; Spreng et al., 2009). How does this fit with our developmental findings that are sketching a more cautious picture with regard to the validity of the account? A reason for this discrepancy could be that in adult studies the theoretical and methodological issues described above might weigh less because additional cognitive factors necessary to solve the tasks (e.g., verbal proficiency, working memory, sustained attention) are typically fully developed in adults. This potentially lessens their influence on task performance with regard to the target ability. Furthermore, many of the adult studies used neuroimaging techniques to investigate the assumptions of the self-projection account. Neuroimaging measures could present an additional advantage because they might enable a more direct capture of the (neural) connections between abilities without being sensitive to possible behavioral confounds. Consequently, investigating the self-projection account in young children with neuroimaging studies might be a fruitful approach to shed further light on potential developmental evidence of the account. However, neuroimaging with preschoolers certainly comes with its own procedural and technical challenges (e.g., Bell & Cuevas, 2012; Raschle et al., 2012). This raises a question about possible future avenues to investigate the self-projection account with behavioral measures.

Based on the presented theoretical and methodological issues, the development of thoroughly validated and largely accepted prototypical tasks that assess EM, EFT, ToM, and spatial navigation would be a promising step. To this end, first, the theoretical framework of the abilities should be further specified, for instance, with regard to assumed core components and possible subconcepts. If this examination yields that specific abilities are indeed not unified concepts but rather consist of multiple subcomponents, a thorough specification of these abilities will be indispensable in order to investigate relationships among these abilities in a meaningful way. With regard to the self-projection hypothesis, this would require subsequently determining in more detail which of these (potential) subcompo-

nents of EFT, EM, ToM, and spatial navigation are suggested to be related to the purported underlying factor self-projection. Subsequently, paradigms testing the identified concepts (or subconcepts) should be developed. In line with recent attempts to increase replicability of findings in our field (e.g., Frank et al., 2017), a thorough validation of these paradigms, ideally across different laboratories, would certainly add to the validity of the tasks and corroborate the reliability of a detected presence (or absence) of relationships between concepts.

In the process of developing these paradigms, the impurity problem should be taken into consideration. Ideally, the novel paradigms should be similar in their surface structure and low in additional verbal and cognitive demands. However, it should be kept in mind that such an alignment of tasks might be desirable only to a certain degree. Specifically, it could be argued that differences between the tasks might present a desired strong test of the purported common mechanism self-projection. After all, the central and innovative aspect of the self-projection account is the assumption that abilities that differ at the surface (EM, EFT, ToM, and spatial navigation) may be based on the same cognitive underpinning. Whereas task differences would leave unexplained variance, we would still expect shared variance by all tasks if self-projection would indeed underlie all abilities. However, if these task differences are so large that they do not allow detecting a potential link, a certain alignment appears to be inevitable. A possible step to reconcile both positions could involve creating for all four self-projection-related abilities paradigms that are aligned with regard to their surface structure but vary in terms of a core component specific to each ability.[1] An example of how such an alignment could look like is given in Appendix C. To summarize, next to the mentioned refinement of the theoretical framework, when developing new paradigms, it would be important to strike a careful balance between tasks that are similar enough in their surface structure to allow associations to be found (if existent) but at the same time still different enough from each other to ensure that the found link is meaningful and not primarily based on a structural similarity.

*Conclusions*

This is the first study to examine the self-projection hypothesis during early childhood using a comprehensive test battery including all four abilities and a large participant cohort with a narrow age range. In sum, the results of our study challenge the idea of a common cognitive mechanism underlying the four abilities EM, EFT, ToM, and spatial navigation during early childhood. How do our results fit previous findings speaking in favor of the self-projection account? Given the low factor loadings, the interrelations among EM, EFT, ToM, and spatial navigation during early development can be questioned. However, these findings could also suggest important theoretical and methodological challenges. First, there might be conceptual shortcomings with regard to the theoretical frameworks of EM, EFT, ToM, and spatial navigation, which consequently would require an adaptation and specification of the self-projection account. Second, it is possible that some of the current tasks are subject to the "task impurity problem," which would ask for a rigorous validation and "purification" of the existing paradigms or ideally the development of new ones. Because both the theoretical and methodological issues are likely related, it would be most fruitful to tackle them in parallel. Given the important implications of the self-projection account for early childhood development, if proven to be true, further investigations of this hypothesis accounting for the above-mentioned theoretical and methodological challenges are essential.

**Author contributions**

A-S.I., M.A., M.M., and S.H. jointly developed the study concept and design. A-S.I. collected the data. A-S.I. and H.M.E. performed data analyses, and all authors interpreted the data. A-S.I. drafted the manuscript, and M.A., M.M., H.M.E., and S.H. contributed critical revisions. All authors approved the final version of the manuscript for submission.

---

[1] We thank the reviewers for this suggestion.

## Acknowledgements

## Appendix A

*Episodic memory*

*Treasure Box*

The Treasure Box task is a subtest of the Wiener Entwicklungstest (WET; Kastner-Koller & Deimann, 2002). In this task, the child needed to learn and retrieve the position of six items hidden in a wooden treasure box consisting of 20 drawers. We followed the standardized protocol of the Wiener Entwicklungstest (WET; Kastner-Koller & Deimann, 2002). During the learning phase, the experimenter showed the position of the six hidden items. Then, immediate retrieval was tested by presenting duplicates of the items while the child needed to indicate in which drawer the respective item was hidden. If the child made a mistake, the retrieval round was stopped and another learning round was started, subsequently followed by another retrieval round. The procedure was repeated until the child correctly remembered the position of all items within one retrieval round. The treasure box was put out of the child's sight, and after a delay of 20 min the final retrieval phase was initiated. This time, regardless of whether the child opened the correct or incorrect drawer, the experimenter continued presenting the following item until all six items were shown. For the scoring, the amount of items correctly retrieved after the very first learning round, the total amount of learning rounds, and the amount of items correctly retrieved after 20 min of delay were taken into consideration and weighted according to the standardized scoring table, yielding possible raw scores from 2 to 26. We used these raw scores for analyses.

