

# Spatial Perspective Taking Is Distinct From Cognitive and Affective Perspective Taking

Maria Brucato, Nora S. Newcombe, and Jason Chein

Department of Psychology, Temple University

Perspective taking (PT) is the ability to imagine viewpoints different from our own. However, the nature of PT as a construct and its underlying cognitive mechanisms are not well established. Some researchers propose that understanding what others believe (cognitive PT), feel (affective PT), and see (spatial PT) forms a single behavioral dimension, relying on the orienting of attention between competing frame-of-reference representations. Others propose that PT mechanisms are dissociable, although there are three different proposals about such dissociations. The present study examined behavioral associations among measures of spatial, cognitive, and affective PT and attentional control in neurotypical young adults. There was a lack of convergent validity for measures of cognitive and affective PT, pointing to the need for more psychometric work on these dimensions. Much better convergence was found for spatial PT measures. There was little to no behavioral association between spatial PT and either social form of PT (cognitive or affective) or attentional control measures. This pattern suggests support for a dissociated model in which spatial PT is a distinct cognitive construct.

## Public Significance Statement

Imagining a different point of view, like what others believe or how they feel, allows us to better navigate social situations, and imagining a different point of view of a physical scene allows us to better navigate spatial environments. Deficits in these different types of perspective-taking skills are apparent in several *Diagnostic and Statistical Manual of Mental Disorders*, fifth edition categorized clinical populations including autism spectrum disorder, schizophrenia, and major depressive disorder, and differences in the severity of perspective-taking impairments may be related to more general skills like attentional control rather than diagnostic categories. This study used data from neurotypical young adults' self-reported and actual perspective-taking performance and measures of attentional control to distinguish between four proposals on the association of perspective-taking abilities.

**Keywords:** perspective taking, theory of mind, empathic accuracy, attentional control

Perspective taking (PT) is the name often given to a psychological construct that includes the ability to imagine another person's experience in one of several ways: what others see (spatial PT), what they believe (cognitive PT), and what they feel (affective PT). Spatial PT supports successful spatial navigation of large-scale environments (Allen et al., 1996; Fields & Shelton, 2006; Kozhevnikov et al., 2006; Nazareth et al., 2018) and may also contribute to

educational achievement and career success in science, technology, engineering, and mathematics disciplines (Oldakowski, 2001; Plummer et al., 2016; Wai et al., 2009). Cognitive and affective PT are important for forming and maintaining high-quality social relationships (Galinsky et al., 2005; Hughes & Leekam, 2004; E. C. Long & Andrews, 1990). Deficits in cognitive and affective PT are also characteristic of many clinical populations (Cotter et al., 2018),

This article was published Online First January 8, 2026.

Rachel Wu served as action editor.

Maria Brucato  <https://orcid.org/0000-0002-7272-2622>

Nora S. Newcombe  <https://orcid.org/0000-0002-7044-6046>

Jason Chein  <https://orcid.org/0000-0002-8430-7899>

Maria Brucato is now at the Jefferson Center for Interprofessional Practice and Education, Thomas Jefferson University.

The ideas and data presented in this article were previously shared in postdoctoral job talks, in a poster presentation at the 63rd annual meeting of the Psychonomic Society, and in a bioRxiv preprint. The authors have no conflicts of interest to disclose. This work was supported by the Psi Chi International Honor Society in Psychology, 2020–2021 Summer Graduate Research Grant to Maria Brucato and the National Science Foundation EHR 1660996 to Nora S. Newcombe. The authors thank Alina Spas, Ria Kazmi,

Re-Al Mitchell, Dara Rogers, Serina Victor, and Kayla Wyatt for their assistance with data collection and Ria and Kayla for their additional assistance with data coding.

Maria Brucato played a lead role in data curation, formal analysis, visualization, and writing—original draft and an equal role in conceptualization and methodology. Nora S. Newcombe played a lead role in funding acquisition and an equal role in conceptualization, resources, supervision, and writing—review and editing. Jason Chein played a supporting role in project administration and an equal role in conceptualization, methodology, resources, supervision, and writing—review and editing.

Correspondence concerning this article should be addressed to Maria Brucato, Jefferson Center for Interprofessional Practice and Education, Thomas Jefferson University, 130 South 9th Street, 1839C Edison, Philadelphia, PA 19107, United States. Email: [maria.brucato@jefferson.edu](mailto:maria.brucato@jefferson.edu)

including autism spectrum disorder (Baron-Cohen, 2000) and major depressive disorder (Bora & Berk, 2016; Nestor et al., 2022; Zobel et al., 2010), and may play a role in development, progression, and responses to treatment (Bora & Zorlu, 2017; Quednow, 2017). Differences in the severity of PT impairments may be related to general mechanisms that support this ability regardless of diagnostic category and may help connect and differentiate between traditional diagnostic categories as in a research domain criteria approach (Cuthbert & Insel, 2013). However, it is currently debated whether PT constitutes a singular psychological construct, and if so, what cognitive mechanisms support it.

### Does PT Constitute a Singular Psychological Construct?

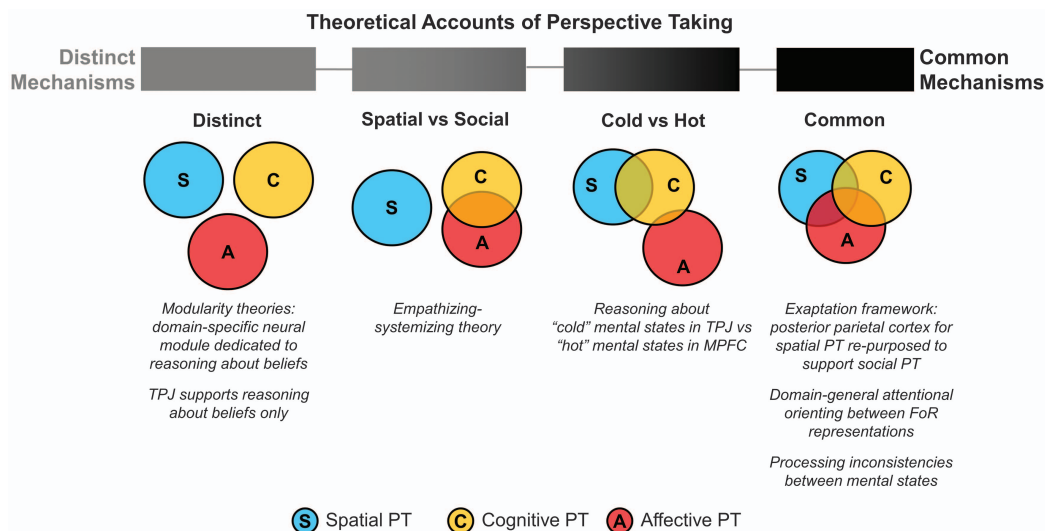
There are four views on the question of whether spatial, cognitive, and affective PT cohere to form a single dimension (see Figure 1). The first view, the “distinct” account, proposes distinct mechanisms for all three kinds of PT. For example, modularity theories regarding social PT claim that there is an innate domain-specific neural mechanism dedicated to social mental state reasoning, encompassing both cognitive and affective PT (e.g., Carruthers, 2013; Leslie et al., 2004). In line with this theory, some researchers hypothesize that the temporoparietal junction is associated with reasoning about beliefs (cognitive PT) but does not support other cognitive tasks (e.g., spatial PT; Kosakowski & Saxe, 2018; Saxe & Kanwisher, 2005; Saxe & Wexler, 2005). A second proposal, the “spatial versus social” account, is drawn in part from studies of individuals with autism spectrum disorder and aligns with Baron-Cohen’s (2009) empathizing-systemizing theory. In this view, cognitive and affective PT (social forms of PT) have distinct cognitive mechanisms from spatial PT. In support of this account, a recent neuroimaging meta-analysis found that cognitive and affective

PT showed large regions of overlap in functional activation, whereas both showed virtually no overlap with spatial PT (Brucato, 2022). A third proposal, the “cold versus hot” account, suggests that distinct mechanisms may support PT tasks that require representation of “cold” versus “hot” mental states (Aichhorn et al., 2006). In this view, spatial and cognitive PT are purported to involve reasoning about nonemotional or “cold” representations, commonly engaging the posterior end of the superior temporal sulcus and temporoparietal junction. Conversely, affective PT and, to a lesser extent, cognitive PT, which both require individuals to represent and predict emotional or “hot” mental states, are thought to rely additionally on neural mechanisms in the medial prefrontal cortex.

A competing set of theoretical and empirical articles propose a fourth, unitary, or “common” account in which spatial, cognitive, and affective PT all rely on a common set of cognitive processes (e.g., Arora et al., 2017; Parkinson & Wheatley, 2013; Quesque & Rossetti, 2020; Sun & Wang, 2014), albeit for varying reasons. Parkinson and Wheatley (2013), for instance, argued that evolutionarily older neural circuits used for reasoning about physical space were later repurposed for social cognition. Others postulate that spatial, cognitive, and affective PT are unified because they all require processing inconsistencies between mental states (e.g., Arora et al., 2017; Quesque & Rossetti, 2020; Schurz et al., 2013) or because each involves the manipulation of frame-of-reference representations, which leads to shared reliance on dorsal and ventral attentional networks (Sun & Wang, 2014). That is, attention orienting could be a domain-general mechanism that supports the comparison or shifting between alternative possible perspectives, regardless of domain.

With clinical populations, research shows a compelling pattern of codeficits among cognitive, affective, and spatial PT, favoring the unitary view. Autism spectrum disorder is characterized by difficulties

**Figure 1**  
*Schematic Overview of Theoretical Accounts of PT*



*Note.* The top bar gradient depicts how closely each account aligns with distinct mechanism accounts (lighter) or common mechanism accounts (darker). PT = perspective taking; FoR = frame-of-reference; TPJ = temporoparietal junction; MPFC = medial prefrontal cortex.

in understanding the mental states of others on measures of cognitive (Baron-Cohen, 2000; Yirmiya et al., 1998), affective (Song et al., 2019), and spatial PT (Hamilton et al., 2009). Further, cognitive PT has been found to significantly predict spatial PT performance for both autism spectrum disorder and typically developing children (Hamilton et al., 2009). Individuals with major depressive disorder are also known to have impairments in cognitive and affective PT, and greater symptom severity is associated with worsening cognitive and affective performance (Bora & Berk, 2016). Further evidence suggests that individuals with subclinical depression may also have codeficits in spatial PT (T. M. Erle et al., 2019).