*Yesterday*

In this task (adapted from Busby & Suddendorf, 2005), the child was asked to report three events he or she had done the day before at school. The child was asked "What did you do yesterday in the morning circle?", "What did you do yesterday while playing outside?", and "What did you do yesterday while playing inside?" If the child answered that he or she did not know or remember, the experimenter prompted the child ("And if you think really hard") and repeated the respective question. If the child still could provide no answer, the experimenter went on to the next question. If the child gave a generic answer such as "I played," the experimenter asked "And what did you play?" Answers were checked with the teacher after the testing session. If the teacher judged the child's answer to be correct, the child could get ½ or 1 point, depending on whether the answer was still generic despite the further questioning (½ point) or the answer was specific (1 point). If the answer was judged as not being correct or if the child had answered that he or she did not remember, 0 points were given. In total, scores from 0 to 3 could be obtained. Examples of children's correct and incorrect answers are provided in Table A1.

**Table A1**
Examples of children's correct and incorrect responses to the Yesterday and Tomorrow tasks and the teacher's explanations.

| Question type | Response | Correct / Incorrect (as judged by teacher) |
|---|---|---|
| Yesterday | Puzzle | Correct |
| Yesterday | Birthday celebration classmate | Incorrect (explanation: birthday celebration was today) |
| Tomorrow | Baking cake in sandbox | Correct |
| Tomorrow | Talking about the weekend | Incorrect (explanation: is only done on Mondays) |

## Cartoon recognition

In this task (adapted from Nigro, Brandimonte, Cicogna, & Cosenza, 2014), the child watched a cartoon for 10 min and afterward needed to judge 22 scenes with regard to whether they had occurred in the cartoon or not. Before the cartoon was presented, the child was asked to watch attentively and was informed about the later retrieval task. Once the cartoon was finished, the child needed to complete a filler task for 5 min. The filler task was the Matrix Reasoning task. After 5 min, 22 pictures were shown to the child accompanied by the question of whether he or she had seen this picture in the cartoon before. Half the presented pictures depicted scenes from the cartoon the child had seen before, whereas the other half depicted scenes from a cartoon the child had not seen before but in which the same characters as in the watched cartoon were present. The child could receive a score from 0 to 22.

## Episodic future thinking

### Picture book

In this task (adapted from Atance & Meltzoff, 2005), the child was shown pictures of different places (a river and rocks scene, a desert, and a mountain scene). Subsequently, the child was asked to imagine going on a trip to the depicted place the day after and was told that it would be time to get ready. The experimenter showed three pictures presenting different items (e.g., a piece of soap, a shell, and a pair of sunglasses in the case of the desert scene) and asked which one the child wanted to take with him or her on the trip. After the child chose one item, the experimenter asked the child to explain his or her decision. Per trial, a maximum of 3 points could be achieved. According to the scoring procedure used by Atance and Meltzoff (2005), 1 point was given for the correctly chosen item, a second point was given if the child named a future term, and a third point was given if the child named an internal state term in his or her explanation. Based on the item choices and answers given for the three trials, a maximum of 9 points could be obtained. Prior to the three experimental trials, the child was acquainted with the task via two practice trials.

### Tomorrow

In this task (Busby & Suddendorf, 2005), the child was asked to report three events he or she could imagine doing the day after. The procedure was similar to the one described for the Yesterday task except that instead of asking what the child did yesterday at school, the child was asked what he or she would do the day after at school during morning circle, playing outside, and playing inside, respectively. Again, the experimenter encouraged the child to be more precise if the answer was not specific enough (e.g., the answer "I will play" was followed by the question "And what will you play?"). Answers were checked with the teacher after the testing session. If the teacher judged the child's answer to be possible, the child could get ½ or 1 point (the former if the answer was still generic despite the further questioning, and the latter if the answer was specific). If the answer was judged as not being possible or if the child had answered that he or she did not know, 0 points were given. In total, scores from 0 to 3 could be obtained.

### Item selection

In this task (adapted from Atance & Sommerville, 2014), a small locked wooden treasure box was presented to the child for inspection. Once the child had realized that the box was locked and that there was no possibility to open it, the experimenter asked the child to leave the box on the table and said that it was time to play another game outside the experimental room. A 5-min distraction game was played in the corridor. Afterward, the experimenter said that it was time to go back into the room where they had played the other games before but that, prior to going back, the child could choose one of the following items to take with him or her: a comb, a color pencil, a sharpener, or a key. After choosing an item, the child was asked to explain the choice. Back in the experimental room, the experimenter encouraged the child to use the key if he or she had chosen it. If the child had chosen another item outside the experimental room, the experimenter presented all four items again and asked the child which one he or she could use in order to open the treasure box (*knowledge question*). The child could keep the content of the treasure box (small bouncing ball). If the child could not

answer the knowledge question correctly, the task was excluded from analysis. In total, the child could receive 1 point for choosing the correct item and 1 point for giving a correct explanation, yielding possible scores from 0 to 2.