The bulk of the research on PT dimensionality, however, involves neurotypical adults, and the evidence here is much more mixed (see Tables A1 and A2). Four early experiments assessed the association of self-reported affective PT and reaction time (RT) on a spatial PT task called the Own Body Transformation Test and found differing results (Gronholm et al., 2012; Mohr et al., 2010; Thakkar et al., 2009; Thakkar & Park, 2010). Two of these experiments found a significant correlation between self-reported affective PT and RT on the spatial PT task for female but not male participants; however, the effect was directionally opposite between the two studies (Mohr et al., 2010; Thakkar et al., 2009). Further, neither study found an association between self-reported affective PT and spatial PT accuracy (Mohr et al., 2010; Thakkar et al., 2009). Thakkar and Park (2010) later reported that more spatial PT errors were correlated with lower self-reported affective PT in female participants but failed to replicate RT results of their earlier study. Across four experiments, T. Erle and Topolinski (2015) found that greater spatial PT accuracy and quicker RT were significantly associated with higher self-reported affective PT in neurotypical adults. However, none of these studies measured objective affective PT performance. Meanwhile, relatively fewer studies have investigated the association of cognitive PT and spatial PT (Table A2). Two recent studies used objective measures of these two types of PT and found contrasting results, either indicating a significant association (Putko & Zlotogórska-Suwińska, 2019) or not (Job et al., 2021).

Importantly, most studies have not tested all forms of PT together in one sample. The fact that all three types of PT are not tested together limits our ability to use this evidence in distinguishing between common and distinct mechanism accounts of PT. Indeed, extant behavioral research on the question of common or distinct dimensions of PT, and the role of attention orienting in supporting PT across different domains, renders a murky picture. Broadly, work in adults produces inconsistent findings, sometimes pointing to shared processing and sometimes not.

Work conducted in children also fails to clarify the situation. One early review on children's PT abilities concluded that performance on PT tasks exhibits little to no between-domain correlations (Ford, 1979). However, more recent work has pointed to behavioral covariance across PT domains, including evidence for an association between cognitive and spatial PT in very young children (Hamilton et al., 2009; Tian et al., 2021; Viana et al., 2016) and evidence of significant associations between cognitive and affective PT in both younger (Bensalah et al., 2016) and older (Cassetta et al., 2018) children. Aslan and Köksal Akyol (2020) developed a PT measure for children designed to assess all three types of PT abilities in one test. The fact that the overall measure exhibited good internal consistency ( $\alpha = .71$ ) suggests that spatial, cognitive, and affective PT items in the test must have covaried

substantially in the sample (see also Tian et al., 2021), and thus, offers some support for a unified view of these abilities.

### Convergent Validity Within PT Domains

Many different assessments exist for each PT construct, and each has an unclear and largely untested relation to others, even within a given PT domain. Few studies have administered more than one task of the same type of PT to establish convergent validity either within or between PT domains, which is crucial when considering previously reported issues of within-domain convergent validity for cognitive and affective PT tasks (also termed "social PT" or "theory of mind" tasks; Ford, 1979; Gallant et al., 2020; Quesque & Rossetti, 2020; Warnell & Redcay, 2019). For example, prior reports investigating the construct validity of social PT tasks have found that there is none (i.e., Gallant et al., 2020; Warnell & Redcay, 2019). Social PT is a psychological construct defined as the ability to understand others' emotional or cognitive mental states and differentiate them from one's own, yet it has recently been argued that some of the most commonly used social PT tasks do not require PT at all (Quesque & Rossetti, 2020). Inclusion of tasks that do not meet the criterion as PT tasks may explain previous absences of convergent validity within PT domains.

Navarro (2022) conducted the first psychometric investigation of social PT tasks, including the Director Task, the Short Stories Questionnaire, and the Reading the Eyes in the Mind Task. There were small significant correlations among all tasks in the battery, but psychometric network analysis indicated that these social PT tasks did not represent a unified cluster. Specifically, the Reading the Eyes in the Mind Task did not share any significant edges with the Director Task. In some ways, this finding is unsurprising, since other work has concluded that the Reading the Eyes in the Mind Task is actually a measure of emotion recognition and not explicitly a PT task, insofar as it does not require representation of mental states or differentiation between one's own and another's perspective (e.g., Oakley et al., 2016; Quesque & Rossetti, 2020).

Intentional investigations into the psychometric properties and convergent validity of popular affective PT tasks in neurotypical adults are sparse. The Yoni Task (Shamay-Tsoory et al., 2007) has recently been shown to demonstrate strong psychometric properties, and affective PT trials of the Yoni Task significantly correlated with the Reading in the Minds Eyes Task ( $r = .29$ ), but not with a gender discrimination task (Isernia, Rossetto, Blasi, et al., 2023; Isernia, Rossetto, Shamay-Tsoory, et al., 2023). Given the claim that the Reading in the Minds Eyes Task may not require PT (Oakley et al., 2016; Quesque & Rossetti, 2020), though these findings are promising with respect to the validity of social skills measures, they do not necessarily provide information about the convergent validity of the Yoni Task with other measures of affective PT.

Meanwhile, in the spatial domain, PT is inconsistently tested as what Flavell et al. (1981) called Level 1 spatial PT (establishing a line of sight from another's eyes) and Level 2 spatial PT (working out the details of what another sees). Promisingly, a recent psychometric analysis of behavioral tasks used to measure Level 2 spatial PT reported that at least three of four commonly used behavioral measures of this construct show strong psychometric properties, including good item discriminability and internal reliability within tasks, convergent validity between tasks, and

divergent validity from other spatial skills like object rotation (Brucato et al., 2023).

In addition, for all three kinds of PT, there are both performance assessments and self-report measures, but the extent to which self-reported PT abilities reflect performance on tests is variable (Baksh et al., 2018; Ickes, 1993; Israelashvili et al., 2019; Murphy & Lilienfeld, 2019; Murray et al., 2017; Sunahara et al., 2022).

### Attentional Control and PT Performance

Despite assertions from researchers working within spatial, cognitive, and affective domains about the potential role of attentional control in supporting PT performance (Amorim, 2003; Wade et al., 2018), there are very few studies that directly test these claims. A small collection of behavioral studies conducted with adults and children have found positive associations between attention switching and Level 2 spatial PT (De Lillo & Ferguson, 2023), cognitive PT (Austin et al., 2014; M. R. Long et al., 2018), and affective PT abilities (Goodhew & Edwards, 2021). There are also several other dimensions of attentional control aside from shifting, including attentional orienting, which may contribute to cross-domain PT performance (Sun & Wang, 2014; Wade et al., 2018). However, little is known about whether behavioral measures of attentional control might play an explanatory role in the occasionally observed behavioral covariance of spatial, cognitive, or affective PT abilities.

### The Present Study

The present study attempted to bridge gaps in behavioral work to more adequately adjudicate between common and distinct mechanisms accounts of PT. The primary aim was to examine the behavioral associations among measures of spatial, cognitive, and affective PT in neurotypical young adults, controlling for general reasoning abilities. Secondly, if behavioral covariance was found, the study also aimed to investigate whether the association of PT abilities may be explained by superficial task features (i.e., spatiality or agency of stimuli) or, alternatively, shared attentional control processes.

### Method

#### Participants

Participants were 221 undergraduates at Temple University and adults living in the surrounding area (139 women, 71 men, 10 nonbinary or another gender, and one did not self-report gender). The mean sample age was 19.48 ( $SD = 3.22$  years) and ranged from 17 to 41 years. All participants had normal or corrected-to-normal vision and hearing. Listwise deletion was used to address missing data. Six participants had missing data on one task or questionnaire in the behavioral battery and were removed from the data set. Thirteen participants failed comprehension checks for at least one task and/or two or more attention checks in questionnaires and were thus also removed from the data set. Univariate and multivariate outlier analyses were also conducted. Specifically, seven participants were identified as univariate outliers as they had a mean score on at least one task that was 3 or greater standard deviations from the sample mean. Two multivariate outliers were also identified based on extreme Mahalanobis distance ( $p < .001$  for the  $\chi^2$  value). In total, nine outliers were removed from the data set. The final sample contained 193 participants (120 women, 63 men, and 10 nonbinary),

with a mean age of 19.44 years ( $SD = 3.07$  years). All participants gave informed consent before completing the protocol, which was approved by Temple University's Institutional Review Board. As compensation, participants received either partial course credit or a \$30 gift card following the study session.

### Procedure

Over a study session that lasted approximately 2 hr, participants completed six PT tasks, an attention task, a reasoning task, three self-report measures, and a demographic survey. The demographic survey asked participants to report their age and gender (options of "man," "woman," and "other," with an open text box to report another gender). After providing informed consent, participants were randomly assigned to one of 10 possible testing orders that counterbalanced tasks to avoid confounding due to task sequencing or fatigue effects. Each testing order was arranged such that PT tasks of the same domain did not directly follow one another, and each task was first and last in the behavioral testing battery exactly once. All measures were administered on a Windows 64-bit computer with an LCD monitor display. Up to two participants were tested at once in the testing room. Testing computers were on opposite sides of the same wall of the testing room with a separator in between, and participants wore headphones during the experiment so that any tasks that contained an audio stimulus did not disrupt the testing session for either participant.

### Materials

#### PT Tasks

Two behavioral tasks were selected to measure individual differences in PT in each domain. In the spatial domain, one task (Perspective-Taking Task for Adults) contained stimuli that resembled a human (i.e., an "agent") and the other did not (Spatial Orientation Task). These tasks were selected under the assumption that spatial PT tasks including a "social" agent might covary more substantially with social PT tasks (see Clements-Stephens et al., 2013; Shelton et al., 2012). Applying the same logic, one cognitive PT (Director Task) and one affective PT (Yoni Task) task were selected because they contained stimuli with more spatial features, and the alternative cognitive (Strange Stories Task) and affective PT (Empathic Accuracy Task) tasks were selected to have stimuli with fewer spatial features. All PT tasks included in the present study have previously been purported to involve mental state representation and require distinguishing between different perspectives, which are two criteria of PT tasks as defined by Quesque and Rossetti (2020).

**Spatial Orientation Task.** The Spatial Orientation Task (Friedman et al., 2020) was used as a measure of spatial PT. The task displayed a configuration of two-dimensional objects on the left side of the screen and a response circle on the right. At the start of each trial, participants were asked to imagine that they were standing at one object and facing a second object. Their imagined position was represented by a dot in the center of a response circle, and their imagined line of sight by a solid arrow pointing upward. Participants were then asked to imagine pointing toward a third object and to drag the arrow in the response circle to indicate their imagined pointing direction. Participants completed four practice items followed by 12 test items in a random order. The angular error between

the correct response and the participant's response was measured for each trial. Participants were allotted 5 min to complete the task. Participants received an angular error score of 90° for each item that was not answered in the allotted time. The total score was equal to the average angular error across all trials.

**Perspective-Taking Task for Adults.** The Perspective-Taking Task for Adults (Brucato et al., 2023) was used as another measure of spatial PT. Participants saw an image of a three-dimensional human-like figurine taking a photograph of an arrangement of three different-colored three-dimensional objects. Eight different perspectives of the arrangement were simultaneously provided below this image, and participants were asked to determine which one looked like the picture the figurine could have taken from where it was standing. There were three practice items followed by 32 test items, presented in a fixed quasirandom order. Participants were allotted 3 min to complete all test items. Four additional items, in which the perspectives of the figurine and participant were the same (first-person perspective), were used as comprehension checks for attentive responding, such that incorrect responses to two or more first-person perspective items resulted in exclusion from the analysis. The total score for the test was the number of correctly answered items per minute, excluding first-person perspective items.

**Empathic Accuracy Task.** The Empathic Accuracy Task (Mackes et al., 2018) was used as a measure of affective PT. Videos of individuals (targets) recounting emotional autobiographical events and those targets' emotional intensity ratings during each recollection were acquired from the Free Emotional Event Library (Mackes et al., 2018). Six video clips were presented to each participant: two happy, two sad, and two angry. One video of each emotion had a male target, and the other had a female target. Video clip presentation was counterbalanced using a balanced Latin square design. One of the six counterbalanced video clip orders was randomly assigned to each participant. While watching each clip, participants were asked to continuously rate the emotional intensity of the target on a 9-point scale (from 1 = *no emotion* to 9 = *very strong emotion*) using the keyboard. After each clip finished, participants were asked to select (a) which emotion the target felt most strongly (options of "happy," "angry," "surprised," "sad," "frightened," and "no emotion") and (b) which emotion they themselves felt most strongly. The analysis outlined in Mackes et al. (2018) was followed to derive empathic accuracy scores for each participant. First, ratings for each video were separated into 2-s bins, and one time-weighted average rating was calculated for each bin. Next, Pearson's correlation coefficients were calculated for the association between participants' and targets' ratings. The resulting Pearson's correlation coefficient for each video clip and each participant was then *r*-to-*Z* transformed to allow comparison between correlation coefficients.

**Yoni Task.** The Yoni Task (Shamay-Tsoory et al., 2007) was also used to measure affective PT. There were 98 test trials that showed a cartoon outline of a face (named "Yoni") surrounded by four pictures in each corner of the computer screen of objects belonging to a single category (e.g., fruits, chairs) or faces. Participants were asked to click on the correct image to which Yoni is referring as quickly and accurately as possible. The stimuli contained two types of cues that participants could use to make their responses: verbal only (i.e., a sentence that appeared at the top of the screen) or verbal cue with eye gaze. On each trial, Yoni's eyes looked at the correct answer (55 trials), straight ahead (32 trials), or at a misleading answer (11 trials). There were three main trial types: cognitive, affective, and physical.

The cognitive and the affective trials involved mental state inferences, whereas physical trials required a choice based on a physical attribute of the character. In addition, each trial type had two levels of difficulty: first order and second order. In first-order trials, participants were only required to understand Yoni's mental state (e.g., "Yoni loves <blank>"), whereas second-order trials required understanding the interaction between Yoni's mental state and each of the four characters around him (e.g., "Yoni loves the animal that <blank> loves"). Participants who scored less than 50% on the physical control trials were removed from analyses involving this task. Yoni accuracy scores have demonstrated stability over time, in contrast to task RTs (Isernia et al., 2024). Thus, the present study used the accuracy of affective trials in all analyses, with correct trials summed to get an affective PT score ranging from 0 to 48.

**Strange Stories Task.** The revised Strange Stories Task (White et al., 2009) was used to measure cognitive PT. During each trial, a short story or a series of unlinked sentences appeared on the screen. After reading the text, participants pressed the space bar to receive a question. On mental state trials, participants were asked a question that required them to reason about the mental state of characters in a short story, and on unlinked sentences trials, participants were asked to report a fact from one of the unlinked sentences. There were eight mental state trials and eight unlinked sentences trials, presented in a randomized order. Participants were instructed to wait until they thought of the correct answer and then to press the space bar to advance. They could then type their response into a text box on the screen and hit the "Enter" key to advance to the next trial. RT was recorded during the think period for each trial. Typed responses to each answer were then coded for accuracy by two independent raters on a scale of 0–2 based on the scoring paradigm from White et al. (2009). Following White and colleagues' protocol, 20% of the scores selected randomly for each story were used in an intraclass correlation analysis to determine the level of agreement between raters, and good agreement was reached (intraclass correlation coefficient = .83). Participants who scored less than 50% on the unlinked sentences stories were removed from analyses involving this task. Accuracy on mental state story trials was summed for a cognitive PT score ranging from 0 to 16.

**Director Task.** The Director Task (Dumontheil et al., 2010) was also used to measure cognitive PT. The stimuli showed shelves (four slots high and four slots wide) with a director standing behind them. Eight objects (some were different types of the same kind of object: e.g., golf ball, tennis ball, basketball) were shown in slots of the shelves, and five slots were occluded such that the participant but not the director could see their contents. The participant was asked to follow the instructions given by the director. For each trial, the director instructed the participant to move one of the eight objects by one slot (up, down, left, or right). Using a computer mouse, participants were required to click on the object they thought the director was referring to and to drag it into the appropriate slot on the shelves. Participants were given 3.6 s to make their response after each instruction.

There was one practice block followed by 16 test blocks, each made up of three trials. There were three types of trials. On filler trials, the director's instructions only referred to objects in non-occluded slots (i.e., visible to both director and participant). There were two filler trials per block, and the third trial was either an experimental or a control trial. On experimental trials, the director instructed participants to move an object for which there were

several types of the same kind on the shelves, with one of them in an occluded slot. For example, on one experimental trial, there were three balls of different sizes in different slots on the shelves, two of which were visible to the director. The director instructed participants to “move the small ball up.” The smallest ball of the three was in an occluded slot of the shelves, so the participant was required to consider the director’s perspective and move the smallest ball that the director could see, rather than the smallest ball that the participant themselves could see. Each experimental trial had a matched control trial in a different block wherein the objects on the shelves were identical to the experimental trial except that there was no longer an object of the same type in the occluded slot. The order of the 32 filler, eight control, and eight experimental trials was counterbalanced between participants.

Participants were not required to imagine what the director saw in its specific orientation from his vantage point to answer correctly (i.e., Level 2 spatial PT); rather, the task likely required establishing a line of sight from the director’s eyes to the object on the open versus occluded shelves (Level 1 spatial PT) to inform what the director believed was the smallest or largest object of its type on the shelves (his cognitive perspective), and then, the participant must differentiate between their own belief and the director’s to correctly infer his instructions’ meaning. A participant was removed from the analysis if they did not attempt a response on any experimental trials. Total scores were the mean accuracy across all eight experimental trials.

### Attention and Reasoning Tasks

**Attentional Network Test.** The Attentional Network Test (Fan et al., 2002) was used to measure multiple aspects of attentional control. Each trial of the task began with an initial fixation period for 400–1,600 ms (black fixation cross centered on a gray screen), followed by a warning cue that was presented for 100 ms. There were four possible kinds of warning cues: (a) an asterisk in place of the fixation cross, (b) two asterisks—one on the top and one on the bottom of the fixation cross, (c) one asterisk above the fixation cross, and (d) one asterisk below the fixation cross. The first two types of warning cues did not give information about the location of an upcoming target, whereas the latter two types of warning cues indicated the location of the upcoming target (above or below the fixation cross, respectively). The target was a black arrow pointing left or right that appeared either above or below the fixation cross. The target was flanked on both sides by two arrows that pointed in the same direction (congruent condition), by two arrows that pointed in the opposite direction (incongruent condition), or by two parallel lines (neutral condition). Participants were asked to indicate the direction of the target arrow by pressing one of two designated keys, using their left and right index fingers. Participants were asked to respond as quickly and accurately as possible to each trial. There were two blocks of 12 practice trials with feedback for accuracy. The first practice block did not have cueing, and the second practice block did have cueing. After the practice, participants completed four test blocks of 48 task trials comprising 12 trials of each cue condition.

Accuracy and RT data were collected from each block. Participants who scored lower than 50% accuracy on two or more blocks were excluded from analyses involving this task. Outlier detection for RTs was conducted within participants, such that trial RTs falling 2

standard deviations above or below the participant’s mean RT were excluded. Indices for attentional orienting and executive attention were calculated according to Fan et al. (2002), with the attentional orienting effect given by (mean RT on trials with a center cue) – (mean RT on trials with a spatial cue) and an executive attention effect given by (mean RT on trials with an incongruent flanker) – (mean RT on trials with a congruent flanker).

**Matrix Reasoning Task.** The color-blind-friendly version of the Matrix Reasoning Task (Chierchia et al., 2019) was used as a control measure of general reasoning ability. The task showed a 3 × 3 matrix on the computer screen. Eight of the nine cells contained a colored abstract shape, and the bottom right cell of the matrix was empty. Participants were asked to complete the matrix by choosing one of four shape options displayed on the bottom of the screen. The shape characteristics varied in shape, color, size, and location. Each trial began with a 500-ms fixation cross, followed by a 100-ms white screen, and participants had a time limit of 30 s to complete each trial. A time warning was given 25 s after the matrix was presented to indicate 5 s remained in the trial. There were three easy practice trials followed by 8 min of test items presented in a fixed order. All items in practice and test had feedback on accuracy. There were 80 possible items in the whole set, but participants were not required nor expected to complete all items in the allotted time. If a participant completed 80 items before the testing time limit, the items were presented again in the same order until 8 min was reached, but responses for repeated items were not analyzed. The total Matrix Reasoning Task score was the number of correctly completed items per minute.

### Self-Report Questionnaires

**The Object–Spatial Imagery Questionnaire.** The Object–Spatial Imagery Questionnaire (Blajenkova et al., 2006) is a 30-item questionnaire with two subscales that assess individual differences in visual imagery preferences and experiences for (a) objects and (b) spatial relations. Representing and processing spatial relations and transformations among objects are two important skills for spatial PT (Kozhevnikov & Hegarty, 2001). Thus, the present study administered the 15 items of the spatial relations subscale as a measure of self-reported spatial PT. The spatial relations subscale demonstrated good internal reliability previously ( $\alpha = .79$ ; Blajenkova et al., 2006) and in the present study ( $\alpha = .79$ ). Participants were asked to read each statement and rate how strongly they agreed or disagreed on a 5-point scale (*totally disagree* to *totally agree*). Total scores were obtained by averaging across the 15 items of the spatial subscale and therefore could range from 1 to 5. Higher scores indicated greater preference and experience for spatial-relational imagery.