### Theory of mind

### False Belief Location (FBL)

This task (e.g., Sabbagh, Bowman, Evraire, & Ito, 2009; Wimmer & Perner, 1983) was presented to the child in the form of short 1- to 2-min videos. The protagonist put an object (e.g., a ball) in a specific location (e.g., a basket) and left the scene. During the protagonist's absence, the object was transferred to another place (e.g., a box) also present in the scene. When the protagonist came back, the child was asked the FBL question (e.g., "Where will she look for her ball?"). The child passed the test if he or she named the correct location or pointed toward it. Three FBL tasks, all following the same rationale, were administered: "Maxi and the chocolate," "Sally and the ball," and "Heidi and the plane." The maximum score that could be received for all three tasks was 3 points, that is, 1 point per correctly answered FBL question.

### False Belief Content (FBC)

In this task (e.g. Perner, Leekam, & Wimmer, 1987), a box (e.g., crayon box) was presented to the child and the experimenter asked what he or she thought was inside (Question 1). The experimenter showed the actual content of the box (e.g., puzzle pieces) and asked the child what it was (Question 2). After the box was closed again, a hand puppet entered the scene and the experimenter said, "That's Mickey. He sees the box for the first time. He has never looked inside. Think hard: What does Mickey think is inside?" (Question 3 = FBC question). If the child answered any of the three questions with anything apart from the usually expected content (e.g., crayons) or the actual content (puzzle pieces), the FBC task was excluded from analysis. Three FBC tasks, all following the same rationale, were administered: crayon box filled with puzzle pieces, milk package filled with water, and chocolate sprinkles box filled with buttons. The maximum score that could be received for all three tasks together was 3 points, that is, 1 point per correctly answered FBC question.

### Animated shape

In this task (adapted from Abell, Happé, & Frith, 2000), the child watched three short films about moving triangles and was asked to describe what was happening while watching. Different from Abell and colleagues, we only presented the theory of mind sequences and adapted the procedure according to the age group we were testing. Every film was presented twice: one time for watching and thereby getting acquainted with the film and a second time for completing the actual task, that is, giving a description of the film. To acquaint the child with the task, the experimenter explained the procedure by giving an example. First, the "surprise" film was shown without giving any descriptions. Afterward, the same film was watched again, but this time the experimenter simultaneously told a story about what was going on. The story contained six mentalization words. For the following three experimental trials ("coaxing," "mocking," and "seducing"), this procedure was repeated except that it was now the child's task to interpret what was going on. If the child did not start talking by himself or herself, the experimenter encouraged the child a maximum of three times by asking, "What is happening now?" The child's answers were scored according to the scoring guideline used by Abell et al. Per experimental trial, the child could receive a mentalization score ranging from 0 to 2, yielding a total score from 0 to 6for all three trials.

### Spatial navigation

### Board & cup

In this task (adapted from Nardini, Burgess, Breckenridge, & Atkinson, 2006), a board, covered with different toys (play houses and plush animals) on two sides of the board and with 12 cups in the middle part, was presented to the child. We tested two of the conditions used by Nardini and colleagues: child-move and array-move (see Fig. A1). In the child-move condition, while the child was standing at
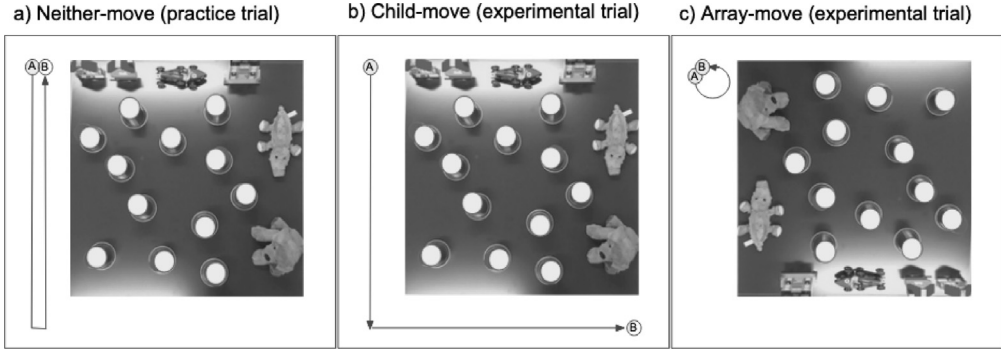
**Fig. A1.** Board & Cup task. Depicted are the arrangements of landmarks (toys) and the 12 cups on the board as well as the child's location relative to the board during encoding and retrieval. Panel A indicates the child's location during encoding, and Panel B indicates the child's location during retrieval. In the neither-move and child-move conditions, the board remained in the same position during encoding and retrieval. However, in the array-move condition (Panel C), the board was subjected to a rotation; during encoding the board was positioned as depicted in Panels A and B, but during retrieval the board was positioned as depicted in Panel C.

one pointy edge of the board, a little toy animal was hidden under one of the cups. Without looking at the board, the child was led to the opposite edge of the board and needed to indicate from this novel position under which cup the toy was hidden. The child could use the toy animals at the side edges as landmarks supporting the encoding and retrieval after the perspective change. The upcoming trial always started from the position at which the previous trial had been finished. We tested this condition three times. In the array-move condition, instead of walking to the other edge of the board after hiding the toy animal, the child needed to turn around and fixate a point on the wall while the experimenter rotated the board (180° rotation). Upon turning back toward the board, the child needed to indicate where the toy animal was hidden. We tested this condition once. Prior to the described experimental trials, the child was introduced to the task by completing two practice trials. Both practice trials were the same as the ones of the child-move condition with the exception that the first one included walking but no perspective change (i.e., the child came back to the same position at which encoding took place; see neither-move condition in Fig. A1). The scoring was conducted the same way as described by Nardini and colleagues. A score of 100 indicated a correct choice, a score of 0 indicated a choice equal to chance level, and a score below 0 indicated a choice below chance level. For the final score, the scores for all four trials were added and divided by 4.