**Questionnaire of Cognitive and Affective Empathy.** The Questionnaire of Cognitive and Affective Empathy (Reniers et al., 2011) is a 31-item questionnaire composed of items from four previously validated empathy questionnaires (Baron-Cohen & Wheelwright, 2004; Davis, 1983; Eysenck & Eysenck, 1978; Hogan, 1969). It has two subscales that measure cognitive empathy and three subscales that measure affective empathy. The authors defined cognitive empathy as comprehension of other people’s experience that requires representing, holding, and manipulating information in mind, whereas affective empathy is “the ability to vicariously experience the emotional experience of others” (Reniers et al., 2011). Thus, the two subscales that made up the cognitive empathy facet were used, as this is the construct of most

relevance to the present study. The two subscales were PT and online simulation, and they contained 10 and nine items, respectively. The Questionnaire of Cognitive and Affective Empathy cognitive empathy scale has demonstrated good internal reliability previously (PT subscale:  $\alpha = .85$ , online simulation subscale:  $\alpha = .83$ , Reniers et al., 2011; total cognitive empathy scale:  $\alpha = .85$ , Gomez et al., 2022) and in the present study (PT subscale:  $\alpha = .82$ , online simulation subscale:  $\alpha = .84$ , total cognitive empathy scale:  $\alpha = .86$ ). Participants were asked to read each statement and indicate their agreement or disagreement on a 4-point scale (*strongly disagree* to *strongly agree*). Ratings for each item were summed for each subscale to get subscale scores. Total cognitive empathy scores of the Questionnaire of Cognitive and Affective Empathy were found by summing the PT and online simulation subscales. Thus, total scores could range from 19 to 76, with higher scores indicating greater self-reported affective PT.

**The Mentalization Scale.** The Mentalization Scale (Dimitrijević et al., 2018) is a 28-item questionnaire with three subscales: self-related mentalization, other-related mentalization, and motivation to mentalize. In the present study, the 10-item other-related mentalization subscale was used as a self-report indicator of cognitive PT. The other-related mentalization subscale demonstrated good internal reliability previously ( $\alpha = .77$ ; Dimitrijević et al., 2018) and in the present study ( $\alpha = .70$ ). Participants were asked to read each of the 10 statements and rate how correct each statement was about themselves on a 5-point scale (*completely incorrect* to *completely correct*). The total score of the Mentalization Scale other-related mentalization subscale could range from 10 to 50, and a higher score indicates a better ability to mentalize about others.

### Transparency and Openness

This study's design, variables, and analysis plan were preregistered on the Open Science Framework ([https://osf.io/3fxwb/?view\\_only=56ad8b5ca2fe4f81ba11bddac6f8e5aa](https://osf.io/3fxwb/?view_only=56ad8b5ca2fe4f81ba11bddac6f8e5aa); Brucato et al., 2021). Our preregistration included a progressive series of hypotheses, wherein Hypotheses 2 and 3 were contingent on Hypothesis 1. Results of the present study did not provide evidence in support of Hypothesis 1 (positive correlation between cognitive, affective, and spatial PT), and we therefore could not pursue subsequent proposed structural equation modeling analyses related to Hypothesis 2 (association of

attention with a general PT factor) or Hypothesis 3 (association of general PT factor with self-reported clinical symptomatology).

Consistent with the Journal Article Reporting Standards, we report how we determined our sample size, data exclusions, and all measures in the study (Appelbaum et al., 2018). Data and analysis scripts are also available on the Open Science Framework ([https://osf.io/xjqcv/?view\\_only=f1009092b5a447878675b4fd3a14c3ae](https://osf.io/xjqcv/?view_only=f1009092b5a447878675b4fd3a14c3ae); Brucato et al., 2025). An a priori power analysis using G power (Faul et al., 2009) indicated that a sample of 211 participants would enable a high level of power (.90) to detect a small effect size ( $r = .2$ ) using the standard error probability ( $\alpha = .05$ ). We retained 193 of 211 participants after listwise deletion for missing data, comprehension check failures, and univariate and multivariate outliers. Post hoc power analysis indicated that the final sample maintained a large level of power (.88) to detect the same effect size.

### Results

Reported results reflect analyses conducted on the accuracy of PT and reasoning tasks, summary scores from self-report measures, and RT indices of attention. Descriptive statistics for all measures are listed in Table 1 and by gender in Table 2. Correlations using task RT data are available in Table A5 (c.f. Table A4 for descriptive statistics of RT data) and are consistent with results on accuracy data.

### Comparison of Performance on Control and Experimental Trials

For tasks that provided a control condition, performance was compared between experimental and control trials to check that typically expected behavioral patterns were observed with the present sample. Perspective-Taking Task for Adults performance was assessed to determine if participants were quicker to respond to trials with 0° angular disparity (first-person perspective control trials) than those with 45°, 90°, 135°, and 180° angular disparities, including only trials for which a correct response was chosen. Paired-samples *t* tests indicated that the mean RT for correct 0° angular disparity trials ( $M = 8.66$  s,  $SD = 3.24$  s) was significantly faster than the mean RTs for correct trials of each other angular disparity ( $M_{45} = 10.80$  s,

**Table 1**  
*Descriptive Statistics for All Behavioral and Self-Report Measures*

| Construct         | Measure                                    | <i>n</i> | <i>M</i> | <i>SD</i> | Min.   | Max.   | Skew  |
|-------------------|--------------------------------------------|----------|----------|-----------|--------|--------|-------|
| Attention         | ANT—orienting index                        | 193      | 30.69    | 22.74     | −47.71 | 107.56 | 0.1   |
| Attention         | ANT—executing index                        | 193      | 56.11    | 24.45     | −23.08 | 128.74 | −0.07 |
| General reasoning | MaRs—number of correct items per minute    | 193      | 3.29     | 0.7       | 1.75   | 5.25   | 0.34  |
| Spatial PT        | SOT—average angular error                  | 193      | 43.18    | 25.18     | 6.67   | 113.08 | 0.52  |
| Spatial PT        | PTT-A—number of correct items per minute   | 193      | 3.95     | 1.44      | 0.67   | 9      | 0.41  |
| Spatial PT        | OSIQ—self-report                           | 193      | 2.48     | 0.58      | 1.07   | 3.93   | 0.15  |
| Cognitive PT      | Strange Stories—mental state stories score | 193      | 12.31    | 2.41      | 3      | 16     | −0.72 |
| Cognitive PT      | Director Task—mean experimental accuracy   | 193      | 0.46     | 0.35      | 0      | 1      | 0.26  |
| Cognitive PT      | MentS—self-report                          | 193      | 40.39    | 4.02      | 26     | 50     | −0.49 |
| Affective PT      | EAT—z-scored empathic accuracy             | 193      | 0.91     | 0.22      | 0.13   | 1.41   | −0.5  |
| Affective PT      | Yoni—number of correct affective trials    | 178      | 41.75    | 4.85      | 24     | 48     | −1.11 |
| Affective PT      | QCAE—self-report                           | 193      | 59.24    | 7.86      | 34     | 74     | −0.53 |

*Note.* ANT = Attentional Network Test; MaRs = Matrix Reasoning; SOT = Spatial Orientation Test; PTT-A = Perspective-Taking Task for Adults; EAT = Empathic Accuracy Task; OSIQ = Object-Spatial Imagery Questionnaire; QCAE = Questionnaire of Cognitive and Affective Empathy; MentS = Mentalization Scale; PT = perspective taking; Min. = minimum; Max. = maximum.

$SD_{45} = 4.14$  s;  $M_{90} = 10.99$  s,  $SD_{90} = 3.93$  s;  $M_{135} = 11.74$  s,  $SD_{135} = 5.34$ ;  $M_{180} = 11.82$  s,  $SD_{180} = 6.19$  s;  $ps < .01$ ; see Figure A1 and Table A3 in the Appendix).

Accuracy on Yoni physical trials was compared with accuracy on Yoni affective PT trials. Notably, 15 participants scored at chance level or lower on Yoni physical accuracy trials but were not identified as outliers and did not fail any other comprehension checks in the experimental battery. These participants were not included in any analyses that involved Yoni Task performance but were retained in the data set for inclusion in other analyses. Participants who passed the Yoni physical trials were significantly more accurate on physical ( $M = 0.92$ ,  $SD = 0.11$ ) than affective PT trials ( $M = 0.87$ ,  $SD = 0.10$ ) according to a paired-samples  $t$  test,  $t(177) = 5.30$ ,  $p < .001$ . Responses were also faster for physical ( $M = 2,616.03$  ms,  $SD = 690.44$  ms) trials than for affective PT trials ( $M = 3,728.51$  ms,  $SD = 981.45$  ms),  $t(177) = -22.22$ ,  $p < .001$ .

Mean scores of mental state stories and unlinked sentences trials were compared for the Strange Stories Task. A paired-samples  $t$  test indicated that participants scored significantly higher on unlinked sentences trials ( $M = 1.94$ ,  $SD = 0.13$ ) than on mental state stories trials ( $M = 1.54$ ,  $SD = 0.30$ ),  $t(192) = 17.23$ ,  $p < .001$ .

To compare the effect of trial type (i.e., filler, control, and experimental) on accuracy for the Director Task, a one-way within-subjects analysis of variance test was used. There was a significant effect of trial type on accuracy,  $F(2, 220) = 330.68$ ,  $p < .0001$ ,  $\eta^2(g) = 0.52$ . Post hoc pairwise comparisons indicated that all pairwise comparisons were significant. Mean accuracy for experimental trials ( $M = 0.46$ ,  $SD = 0.35$ ) was significantly lower than mean accuracy for filler trials ( $M = 0.95$ ,  $SD = 0.04$ ),  $t(192) = -19.4$ ,  $p < .001$ , and mean accuracy for control trials ( $M = 0.90$ ,  $SD = 0.10$ ),  $t(192) = -17.4$ ,  $p < .001$ . In addition, participants were significantly more accurate on filler than control trials,  $t(192) = 6.4$ ,  $p < .001$  (see also Figure A2 in the Appendix).

Overall, comparison of performance on experimental and control trials for the tasks used in the present study aligned with expectations based on their original use (i.e., better performance or faster response times in control as opposed to experimental trials).

## Gender Differences in Performance

Group differences in performance by self-reported gender were investigated only for participants who identified as a man or woman, as there were too few nonbinary participants ( $n = 10$ ) to include for comparison. Descriptive statistics for men and women participants are listed in Table 3. Mean differences for all measures and results of independent-samples  $t$  tests between genders are listed in Table 4. Men significantly outperformed women on the two behavioral measures and one self-report measure of spatial PT ( $ps < .001$ ), and the effect of gender was medium to large (see Table 4). Men also had significantly higher accuracy on the Director Task than female participants, with a small effect size ( $d = -0.33$ ; see Table 4). Conversely, women had higher Attentional Network Test orienting indices than men, again with a small effect size ( $d = 0.32$ ). There were no significant gender differences on any other measures.