*Map*

In this task (adapted from Shusterman, Lee, & Spelke, 2008), a map lying on a table and depicting three circles arranged in either a right triangular or a linear shape was presented to the child. One of the three circles had a star in its center. The child was told that Kikker (the name of a small toy frog) would like to sit where the star was. Behind the child's back, three buckets were arranged according to the shape depicted on the map. The child was asked to turn around. During the first trial, a practice trial, the experimenter explained that the arrangement of the circles in the map corresponded to the arrangement of the buckets on the ground. The map was placed back on the table, and the child was asked to put Kikker into the bucket according to the position indicated via the star on the map. The procedure of the following six experimental trials was similar to the one described for the exercise trial except that the map was left on the table before turning around and there was no additional explanation or feedback on whether the child had chosen the correct bucket. In total, the experimental trials encompassed three trials with a right triangle and three trials with a linear arrangement of circles and buckets, respectively. For each of the three trials within one shape, the star was placed into another circle for every trial. Like this, all possible placing positions within one shape were tested. The maps depicting the two shapes were alternated for each trial. Per trial, the child could receive 0 to 1

point: 1 point if the correct bucket had been chosen after having seen the map only once, ½ point if the correct bucket was chosen but the child needed to look back at the map before placing Kikker, and 0 points if the wrong bucket was chosen. Taking all trials into consideration, scores from 0 to 6 could be obtained.

*Turning table*

In this task, two pictures showing three objects on a turning table were presented one after the other, each showing the object arrangement from a different perspective. In the second picture, not only the perspective but also the position of one of the objects was changed. The child needed to indicate which object was displaced in relation to the first picture. The paradigm was based on the ones used by Wang and Spelke (2002) and Lambrey, Doeller, Berthoz, and Burgess (2012). After piloting, we adapted and simplified the task in the following ways to match the capabilities of our age group. To familiarize the child with this complex task, the procedure was explained during two practice trials. The first practice trial explained the change of position of one of the objects. A picture depicting three items on a round table was presented, and the child was asked to name the objects and to encode their positions (see Fig. A2, Picture A1). While the child closed his or her eyes, the experimenter replaced the first picture with the second one (see Fig. A2, Picture B1). For this first exercise trial, only the position of one of the objects was changed; the perspective was the same. The child needed to indicate
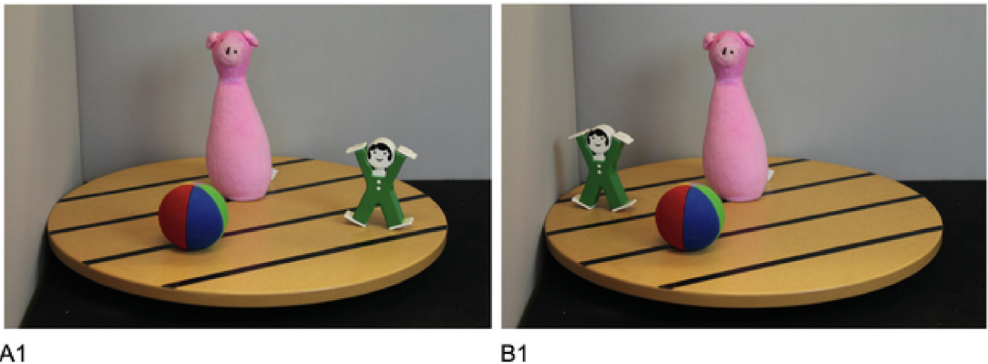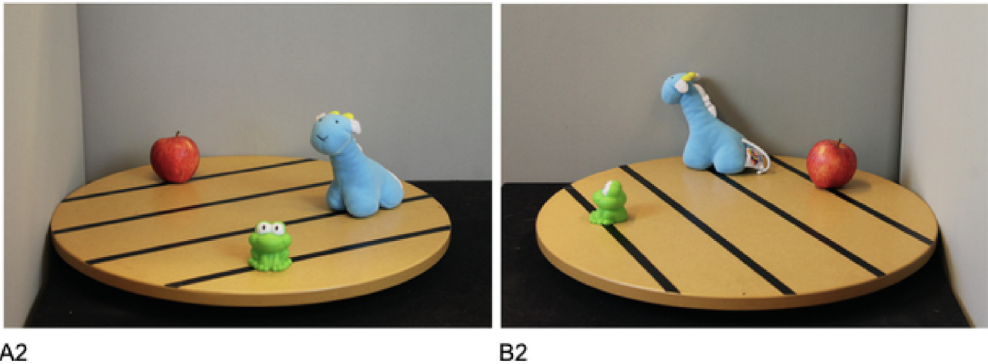


**Fig. A2.** Turning Table task. Examples of stimuli material showing the setup of the practice trials (Pictures A1 and B1), as well as the experimental trials (Pictures A2 and B2), are shown.

which object was displaced. The second practice trial introduced the child to the perspective change. On four small pictures depicting the table from different perspectives, the experimenter demonstrated from which perspective the first picture was taken and that the photographer walked a quarter of a circle around the table in order to take the second picture from a new perspective (90° change in perspective). The same procedure as described for the first practice trial was executed except that now the second picture included an object as well as a perspective change (see Fig. A2, Pictures A2 and B2). The two practice trials were followed by four experimental trials, all of which adhered to the same procedure as described for the second practice trial (i.e., with pictures including object and perspective changes). The three objects depicted in the picture were different for every experimental trial. Per trial, the child could receive 1 point if he or she correctly spotted the object that changed position, yielding total scores from 0 to 4.

### Control variables

To control for reasoning and verbal ability skills, we administered the Matrix Reasoning and Vocabulary tasks, respectively, both taken from the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-III-NL; Wechsler, 2002). Instructions in accordance with our tested age range were used.

### Matrix reasoning

In this task, a matrix of three items that needed to be completed by a fourth logically fitting item was presented to the child. The child needed to choose one of four possible items to complete the matrix. For the analyses, we used the raw scores, yielding possible total scores ranging from 0 to 29.

### Vocabulary

In this task, the child was asked to define words that were read aloud by the experimenter ("What is [verbal item]?"). Per verbal item, 0, 1, or 2 points could be received depending on the given explanation. Judgment of explanations and related ratings were determined by the official protocol of the task. For instance, if the child needed to explain the word "letter," he or she would receive 2 points by naming either two characteristics of a letter (e.g., "envelope," "reading," "there is your name on it") or two related actions (e.g., "putting a stamp on it," "you get it by post," "sending," "found in the letter box") or a combination of a characteristic and an action. The child was given 1 point if he or she named only one of these characteristics or actions and was given 0 points if the answer was not related to or decisive for the item (e.g., "paper," "coloring"). For our target age group (4-year-olds), the task started with the explanation of words. If the child failed to receive a perfect score (2 points) for at least one of the first two verbal items ("umbrella" and "dog"), five pictures depicting objects were presented and the child was asked to name these. If this was successful, the task continued with the explanation of verbal items until the stopping criterion was reached (five 0 scores in succession). In total, there were 20 verbal items, which increased in difficulty from one item to the next. For the analysis, we used the raw scores, yielding possible total scores ranging from 0 to 45.

## Appendix B

See Table B1.

**Table B1**

Distributional details and results of Shapiro–Wilk test of normal distribution for the 12 experimental tasks and 2 control measures.

|  |  | Skewness | *SE* | Kurtosis | *SE* | Shapiro–Wilk test | *p* |
|---|---|---|---|---|---|---|---|
| EM | Treasure Box | 0.301 | 0.205 | 0.220 | 0.407 | 0.979 | .027 |
|  | Yesterday | −0.191 | 0.210 | −1.092 | 0.417 | 0.908 | <.001 |
|  | Cartoon Recognition | −0.980 | 0.208 | 0.725 | 0.413 | 0.914 | <.001 |
| EFT | Picture Book | 0.372 | 0.203 | −0.858 | 0.403 | 0.935 | <.001 |
|  | Tomorrow | −0.567 | 0.205 | −0.594 | 0.407 | 0.887 | <.001 |
|  | Item Selection | −0.025 | 0.209 | −1.271 | 0.414 | 0.808 | <.001 |
| ToM | False Belief Location | 0.208 | 0.209 | −1.495 | 0.416 | 0.823 | <.001 |
|  | False Belief Content | −0.128 | 0.218 | −1.679 | 0.433 | 0.790 | <.001 |
|  | Animated Shape | 1.001 | 0.208 | 0.211 | 0.413 | 0.783 | <.001 |
| SpNa | Board & Cup | −0.434 | 0.205 | −0.225 | 0.407 | 0.968 | .002 |
|  | Map | −0.044 | 0.206 | −0.883 | 0.408 | 0.950 | <.001 |
|  | Turning Table | 0.367 | 0.206 | −0.847 | 0.410 | 0.896 | <.001 |
| CV | Matrix Reasoning | −0.258 | 0.204 | −0.613 | 0.406 | 0.974 | .009 |
|  | Vocabulary | −0.002 | 0.206 | −0.075 | 0.408 | 0.984 | .117 |

*Note.* EM, episodic memory; EFT, episodic future thinking; ToM, theory of mind; SpNa, spatial navigation; CV, control variables.

## Appendix C

*Aligning surface structure demands across paradigms while manipulating the core component relevant to the respective ability*

Below, we give an examples of what aligning task demands across abilities could look like. A reasonable approach would entail choosing one of the existing paradigms as a template and designing, along the cognitive and verbal demands of that paradigm, tasks representing the other three abilities. The Picture Book task, which measures episodic future thinking, would be one such suitable template.

*Episodic future thinking*

The Picture Book task (Atance & Meltzoff, 2005; see also Appendix A) consists of the presentation of photographs depicting various landscapes. In preparation for an imaginary trip to that place, three items are presented, one of which is suited for the depicted environment. Children are asked to choose one of the items and to explain their choice. The main elements of the Picture Book task thereby are the presentation of pictures, the (possibly nonverbal) choice of an item, and a verbal explanation of the choice made.

*Episodic memory*

An analogous episodic memory task could consist of presenting a short engaging movie in which different characters each play with a different object. In a subsequent retrieval phase, scenes depicting each of the characters in turn (but without their respective object) could be presented; for each one, children would be asked to choose which of three items they think the character had previously played with and to explain their choice.

*Theory of mind*

With regard to theory of mind, the classic Sally Ann task assessing false belief understanding (Baron-Cohen, Leslie, & Frith, 1985; see also False Belief Location task in Appendix A) could be told alongside photographs. However, to align with the other paradigms, three containers (e.g., a basket, a box, and a treasure box) for the "storage" of Sally's ball could be included. After completion of the story, children would be presented with three pictures, each depicting one of the containers, and would be asked to answer the false belief question ("Where will Sally look for her ball?") by pointing to the picture and subsequently explaining their choice.