## Convergent Validity of PT Within Domains

Convergent validity within PT domains was investigated by conducting Pearson's correlations between each task and self-report measure within each domain (see Table A5). The two spatial PT tasks had a highly significant correlation with each other ( $r = -.5$ ,  $p < .001$ ), and both were correlated with the Object-Spatial Imagery Questionnaire self-report measure of spatial PT ( $ps < .001$ ). Cognitive PT tasks had a small correlation that did not survive false discovery rate (FDR) correction for multiple comparisons, and neither task correlated with the Mentalization Scale self-report subscale for other-related mentalization ( $ps < .05$ ). Finally, the two affective PT tasks were not significantly correlated with each other, nor was either task correlated with the Questionnaire of Cognitive and Affective Empathy self-report of cognitive empathy ( $ps > .05$ ).

## Association of PT Between Domains

Pearson's correlations of cognitive PT (Strange Stories, mean mental state stories score; Director Task, mean accuracy on experimental trials) and affective PT (Empathic Accuracy Task,  $z$ -scored mean empathic accuracy; Yoni, number of correct affective trials)

**Table 2**  
*Descriptive Statistics by Self-Reported Gender*

| Construct         | Measure                        | Woman    |          |           |        |        | Man      |          |           |        |        |
|-------------------|--------------------------------|----------|----------|-----------|--------|--------|----------|----------|-----------|--------|--------|
|                   |                                | <i>n</i> | <i>M</i> | <i>SD</i> | Min.   | Max.   | <i>n</i> | <i>M</i> | <i>SD</i> | Min.   | Max.   |
| Attention         | ANT—orienting index            | 120      | 33.23    | 22.97     | -30.52 | 107.56 | 63       | 25.97    | 21.82     | -47.71 | 74.11  |
| Attention         | ANT—executing index            | 120      | 54.7     | 25.95     | -23.08 | 114.88 | 63       | 59.36    | 19.65     | 15.72  | 128.74 |
| General reasoning | MaRs—correct items per minute  | 120      | 3.29     | 0.66      | 1.75   | 5.25   | 63       | 3.27     | 0.76      | 2      | 5      |
| Spatial PT        | SOT—average angular error      | 120      | 50.22    | 23.88     | 7.42   | 108.33 | 63       | 30.84    | 23.44     | 6.67   | 113.08 |
| Spatial PT        | PTT-A—correct items per minute | 120      | 3.6      | 1.33      | 0.67   | 8      | 63       | 4.6      | 1.42      | 1.33   | 9      |
| Spatial PT        | OSIQ—self-report               | 120      | 2.3      | 0.53      | 1.07   | 3.93   | 63       | 2.86     | 0.51      | 1.8    | 3.87   |
| Cognitive PT      | Strange Stories—score          | 120      | 12.25    | 2.45      | 3      | 16     | 63       | 12.32    | 2.38      | 6      | 16     |
| Cognitive PT      | Director Task—accuracy         | 120      | 0.42     | 0.32      | 0      | 1      | 63       | 0.53     | 0.38      | 0      | 1      |
| Cognitive PT      | MentS—self-report              | 120      | 40.62    | 3.79      | 29     | 49     | 63       | 39.79    | 4.31      | 26     | 49     |
| Affective PT      | EAT—empathic accuracy          | 120      | 0.9      | 0.23      | 0.13   | 1.36   | 63       | 0.93     | 0.2       | 0.34   | 1.41   |
| Affective PT      | Yoni—correct affective trials  | 110      | 41.75    | 4.21      | 29     | 48     | 59       | 41.56    | 5.97      | 24     | 48     |
| Affective PT      | QCAE—self-report               | 120      | 59.7     | 7.56      | 34     | 74     | 63       | 58.46    | 8.44      | 39     | 73     |

*Note.* ANT = Attentional Network Test; MaRs = Matrix Reasoning; SOT = Spatial Orientation Test; PTT-A = Perspective-Taking Task for Adults; EAT = Empathic Accuracy Task; OSIQ = Object-Spatial Imagery Questionnaire; QCAE = Questionnaire of Cognitive and Affective Empathy; MentS = Mentalization Scale; PT = perspective taking; Min. = minimum; Max. = maximum.

**Table 3**  
*Group Differences by Self-Reported Gender*

| Construct         | Measure                        | Mean difference | df  | <i>p</i> | Cohen's <i>d</i> | Confidence interval |
|-------------------|--------------------------------|-----------------|-----|----------|------------------|---------------------|
| Attention         | ANT—orienting index            | 7.26            | 181 | .04      | 0.32             | [0.02, 0.63]        |
| Attention         | ANT—executing index            | −4.66           | 181 | .21      |                  |                     |
| General reasoning | MaRs—correct items per minute  | 0.02            | 181 | .87      |                  |                     |
| Spatial PT        | SOT—average angular error      | 19.38           | 181 | <.001    | 0.82             | [0.50, 1.14]        |
| Spatial PT        | PTT-A—correct items per minute | −1              | 181 | <.001    | −0.74            | [−1.05, −0.43]      |
| Spatial PT        | OSIQ—self-report               | −0.56           | 181 | <.001    | −1.09            | [−1.41, −0.76]      |
| Cognitive PT      | Strange Stories—score          | −0.07           | 181 | .86      |                  |                     |
| Cognitive PT      | Director Task—accuracy         | −0.11           | 181 | .04      | −0.33            | [−0.64, −0.02]      |
| Cognitive PT      | MentS—self-report              | 0.83            | 181 | .18      |                  |                     |
| Affective PT      | EAT—empathic accuracy          | −0.03           | 181 | .25      |                  |                     |
| Affective PT      | Yoni—correct affective trials  | 0.19            | 167 | .52      |                  |                     |
| Affective PT      | QCAE—self-report               | 1.24            | 181 | .31      |                  |                     |

*Note.* Degrees of freedom and *p* values are reported for independent-samples *t* tests comparing performance of each measure between men and women participants. Significant differences are accompanied by Cohen's *d* effect size and a 95% confidence interval. Mean difference = women's mean score − men's mean score. ANT = Attentional Network Test; MaRs = Matrix Reasoning; SOT = Spatial Orientation Test; PTT-A = Perspective-Taking Task for Adults; EAT = Empathic Accuracy Task; OSIQ = Object-Spatial Imagery Questionnaire; QCAE = Questionnaire of Cognitive and Affective Empathy; MentS = Mentalization Scale; PT = perspective taking.

tasks resulted in no significant correlations between the two domains (see Table A5). However, self-report measures of cognitive PT (Mentalization Scale, other-related mentalization subscale) and affective PT (Questionnaire of Cognitive and Affective Empathy, cognitive empathy scale) were strongly correlated, even after removing the effect of the Matrix Reasoning Task ( $r = .7, p < .001$ ). The Mentalization Scale other-related mentalization subscale and the Questionnaire of Cognitive and Affective Empathy cognitive empathy scale were not correlated with cognitive nor affective PT tasks.

Pearson's correlations of cognitive and spatial PT (Spatial Orientation Task, angular error; Perspective-Taking Task for Adults, number of correct items per minute) behavioral tasks also indicated no significant associations (see Table A5). There was a small significant correlation between mean accuracy on the Director Task and self-reported spatial PT as assessed by the Object-Spatial Imagery Questionnaire spatial subscale ( $r = .3, p < .001$ ). The correlation between the Director Task and the Object-Spatial Imagery

Questionnaire spatial subscale remained significant after removing the effect of the Matrix Reasoning Task and FDR correction for multiple comparisons ( $r = .25, p < .01$ ). RT on correct experimental trials of the Director Task was not associated with Object-Spatial Imagery Questionnaire spatial subscale scores (cf. Table A5). No significant correlations were found between RT on the Perspective-Taking Task for Adults and cognitive PT tasks ( $ps < .05$ ).

Pearson's correlations of affective and spatial PT tasks resulted in two small correlations between Spatial Orientation Task angular error and performance on the two affective PT tasks. The correlation between the Spatial Orientation Task and Empathic Accuracy Task *z*-scored mean empathic accuracy remained significant after controlling for the Matrix Reasoning Task and FDR correction for multiple comparisons ( $r = -.27, p < .01$ ). The number of correct items per minute on the Perspective-Taking Task for Adults was not significantly correlated with performance on the Yoni Task nor the Empathic Accuracy Task.

**Table 4**  
*Results of Pearson's Correlations Between All Measures*

| Measure        | Construct         | 1            | 2           | 3          | 4          | 5   | 6           | 7  | 8  | 9   | 10  | 11  | 12 |
|----------------|-------------------|--------------|-------------|------------|------------|-----|-------------|----|----|-----|-----|-----|----|
| 1. SOT (error) | Spatial PT        | —            |             |            |            |     |             |    |    |     |     |     |    |
| 2. PTT-A       | Spatial PT        | <b>−.5**</b> | —           |            |            |     |             |    |    |     |     |     |    |
| 3. OSIQ        | Spatial PT        | <b>−.3**</b> | <b>.4**</b> | —          |            |     |             |    |    |     |     |     |    |
| 4. Director    | Cognitive PT      | −.1          | .1          | <b>.3*</b> | —          |     |             |    |    |     |     |     |    |
| 5. Stories     | Cognitive PT      | −.1          | .0          | .0         | <b>.2*</b> | —   |             |    |    |     |     |     |    |
| 6. MentS       | Cognitive PT      | .0           | .0          | .1         | .1         | −.1 | —           |    |    |     |     |     |    |
| 7. EAT         | Affective PT      | <b>−.3**</b> | .1          | .1         | .0         | .0  | .0          | —  |    |     |     |     |    |
| 8. Yoni        | Affective PT      | −.2          | .1          | .0         | .0         | .0  | .1          | .0 | —  |     |     |     |    |
| 9. QCAE        | Affective PT      | .0           | .0          | .2         | .0         | .0  | <b>.7**</b> | .0 | .1 | —   |     |     |    |
| 10. Orienting  | Attention         | .0           | −.1         | .0         | .1         | .1  | .1          | .1 | .1 | .2  | —   |     |    |
| 11. Executing  | Attention         | .0           | .0          | .1         | .0         | −.1 | .1          | .2 | .0 | .0  | −.1 | —   |    |
| 12. MaRs       | General reasoning | −.1          | <b>.3**</b> | .1         | −.1        | .1  | .0          | .1 | .1 | −.1 | .0  | −.1 | —  |

*Note.* Statistically significant correlations appear in bold. False discovery rate corrections were applied to adjust for multiple comparisons.  $n = 178$  for correlations with Yoni.  $n = 193$  for all other correlations. MaRs = Matrix Reasoning; SOT = Spatial Orientation Test; PTT-A = Perspective-Taking Task for Adults; EAT = Empathic Accuracy Task; OSIQ = Object-Spatial Imagery Questionnaire; QCAE = Questionnaire of Cognitive and Affective Empathy; MentS = Mentalization Scale; PT = perspective taking.

\* $p < .05$ . \*\* $p < .001$ .