*Spatial navigation*

To measure spatial navigation, the Turning Table task (Lambrey, Doeller, Berthoz, & Burgess, 2012; Wang & Spelke, 2002; see also Appendix A) could be adapted. Children would be presented with two pictures depicting the same three objects from different vantage points. In the second picture, taken from an alternative perspective, one of the three objects would be in a different position. Children would then be shown three pictures, each representing one of the objects, and would be asked to point to the object that had changed position and to explain their choice.

All tasks are aligned with regard to three main elements: picture presentation, nonverbal item choice, and verbal explanation of choice. Moreover, they are low in additional demands other than the ones specific to that ability (e.g., encoding for the episodic memory task, mental rotation for the spatial navigation task). In addition, thanks to the inclusion of a nonverbal choice and an explicit verbal explanation of the choice, performance effects linked to children's expressive verbal ability would be clearly detectable. Notably, this is just an example, and its validity would depend largely on whether the Picture Book is regarded as a prototypical task of EFT. Furthermore, a thorough validation of the template-aligned paradigms would be indispensable.

# References

Abell, F., Happé, F., & Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cognitive Development, 15*, 1–16.

Acredolo, L. P. (1978). Development of spatial orientation in infancy. *Developmental Psychology, 14*, 224–234.

Addis, D. R., Wong, A. T., & Schacter, D. L. (2007). Remembering the past and imagining the future: Common and distinct neural substrates during event construction and elaboration. *Neuropsychologia, 45*, 1363–1377.

Addis, D. R., Wong, A. T., & Schacter, D. L. (2008). Age-related changes in the episodic simulation of future events. *Psychological Science, 19*, 33–41.

Adornetti, I., Chiera, A., Altavilla, D., Deriu, V., Marini, A., Valeri, G., ... Ferretti, F. (2021). Self-projection in middle childhood: A study on the relationship between theory of mind and episodic future thinking. *Cognitive Processing, 22*, 321–332.

Andrews-Hanna, J. R., Saxe, R., & Yarkoni, T. (2014). Contributions of episodic retrieval and mentalizing to autobiographical thought: Evidence from functional neuroimaging, resting-state connectivity, and fMRI meta-analyses. *NeuroImage, 91*, 324–335.

Astington, J. W., & Jenkins, J. M. (1995). Theory of mind development and social understanding. *Cognition and Emotion, 9*, 151–165.

Atance, C. M., Louw, A., & Clayton, N. S. (2015). Thinking ahead about where something is needed: New insights about episodic foresight in preschoolers. *Journal of Experimental Child Psychology, 129*, 98–109.

Atance, C. M., & Meltzoff, A. N. (2005). My future self: Young children's ability to anticipate and explain future states. *Cognitive Development, 20*, 341–361.

Atance, C. M., & O'Neill, D. K. (2001). Episodic future thinking. *Trends in Cognitive Sciences, 5*, 533–539.

Atance, C. M., & O'Neill, D. K. (2005). The emergence of episodic future thinking in humans. *Learning and Motivation, 36*, 126–144.

Atance, C. M., & Sommerville, J. A. (2014). Assessing the role of memory in preschoolers' performance on episodic foresight tasks. *Memory, 22*, 118–128.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21*, 37–46.

Bell, M. A., & Cuevas, K. (2012). Using EEG to study cognitive development: Issues and practices. *Journal of Cognition and Development, 13*, 281–294.

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences, 11*, 49–57.

Bullens, J., Iglói, K., Berthoz, A., Postma, A., & Rondi-Reig, L. (2010). Developmental time course of the acquisition of sequential egocentric and allocentric navigation strategies. *Journal of Experimental Child Psychology, 107*, 337–350.

Burgess, N. (2006). Spatial memory: How egocentric and allocentric combine. *Trends in Cognitive Sciences, 10*, 551–557.

Burgess, N. (2008). Spatial cognition and the brain. *Annals of the New York Academy of Sciences, 1124*, 77–97.

Burgess, P. W. (1997). Theory and methodology in executive function research. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 81–116). Hove, UK: Psychology Press.

Busby, J., & Suddendorf, T. (2005). Recalling yesterday and predicting tomorrow. *Cognitive Development, 20*, 362–372.

Causey, K. B., & Bjorklund, D. F. (2014). Prospective memory in preschool children: Influences of agency, incentive, and underlying cognitive mechanisms. *Journal of Experimental Child Psychology, 127*, 36–51.

Crane, L., & Goddard, L. (2008). Episodic and semantic autobiographical memory in adults with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38*, 498–506.

Dafni-Merom, A., & Arzy, S. (2020). The radiation of autonoetic consciousness in cognitive neuroscience: A functional neuroanatomy perspective. *Neuropsychologia, 143*, 107477.

Denckla, M. B. (1994). Measurement of executive function. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 117–142). Baltimore, MD: Brookes.

Fair, D. A., Cohen, A. L., Dosenbach, N. U., Church, J. A., Miezin, F. M., Barch, D. M., ... Schlaggar, B. L. (2008). The maturing architecture of the brain's default network. *Proceedings of the National Academy of Sciences of the United States of America, 105*, 4028–4032.

Ford, R. M., Driscoll, T., Shum, D., & Macaulay, C. E. (2012). Executive and theory-of-mind contributions to event-based prospective memory in children: Exploring the self-projection hypothesis. *Journal of Experimental Child Psychology, 111*, 468–489.

Frank, M. C., Bergelson, E., Bergmann, C., Cristia, A., Floccia, C., Gervain, J., ... Yurovsky, D. (2017). A collaborative approach to infant research: Promoting reproducibility, best practices, and theory-building. *Infancy, 22*, 421–435.

Frith, C., & Frith, U. (2005). Theory of mind. *Current Biology, 15*, 644–645.