### Association of PT With Attention

The Attentional Network Test executing and orienting indices had no significant correlations with any PT measures in the experiment following FDR correction. Attentional Network Test executing and orienting indices were not significantly correlated with each other, which is in alignment with original reports from this paradigm (Fan et al., 2002).

### Discussion

The fundamental aim of this study was to test the behavioral covariance of spatial, cognitive, and affective PT as a way to adjudicate between common and distinct mechanism accounts. Pairwise correlations between PT measures of each domain, controlling for general reasoning ability, showed no significant correlations between spatial and cognitive PT tasks, or cognitive and affective PT tasks, and only a very small (although significant) correlation between one spatial PT task (Spatial Orientation Task) and one affective PT task (Empathic Accuracy Task). However, the affective PT tasks were not correlated with each other. We also tested one version of a common mechanism approach by investigating if attentional control or task stimulus features might explain behavioral covariance among PT tasks. Attentional orienting and executing were not significantly correlated with any PT tasks.

Overall, this pattern of results is most consonant with two of the distinct mechanisms accounts of PT abilities: the “distinct” account and the “spatial versus social” account. First, no significant correlations were found between spatial PT and cognitive PT tasks in the present study. This result stands in contrast to a recent study, which found that RT on a novel spatial PT task was significantly correlated with Strange Stories Task performance (Putko & Zlotogórska-Suwińska, 2019). One explanation for the apparent inconsistency of these findings relates to the validity of tasks used to measure spatial PT. Putko and Zlotogórska-Suwińska (2019) devised a novel spatial PT task for use in their study. Despite its face validity as a measure of spatial PT, no psychometric analyses were provided to support its construct validity. Notably, the authors chose to use RTs in place of accuracy in their correlation analyses due to ceiling effects in their sample. This may suggest that their measure lacked items of sufficient difficulty to capture a range of individual differences in spatial PT ability, and leaves to question what psychological construct their RT data is an indicator of, especially as response times to each question averaged less than 1 s. Conversely, the spatial PT tasks used in the present study have previously demonstrated wide ranges of item difficulty and high discriminability, high convergent validity with other spatial PT tasks, divergent validity from other closely related constructs like object rotation (Brucato et al., 2023), and additionally demonstrated convergent validity in the present study.

Second, self-reported affective PT was uncorrelated with spatial PT performance. Several previous studies have found that self-reported affective PT as measured by the Interpersonal Reactivity Index (Davis, 1983) or Empathy Quotient (Baron-Cohen & Wheelwright, 2004) was significantly associated with greater accuracy or quicker RTs on spatial PT tasks (T. Erle & Topolinski, 2015; Gronholm et al., 2012; Mohr et al., 2010; Sulpizio et al., 2015; Thakkar & Park, 2010; Tomei et al., 2017). In contrast to previous studies, the Questionnaire of Cognitive and Affective Empathy cognitive empathy scale was not significantly associated with objective performance on either spatial

PT task in the present study. This is particularly surprising because the Questionnaire of Cognitive and Affective Empathy cognitive empathy scale contains items from both the Interpersonal Reactivity Index and the Empathy Quotient. However, the Questionnaire of Cognitive and Affective Empathy was created by conducting a principal components analysis of items from four individually validated empathy questionnaires, and only the items that successfully loaded onto one of its five factors were retained (Reniers et al., 2011). As a result, only five out of 14 original items measuring cognitive empathy on the Interpersonal Reactivity Index were considered valid items according to Reniers et al.’s (2011) analysis. Therefore, the results of this study suggest that when using a self-report measure of cognitive empathy that only included the most convergently valid items from previous questionnaires, self-reported affective PT is no longer associated with spatial PT.

Third, self-reported mentalizing was also uncorrelated with spatial PT performance. A handful of previous studies have found associations between self-reported autistic traits and “social ineptitude” scores as measured via the Autism Spectrum Quotient and performance on spatial PT tasks that have agents in their stimuli (Brunyé et al., 2012; Clements-Stephens et al., 2013; Shelton et al., 2012) but not with spatial PT tasks without agents (Clements-Stephens et al., 2013; Job et al., 2021). In the present study, we aimed to directly test the association of spatial PT with self-reported cognitive PT via the Mentalization Scale other-related mentalization subscale (Dimitrijević et al., 2018; Richter et al., 2021) rather than indirectly via self-reported autistic symptomatology alone. Given that the Mentalization Scale other-related mentalization subscale was not associated with the Perspective-Taking Task for Adults or the Spatial Orientation Task, results demonstrate that objective spatial PT performance on a task with or without agents is uncorrelated with self-reported cognitive PT abilities.

Fourth, despite the high correlation between self-reported cognitive and affective PT measures used in the present study, no significant correlations were found between the cognitive and affective PT tasks. A brief review of behavioral correlations among various affective and cognitive PT tasks provided in the introduction reported inconsistent results regarding their behavioral associations. However, many of these correlational studies were conducted between cognitive PT tasks (e.g., Strange Stories Task) and emotion recognition or affective empathy tasks and, therefore, do not necessarily speak to their association with affective PT tasks. Future studies that directly compare behavioral performance on validated cognitive and affective PT tasks may be useful for understanding the association of these abilities.

An alternative explanation for the high correlations between self-report measures may be that each instrument includes a blend of items assessing both cognitive and affective PT, despite scale titles that suggest a greater distinction. For example, the Questionnaire of Cognitive and Affective Empathy’s cognitive empathy subscale includes items with face validity for affective PT (e.g., “I am good at predicting how someone will feel”), and the Mentalization Scale’s other-related mentalization subscale includes items that show face validity for cognitive PT (e.g., “people tell me that I understand them and give them sound advice”). However, the affective PT measure also included items with face validity for cognitive PT (“other people tell me I am good at understanding how they are feeling and what they are thinking”), and vice versa, the measure intended to capture cognitive PT also included items with face validity for affective PT (“sometimes I can understand someone’s feelings

before s/he tells me anything”). Still, each subscale demonstrated high internal consistency, suggesting that items shared covariance. Given prior evidence that self-reports may be poor proxies for behaviorally assessed abilities (e.g., Sunahara et al., 2022), the present results do not provide strong evidence to support accounts of social PT as a unified construct.

The conclusion that spatial PT is distinct leaves open to question why codeficits of each type of PT have been documented in an array of clinical populations (Baron-Cohen, 2000; Bora & Berk, 2016; Hamilton et al., 2009; Song et al., 2019; Yirmiya et al., 1998). One possibility is that codeficits of spatial, cognitive, and affective PT in populations such as those with autism spectrum disorder and major depressive disorder arise from diverse rather than common cognitive impairments. Future work may benefit from focusing efforts on understanding what varied general cognitive processes may contribute to each type of PT ability to devise interventions for the improvement of spatial, cognitive, or affective PT abilities in clinical populations.

### Is There Any Support for Shared Cognitive Processes Among PT Domains?

A few findings from the present study may be interpreted as providing partial, though clearly very limited, support for shared cognitive processes among some PT tasks in different domains. In particular, a small significant correlation was found among one spatial PT task (i.e., Spatial Orientation Task) and one affective PT task (i.e., Empathic Accuracy Task), even after removing the effect of general reasoning ability. This finding is somewhat unexpected given that performance on emotion recognition tasks has previously been found to be uncorrelated with spatial PT performance (Chiu & Yeh, 2018; Martin et al., 2019). This apparent lack of accord with earlier studies is not entirely problematic, however, since the tasks used in the present study were not pure emotion recognition tasks but rather are thought to additionally require mentalizing and differentiating between emotional mental states (Mackes et al., 2018; Shamay-Tsoory et al., 2007). Such qualities might have driven the observed relationship with the Spatial Orientation Task. However, neither affective PT measure was correlated with the alternative spatial PT task (i.e., Perspective-Taking Task for Adults), despite its high convergent validity with the Spatial Orientation Task. The overall results provide little support for the idea that spatial PT and affective PT rely on shared processing resources.

One cognitive process that proponents of unitary accounts have argued could be shared among these tasks is attentional control (Sun & Wang, 2014). However, the present study found that the Attentional Network Test orienting and executing indices were not associated with PT task performance. Therefore, individual differences in attentional control do not appear to account for the association of the Spatial Orientation Task with the Empathic Accuracy Task. An alternative explanation for why smaller angular errors in PT on the Spatial Orientation Task showed a modest negative correlation with accuracy in emotional intensity ratings of others on the Empathic Accuracy Task may be that both required magnitude judgments. It has previously been suggested that a generalized magnitude system serves as a basis for the development of accurate spatial judgments (Newcombe et al., 2015). Some researchers have also proposed that adults use a generalized magnitude system to organize representations of increasing emotional intensity

(Holmes & Lourenco, 2011; though, see Pitt & Casasanto, 2018, for critiques). Thus, future behavioral work may investigate if accuracy of magnitude judgments more generally can explain the association of performance on these tasks.

A small positive correlation was also observed between accuracy on one cognitive PT task (i.e., Director Task) and self-reported preference/experience for spatial-relational imagery, as measured by the Object-Spatial Imagery Questionnaire spatial subscale. However, the association did not hold when considering RT on the Director Task. Further, no significant associations were found between the Director Task and objective behavioral measures of spatial PT, and the Object-Spatial Imagery Questionnaire spatial subscale was not correlated with the other cognitive PT task (i.e., Strange Stories). We suspect that the small correlation between the Director Task and the Object-Spatial Imagery Questionnaire spatial subscale may be best attributed to the spatial features of Director Task stimuli, as compared to Strange Stories that had no visuospatial task demands.

### The Need for Convergent Validity Within Cognitive and Affective PT Domains

The measures of spatial PT used in this study have previously shown strong psychometric properties and convergent validity (e.g., Blajenkova et al., 2006; Brucato et al., 2023), and the coherence of measurement in this domain is corroborated in the present study. However, intentional investigations into convergent validity among measures within the cognitive PT and affective PT domains, such as those used in the present study, are sparse. In this data set, convergent validity was minimal among measures of cognitive PT and poor among affective PT measures. Specifically, the two behavioral measures of cognitive PT, the Director Task and the Strange Stories Task, were weakly correlated after controlling for general reasoning abilities (consistent with Navarro, 2021), but neither cognitive PT task exhibited a significant relationship with the Mentalization Scale other-related mentalization subscale self-report. Future research that carefully assesses the psychometric properties of existing measures of individual differences in cognitive and affective PT in neurotypical adults might help us to better understand what psychological constructs these tasks measure and what general cognitive processes may additionally support performance.

### Conclusion

Despite weaknesses in convergent validity identified in the present study, the present findings do not support the “common” or “cold versus hot” accounts of PT. Further establishment of convergent validity within cognitive and affective PT domains needs to be addressed before conclusions can be drawn about the tenability of “spatial versus social” and “distinct” accounts of PT. Future work should establish construct validity for tasks of cognitive and affective PT before returning to the question of their association with spatial PT and one another.

### Constraints on Generality

This study investigated the association among spatial, cognitive, and affective PT. Many studies have investigated codeficits among cognitive, affective, and spatial PT in clinical populations diagnosed with autism spectrum disorder (Baron-Cohen, 2000; Hamilton et al., 2009;

Muth et al., 2014; Pearson et al., 2013; Song et al., 2019; Yirmiya et al., 1998), schizophrenia (Achim et al., 2011; Derntl et al., 2009; Landgraf et al., 2010; Langdon et al., 2001; Sprong et al., 2007; Wilkins et al., 2013), and major depressive disorder (Bora & Berk, 2016; T. M. Erle et al., 2019). Several studies have also investigated this question in child and adolescent populations (Aslan & Köksal Akyol, 2020; Bensalah et al., 2016; Cassetta et al., 2018; Ford, 1979; Tian et al., 2021; Viana et al., 2016). In the present study, we found minimal evidence for the relation of different PT abilities in a sample of neurotypical adults. Follow-up investigations to the present analysis for construct validation or otherwise may benefit from using a similar sample, as the current results may not generalize to clinical populations or to young children who may not have intact Level 2 PT abilities.

## References

- Achim, A. M., Ouellet, R., Roy, M. A., & Jackson, P. L. (2011). Assessment of empathy in first-episode psychosis and meta-analytic comparison with previous studies in schizophrenia. *Psychiatry Research*, *190*(1), 3–8. <https://doi.org/10.1016/j.psychres.2010.10.030>
- Aichhorn, M., Perner, J., Kronbichler, M., Staffen, W., & Ladurner, G. (2006). Do visual perspective tasks need theory of mind? *NeuroImage*, *30*(3), 1059–1068. <https://doi.org/10.1016/j.neuroimage.2005.10.026>
- Allen, G. L., Kirasic, K. C., Dobson, S. H., Long, R. G., & Beck, S. (1996). Predicting environmental learning from spatial abilities: An indirect route. *Intelligence*, *22*(3), 327–355. [https://doi.org/10.1016/S0160-2896\(96\)90026-4](https://doi.org/10.1016/S0160-2896(96)90026-4)
- Amorim, M. A. (2003). “What is my avatar seeing?”: The coordination of “out-of-body” and “embodied” perspectives for scene recognition across views. *Visual Cognition*, *10*(2), 157–199. <https://doi.org/10.1080/17473756678>
- Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, *73*(1), 3–25. <https://doi.org/10.1037/amp0000191>
- Arora, A., Schurz, M., & Perner, J. (2017). Systematic comparison of brain imaging meta-analyses of ToM with vPT. *BioMed Research International*, *2017*(1), Article 6875850. <https://doi.org/10.1155/2017/6875850>
- Aslan, D., & Köksal Akyol, A. (2020). Impact of an empathy training program on children’s perspective-taking abilities. *Psychological Reports*, *123*(6), 2394–2409. <https://doi.org/10.1177/0033294119868785>
- Austin, G., Groppe, K., & Elsner, B. (2014). The reciprocal relationship between executive function and theory of mind in middle childhood: A 1-year longitudinal perspective. *Frontiers in Psychology*, *5*, Article 655. <https://doi.org/10.3389/fpsyg.2014.00655>
- Baksh, R. A., Abrahams, S., Auyeung, B., & MacPherson, S. E. (2018). The Edinburgh Social Cognition Test (ESCoT): Examining the effects of age on a new measure of theory of mind and social norm understanding. *PLOS ONE*, *13*(4), Article e0195818. <https://doi.org/10.1371/journal.pone.0195818>
- Baron-Cohen, S. (2000). Theory of mind and autism: A review. *International Review of Research in Mental Retardation*, *23*, 169–184. [https://doi.org/10.1016/S0074-7750\(00\)80010-5](https://doi.org/10.1016/S0074-7750(00)80010-5)
- Baron-Cohen, S. (2009). Autism: The empathizing–systemizing (E–S) theory. *Annals of the New York Academy of Sciences*, *1156*(1), 68–80. <https://doi.org/10.1111/j.1749-6632.2009.04467.x>
- Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, *34*(2), 163–175. <https://doi.org/10.1023/B:JADD.0000022607.19833.00>
- Bensalah, L., Caillies, S., & Anduze, M. (2016). Links among cognitive empathy, theory of mind, and affective perspective taking by young children. *The Journal of Genetic Psychology: Research and Theory on Human Development*, *177*(1), 17–31. <https://doi.org/10.1080/00221325.2015.1106438>
- Blajenkova, O., Kozhevnikov, M., & Motes, M. A. (2006). Object-spatial imagery: A new self-report imagery questionnaire. *Applied Cognitive Psychology*, *20*(2), 239–263. <https://doi.org/10.1002/acp.1182>
- Bora, E., & Berk, M. (2016). Theory of mind in major depressive disorder: A meta-analysis. *Journal of Affective Disorders*, *191*, 49–55. <https://doi.org/10.1016/j.jad.2015.11.023>
- Bora, E., & Zorlu, N. (2017). Social cognition in alcohol use disorder: A meta-analysis. *Addiction*, *112*(1), 40–48. <https://doi.org/10.1111/add.13486>
- Brucato, M. (2022). *Investigating the neural and behavioral association of spatial, cognitive, and affective perspective taking* [Doctoral dissertation, Temple University]. ProQuest. <https://www.proquest.com/dissertations-theses/investigating-neural-behavioral-association/docview/2766088385/se-2?accountid=28402>
- Brucato, M., Frick, A., Pichelmann, S., Nazareth, A., & Newcombe, N. S. (2023). Measuring spatial perspective taking: Analysis of four measures using item response theory. *Topics in Cognitive Science*, *15*(1), 46–74. <https://doi.org/10.1111/tops.12597>
- Brucato, M., Newcombe, N., & Chein, J. (2021, August 10). *Cognitive, affective, and spatial perspective taking*. <https://doi.org/10.17605/OSF.IO/3FXWB>
- Brucato, M., Newcombe, N., & Chein, J. (2025, March 23). *Behavioral association or dissociation of spatial, cognitive, and affective PT*. [https://osf.io/3fxwb/?view\\_only=56ad8b5ca2fe4f81ba11bddac6fbc5aa](https://osf.io/3fxwb/?view_only=56ad8b5ca2fe4f81ba11bddac6fbc5aa)
- Brunyé, T. T., Ditman, T., Giles, G. E., Mahoney, C. R., Kessler, K., & Taylor, H. A. (2012). Gender and autistic personality traits predict perspective-taking ability in typical adults. *Personality and Individual Differences*, *52*(1), 84–88. <https://doi.org/10.1016/j.paid.2011.09.004>
- Carruthers, P. (2013). Mindreading in infancy. *Mind & Language*, *28*(2), 141–172. <https://doi.org/10.1111/mila.12014>
- Cassetta, B. D., Pexman, P. M., & Goghari, V. M. (2018). Cognitive and affective theory of mind and relations with executive functioning in middle childhood. *Merrill-Palmer Quarterly*, *64*(4), 514–538. <https://doi.org/10.13110/merrpalmquar1982.64.4.0514>
- Chierchia, G., Fuhrmann, D., Knoll, L. J., Pi-Sunyer, B. P., Sakhardande, A. L., & Blakemore, S.-J. (2019). The matrix reasoning item bank (MaRS-IB): Novel, open-access abstract reasoning items for adolescents and adults. *Royal Society Open Science*, *6*(10), Article 190232. <https://doi.org/10.1098/rsos.190232>
- Chiu, C.-D., & Yeh, Y.-Y. (2018). In your shoes or mine? Shifting from other to self perspective is vital for emotional empathy. *Emotion*, *18*(1), 39–45. <https://doi.org/10.1037/emo0000346>
- Clements-Stephens, A. M., Vasiljevic, K., Murray, A. J., & Shelton, A. L. (2013). The role of potential agents in making spatial perspective taking social. *Frontiers in Human Neuroscience*, *7*, Article 497. <https://doi.org/10.3389/fnhum.2013.00497>
- Cotter, J., Granger, K., Backx, R., Hobbs, M., Looi, C. Y., & Barnett, J. H. (2018). Social cognitive dysfunction as a clinical marker: A systematic review of meta-analyses across 30 clinical conditions. *Neuroscience and Biobehavioral Reviews*, *84*, 92–99. <https://doi.org/10.1016/j.neubiorev.2017.11.014>
- Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine*, *11*(1), Article 126. <https://doi.org/10.1186/1741-7015-11-126>
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*(1), 113–126. <https://doi.org/10.1037/0022-3514.44.1.113>
- De Lillo, M., & Ferguson, H. J. (2023). Perspective-taking and social inferences in adolescents, young adults, and older adults. *Journal of Experimental Psychology: General*, *152*(5), 1420–1438. <https://doi.org/10.1037/xge0001337>