Galati, G., Pelle, G., Berthoz, A., & Committeri, G. (2010). Multiple reference frames used by the human brain for spatial perception and memory. *Experimental Brain Research, 206*, 109–120.

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of "theory of mind". *Trends in Cognitive Sciences, 7*, 77–83.

Goldman, A. I. (2006). *Simulating minds: The philosophy, psychology, and neuroscience of mindreading (Philosophy of Mind series)*. New York: Oxford University Press.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis* (7th ed.). Harlow, UK: Pearson.

Hanson, L. K., Atance, C. M., & Paluck, S. W. (2014). Is thinking about the future related to theory of mind and executive function? Not in preschoolers. *Journal of Experimental Child Psychology, 128*, 120–137.

Hassabis, D., Kumaran, D., & Maguire, E. A. (2007). Using imagination to understand the neural basis of episodic memory. *Journal of Neuroscience, 27*, 14365–14374.

Hassabis, D., & Maguire, E. A. (2007). Deconstructing episodic memory with construction. *Trends in Cognitive Sciences, 11*, 299–306.

Hassabis, D., & Maguire, E. A. (2009). The construction system of the brain. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*, 1263–1271.

Hayne, H., Gross, J., McNamee, S., Fitzgibbon, O., & Tustin, K. (2011). Episodic memory and episodic foresight in 3- and 5-year-old children. *Cognitive Development, 26*, 343–355.

Kastner-Koller, U., & Deimann, P. (2002). *Wiener Entwicklungstest (WET): Ein allgemeines Verfahren zur Erfassung des allgemeinen Entwicklungsstandes bei Kinder von 3 bis 6 Jahren* (2nd ed.). Göttingen, Germany: Hogrefe.

Klein, S. B. (2016). Autonoetic consciousness: Reconsidering the role of episodic memory in future-oriented self-projection. *Quarterly Journal of Experimental Psychology, 69*, 381–401.

Klein, S. B., Loftus, J., & Kihlstrom, J. F. (2002). Memory and temporal experience: The effects of episodic memory loss on an amnesic patient's ability to remember the past and imagine the future. *Social Cognition, 20*, 353–379.

Kretschmer-Trendowicz, A., Ellis, J. A., & Altgassen, M. (2016). Effects of episodic future thinking and self-projection on children's prospective memory performance. *PLoS ONE, 11*(6), e158366.

Kurczek, J., Wechsler, E., Ahuja, S., Jensen, U., Cohen, N. J., Tranel, D., & Duff, M. (2015). Differential contributions of hippocampus and medial prefrontal cortex to self-projection and self-referential processing. *Neuropsychologia, 73*, 116–126.

Lambrey, S., Doeller, C., Berthoz, A., & Burgess, N. (2012). Imagining being somewhere else: Neural basis of changing perspective in space. *Cerebral Cortex, 22*, 166–174.

Lind, S. E., Bowler, D. M., & Raber, J. (2014). Spatial navigation, episodic memory, episodic future thinking, and theory of mind in children with autism spectrum disorder: Evidence for impairments in mental simulation? *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.01411.

Lind, S. E., Williams, D. M., Bowler, D. M., & Peel, A. (2014). Episodic memory and episodic future thinking impairments in high-functioning autism spectrum disorder: An underlying difficulty with scene construction or self-projection? *Neuropsychology, 28*, 55–67.

Lind, S. E., Williams, D. M., Raber, J., Peel, A., & Bowler, D. M. (2013). Spatial navigation impairments among intellectually high-functioning adults with autism spectrum disorder: Exploring relations with theory of mind, episodic memory, and episodic future thinking. *Journal of Abnormal Psychology, 122*, 1189–1199.

Mahr, J. B. (2020). The dimensions of episodic simulation. *Cognition, 196*, 104085.

McNamara, T. P., Rump, B., & Werner, S. (2003). Egocentric and geocentric frames of reference in memory of large-scale space. *Psychonomic Bulletin & Review, 10*, 589–595.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49–100.

Naito, M. (2003). The relationship between theory of mind and episodic memory: Evidence for the development of autonoetic consciousness. *Journal of Experimental Child Psychology, 85*, 312–336.

Nardini, M., Burgess, N., Breckenridge, K., & Atkinson, J. (2006). Differential developmental trajectories for egocentric, environmental and intrinsic frames of reference in spatial memory. *Cognition, 101*, 153–172.

Nigro, G., Brandimonte, M. A., Cicogna, P., & Cosenza, M. (2014). Episodic future thinking as a predictor of children's prospective memory. *Journal of Experimental Child Psychology, 69*, 82–94.

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford, UK: Clarendon.

Okuda, J., Fujii, T., Ohtake, H., Tsukiura, T., Tanji, K., Suzuki, K., ... Yamadori, A. (2003). Thinking of the future and past: The roles of the frontal pole and the medial temporal lobes. *NeuroImage, 19*, 1369–1380.

Perner, J. (2000). Memory and theory of mind. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 297–312). Oxford, UK: Oxford University Press.

Perner, J., Kloo, D., & Gornik, E. (2007). Episodic memory development: Theory of mind is part of re-experiencing experienced events. *Infant and Child Development, 16*, 471–490.

Perner, J., Kloo, D., & Rohwer, M. (2010). Retro- and prospection for mental time travel: Emergence of episodic remembering and mental rotation in 5- to 8-year-old children. *Consciousness and Cognition, 19*, 802–815.

Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology, 5*, 125–137.

Piaget, J., & Inhelder, B. (1967). The coordination of perspectives. In *The child's conception of space* (pp. 209–246). New York: Norton.

Prabhakar, J., & Hudson, J. A. (2019). Past is prologue: The role of memory retrieval in young children's episodic prospection. *Journal of Experimental Child Psychology, 177*, 17–35.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*, 515–526.

Quon, E., & Atance, C. M. (2010). A comparison of preschoolers' memory, knowledge, and anticipation of events. *Journal of Cognition and Development, 11*, 37–60.

R Development Core Team (2016). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America, 98*, 676–682.

Raschle, N., Zuk, J., Ortiz-Mantilla, S., Sliva, D. D., Franceschi, A., Grant, P. E., ... Gaab, N. (2012). Pediatric neuroimaging in early childhood and infancy: Challenges and practical guidelines. *Annals of the New York Academy of Sciences, 1252*, 43–50.

Richmond, J. L., & Pan, R. (2013). Thinking about the future early in life: The role of relational memory. *Journal of Experimental Child Psychology, 114*, 510–521.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5-12 (BETA). *Journal of Statistical Software, 48*(2), 1–36.

Ruggiero, G., D'Errico, O., & Iachini, T. (2016). Development of egocentric and allocentric spatial representations from childhood to elderly age. *Psychological Research Psychologische Forschung, 80*, 259–272.

Sabbagh, M. A., Bowman, L. C., Evraire, L. E., & Ito, J. M. (2009). Neurodevelopmental correlates of theory of mind in preschool children. *Child Development, 80*, 1147–1162.

Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*, 773–786.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2008). Episodic simulation of future events: Concepts, data, and applications. *Annals of the New York Academy of Sciences, 1124*, 39–60.

Schacter, D. L., Benoit, R. G., & Szpunar, K. K. (2017). Episodic future thinking: Mechanisms and functions. *Current Opinion in Behavioral Sciences, 17*, 41–50.

Shusterman, A., Lee, S. A., & Spelke, E. S. (2008). Young children's spontaneous use of geometry in maps. *Developmental Science, 11*, F1–F7.

Spek, A. A., Scholte, E. M., & Van Berckelaer-Onnes, I. A. (2010). Theory of mind in adults with HFA and Asperger syndrome. *Journal of Autism and Developmental Disorders, 40*, 280–289.

Spreng, R. N., & Grady, C. L. (2010). Patterns of brain activity supporting autobiographical memory, prospection, and theory of mind, and their relationship to the default mode network. *Journal of Cognitive Neuroscience, 22*, 1112–1123.

Spreng, R. N., Mar, R. A., & Kim, A. S. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: A quantitative meta-analysis. *Journal of Cognitive Neuroscience, 21*, 489–510.

Suddendorf, T. (2010). Linking yesterday and tomorrow: Preschoolers' ability to report temporally displaced events. *British Journal of Developmental Psychology, 28*, 491–498.

Suddendorf, T., & Corballis, M. C. (1997). Mental time travel and the evolution of the human mind. *Genetic Social and General Psychology Monographs, 123*, 133–167.

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences, 30*, 299–313.

Suddendorf, T., Nielsen, M., & Von Gehlen, R. (2011). Children's capacity to remember a novel problem and to secure its future solution. *Developmental Science, 14*, 26–33.

Supekar, K., Uddin, L. Q., Prater, K., Amin, H., Greicius, M. D., & Menon, V. (2010). Development of functional and structural connectivity within the default mode network in young children. *NeuroImage, 52*, 290–301.

Szpunar, K. K., Spreng, R. N., & Schacter, D. L. (2014). A taxonomy of prospection: Introducing an organizational framework for future-oriented cognition. *Proceedings of the National Academy of Sciences of the United States of America, 111*, 18414–18421.

Szpunar, K. K., Watson, J. M., & McDermott, K. B. (2007). Neural substrates of envisioning the future. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 642–647.

Terrett, G., Rendell, P. G., Raponi-Saunders, S., Henry, J. D., Bailey, P. E., & Altgassen, M. (2013). Episodic future thinking in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 43*, 2558–2568.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). New York: Academic Press.

Tulving, E. (1983). *Elements of episodic memory.* Oxford, UK: Oxford University Press.

Tulving, E. (1985a). How many memory systems are there? *American Psychologist, 40*, 385–398.

Tulving, E. (1985b). Memory and consciousness. *Canadian Psychology/Psychologie Canadienne, 26*, 1–12.

Tulving, E. (1987). Multiple memory systems and consciousness. *Human Neurobiology, 6*(2), 67–80.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology, 53*, 1–25.

Ünal, G., & Hohenberger, A. (2017). The cognitive bases of the development of past and future episodic cognition in preschoolers. *Journal of Experimental Child Psychology, 162*, 242–258.

van der Sluis, S., de Jong, P. F., & van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence, 35*, 427–449.

van Hoogmoed, A. (2014). *Finding your way around: The development of landmark use and its precursors.* Nijmegen, Netherlands: Radboud University Nijmegen.

Wang, R. F., & Spelke, E. S. (2002). Human spatial representation: Insights from animals. *Trends in Cognitive Sciences, 6*, 376–382.

Wechsler, D. (2002). *WPPSI-III-NL Nederlandse bewerking: Technische Handleiding [Dutch version of the WPPSI-III-NL: Technical and interpretive manual]* (2nd ed.). Amsterdam, Netherlands: Pearson Assessment and Information.

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

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1997). Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness. *Psychological Bulletin, 121*, 331–354.

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

Wu, H., Liu, X., Hagan, C. C., & Mobbs, D. (2020). Mentalizing during social InterAction: A four component model. *Cortex, 126*, 242–252.

Yang, Y., Merrill, E. C., & Wang, Q. (2019). Children's response, landmark, and metric strategies in spatial navigation. *Journal of Experimental Child Psychology, 181*, 75–101.