- Derntl, B., Finkelmeyer, A., Toygar, T. K., Hülsmann, A., Schneider, F., Falkenberg, D. I., & Habel, U. (2009). Generalized deficit in all core components of empathy in schizophrenia. *Schizophrenia Research*, 108(1–3), 197–206. <https://doi.org/10.1016/j.schres.2008.11.009>
- Dimitrijević, A., Hanak, N., Altaras Dimitrijević, A., & Jolić Marjanović, Z. (2018). The Mentalization Scale (MentS): A self-report measure for the assessment of mentalizing capacity. *Journal of Personality Assessment*, 100(3), 268–280. <https://doi.org/10.1080/00223891.2017.1310730>
- Dumontheil, I., Küster, O., Apperly, I. A., & Blakemore, S.-J. (2010). Taking perspective into account in a communicative task. *NeuroImage*, 52(4), 1574–1583. <https://doi.org/10.1016/j.neuroimage.2010.05.056>
- Erle, T., & Topolinski, S. (2015). Spatial and empathic perspective-taking correlate on a dispositional level. *Social Cognition*, 33(3), 187–210. <https://doi.org/10.1521/soco.2015.33.3.187>
- Erle, T. M., Barth, N., & Topolinski, S. (2019). Egocentrism in sub-clinical depression. *Cognition and Emotion*, 33(6), 1239–1248. <https://doi.org/10.1080/02699931.2018.1552120>
- Eysenck, S. B., & Eysenck, H. J. (1978). Impulsiveness and venturesomeness: Their position in a dimensional system of personality description. *Psychological Reports*, 43(3\_suppl), 1247–1255. <https://doi.org/10.2466/pr0.1978.43.3f.1247>
- Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, 14(3), 340–347. <https://doi.org/10.1162/089892902317361886>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G\*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Fields, A. W., & Shelton, A. L. (2006). Individual skill differences and large-scale environmental learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(3), 506–515. <https://doi.org/10.1037/0278-7393.32.3.506>
- Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the Level 1–Level 2 distinction. *Developmental Psychology*, 17(1), 99–103. <https://doi.org/10.1037/0012-1649.17.1.99>
- Ford, M. E. (1979). The construct validity of egocentrism. *Psychological Bulletin*, 86(6), 1169–1188. <https://doi.org/10.1037/0033-2909.86.6.1169>
- Friedman, A., Kohler, B., Gunalp, P., Boone, A. P., & Hegarty, M. (2020). A computerized spatial orientation test. *Behavior Research Methods*, 52(2), 799–812. <https://doi.org/10.3758/s13428-019-01277-3>
- Galinsky, A. D., Ku, G., & Wang, C. S. (2005). Perspective-taking and self-other overlap: Fostering social bonds and facilitating social coordination. *Group Processes & Intergroup Relations*, 8(2), 109–124. <https://doi.org/10.1177/1368430205051060>
- Gallant, C. M. M., Lavis, L., & Mahy, C. E. V. (2020). Developing an understanding of others' emotional states: Relations among affective theory of mind and empathy measures in early childhood. *British Journal of Developmental Psychology*, 38(2), 151–166. <https://doi.org/10.1111/bjdp.12322>
- Gomez, R., Brown, T., Watson, S., & Stavropoulos, V. (2022). Confirmatory factor analysis and exploratory structural equation modeling of the factor structure of the Questionnaire of Cognitive and Affective Empathy (QCAE). *PLOS ONE*, 17(2), Article e0261914. <https://doi.org/10.1371/journal.pone.0261914>
- Goodhew, S. C., & Edwards, M. (2021). Attentional control both helps and harms empathy. *Cognition*, 206, Article 104505. <https://doi.org/10.1016/j.cognition.2020.104505>
- Gronholm, P. C., Flynn, M., Edmonds, C. J., & Gardner, M. R. (2012). Empathic and non-empathic routes to visuospatial perspective-taking. *Consciousness and Cognition*, 21(1), 494–500. <https://doi.org/10.1016/j.concog.2011.12.004>
- Hamilton, A. F. de C., Brindley, R., & Frith, U. (2009). Visual perspective taking impairment in children with autistic spectrum disorder. *Cognition*, 113(1), 37–44. <https://doi.org/10.1016/j.cognition.2009.07.007>
- Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology*, 33(3), 307–316. <https://doi.org/10.1037/h0027580>
- Holmes, K. J., & Lourenco, S. F. (2011). Common spatial organization of number and emotional expression: A mental magnitude line. *Brain and Cognition*, 77(2), 315–323. <https://doi.org/10.1016/j.bandc.2011.07.002>
- Hughes, C., & Leekam, S. (2004). What are the links between theory of mind and social relations? Review, reflections and new directions for studies of typical and atypical development. *Social Development*, 13(4), 590–619. <https://doi.org/10.1111/j.1467-9507.2004.00285.x>
- Ickes, W. (1993). Empathic accuracy. *Journal of Personality*, 61(4), 587–610. <https://doi.org/10.1111/j.1467-6494.1993.tb00783.x>
- Isernia, S., Cacciatore, D. M., Rossetto, F., Ricci, C., & Baglio, F. (2024). Reliability and minimal detectable change of the Yoni task for the theory of mind assessment. *Frontiers in Psychology*, 15, Article 1412560. <https://doi.org/10.3389/fpsyg.2024.1412560>
- Isernia, S., Rossetto, F., Blasi, V., Massaro, D., Castelli, I., Ricci, C., Shamay-Tsoory, S., Marchetti, A., & Baglio, F. (2023). Measuring cognitive and affective theory of mind with the Italian Yoni task: Normative data and short versions. *Current Psychology*, 42(27), 23519–23530. <https://doi.org/10.1007/s12144-022-03457-5>
- Isernia, S., Rossetto, F., Shamay-Tsoory, S., Marchetti, A., & Baglio, F. (2023). Standardization and normative data of the 48-item Yoni short version for the assessment of theory of mind in typical and atypical conditions. *Frontiers in Aging Neuroscience*, 14, Article 1048599. <https://doi.org/10.3389/fnagi.2022.1048599>
- Israelashvili, J., Sauter, D., & Fischer, A. (2019). How well can we assess our ability to understand others' feelings? Beliefs about taking others' perspectives and actual understanding of others' emotions. *Frontiers in Psychology*, 10, Article 2475. <https://doi.org/10.3389/fpsyg.2019.02475>
- Job, X., Kirsch, L., Inard, S., Arnold, G., & Auvray, M. (2021). Spatial perspective taking is related to social intelligence and attachment style. *Personality and Individual Differences*, 168, Article 109726. <https://doi.org/10.1016/j.paid.2019.109726>
- Kosakowski, H. L., & Saxe, R. (2018). “Affective theory of mind” and the function of the ventral medial prefrontal cortex. *Cognitive and Behavioral Neurology*, 31(1), 36–50. <https://doi.org/10.1097/WNN.0000000000000150>
- Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & Cognition*, 29(5), 745–756. <https://doi.org/10.3758/BF03200477>
- Kozhevnikov, M., Motes, M. A., Rasch, B., & Blajenkova, O. (2006). Perspective-taking vs. mental rotation transformations and how they predict spatial navigation performance. *Applied Cognitive Psychology*, 20(3), 397–417. <https://doi.org/10.1002/acp.1192>
- Landgraf, S., Krebs, M. O., Olić, J. P., Committer, G., van der Meer, E., Berthoz, A., & Amado, I. (2010). Real world referencing and schizophrenia: Are we experiencing the same reality? *Neuropsychologia*, 48(10), 2922–2930. <https://doi.org/10.1016/j.neuropsychologia.2010.05.034>
- Langdon, R., Coltheart, M., Ward, P. B., & Catts, S. V. (2001). Visual and cognitive perspective-taking impairments in schizophrenia: A failure of allocentric simulation? *Cognitive Neuropsychiatry*, 6(4), 241–269. <https://doi.org/10.1080/13546800143000005>
- Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in “theory of mind”. *Trends in Cognitive Sciences*, 8(12), 528–533. <https://doi.org/10.1016/j.tics.2004.10.001>
- Long, E. C., & Andrews, D. W. (1990). Perspective taking as a predictor of marital adjustment. *Journal of Personality and Social Psychology*, 59(1), 126–131. <https://doi.org/10.1037/0022-3514.59.1.126>
- Long, M. R., Horton, W. S., Rohde, H., & Sorace, A. (2018). Individual differences in switching and inhibition predict perspective-taking across the lifespan. *Cognition*, 170, 25–30. <https://doi.org/10.1016/j.cognition.2017.09.004>

- Mackes, N. K., Golm, D., O'Daly, O. G., Sarkar, S., Sonuga-Barke, E. J. S., Fairchild, G., & Mehta, M. A. (2018). Tracking emotions in the brain—Revisiting the empathic accuracy task. *NeuroImage*, 178, 677–686. <https://doi.org/10.1016/j.neuroimage.2018.05.080>
- Martin, A. K., Perceval, G., Davies, I., Su, P., Huang, J., & Meinzer, M. (2019). Visual perspective taking in young and older adults. *Journal of Experimental Psychology: General*, 148(11), 2006–2026. <https://doi.org/10.1037/xge0000584>
- Mattan, B. D., Rotshtein, P., & Quinn, K. A. (2016). Empathy and visual perspective-taking performance. *Cognitive Neuroscience*, 7(1–4), 170–181. <https://doi.org/10.1080/17588928.2015.1085372>
- Mohr, C., Rowe, A. C., & Blanke, O. (2010). The influence of sex and empathy on putting oneself in the shoes of others. *British Journal of Psychology*, 101(2), 277–291. <https://doi.org/10.1348/000712609X457450>
- Murphy, B. A., & Lilienfeld, S. O. (2019). Are self-report cognitive empathy ratings valid proxies for cognitive empathy ability? Negligible meta-analytic relations with behavioral task performance. *Psychological Assessment*, 31(8), 1062–1072. <https://doi.org/10.1037/pas0000732>
- Murray, K., Johnston, K., Cunnane, H., Kerr, C., Spain, D., Gillan, N., Hammond, N., Murphy, D., & Happé, F. (2017). A new test of advanced theory of mind: The “Strange Stories Film Task” captures social processing differences in adults with autism spectrum disorders. *Autism Research*, 10(6), 1120–1132. <https://doi.org/10.1002/aur.1744>
- Muth, A., Hönekopp, J., & Falter, C. M. (2014). Visuo-spatial performance in autism: A meta-analysis. *Journal of Autism and Developmental Disorders*, 44(12), 3245–3263. <https://doi.org/10.1007/s10803-014-2188-5>
- Navarro, E. (2021). *What is theory of mind? A psychometric study of theory of mind and intelligence*. PsyArXiv. <https://doi.org/10.31234/osf.io/dhpk4>
- Navarro, E. (2022). What is theory of mind? A psychometric study of theory of mind and intelligence. *Cognitive Psychology*, 136, Article 101495. <https://doi.org/10.1016/j.cogpsych.2022.101495>
- Nazareth, A., Weisberg, S. M., Margulis, K., & Newcombe, N. S. (2018). Charting the development of cognitive mapping. *Journal of Experimental Child Psychology*, 170, 86–106. <https://doi.org/10.1016/j.jecp.2018.01.009>
- Nestor, B. A., Sutherland, S., & Garber, J. (2022). Theory of mind performance in depression: A meta-analysis. *Journal of Affective Disorders*, 303, 233–244. <https://doi.org/10.1016/j.jad.2022.02.028>
- Newcombe, N. S., Levine, S. C., & Mix, K. S. (2015). Thinking about quantity: The intertwined development of spatial and numerical cognition. *WIREs Cognitive Science*, 6(6), 491–505. <https://doi.org/10.1002/wcs.1369>
- Nielsen, M. K., Slade, L., Levy, J. P., & Holmes, A. (2015). Inclined to see it your way: Do allocentric intrusion effects in visual perspective taking reflect an intrinsically social process? *Quarterly Journal of Experimental Psychology*, 68(10), 1931–1951. <https://doi.org/10.1080/17470218.2015.1023206>
- Oakley, B. F. M., Brewer, R., Bird, G., & Catmur, C. (2016). Theory of mind is not theory of emotion: A cautionary note on the Reading the Mind in the Eyes Test. *Journal of Abnormal Psychology*, 125(6), 818–823. <https://doi.org/10.1037/abn0000182>
- Oldakowski, R. K. (2001). Activities to develop a spatial perspective among students in introductory geography courses. *The Journal of Geography*, 100(6), 243–250. <https://doi.org/10.1080/00221340108978451>
- Parkinson, C., & Wheatley, T. (2013). Old cortex, new contexts: Re-purposing spatial perception for social cognition. *Frontiers in Human Neuroscience*, 7, Article 645. <https://doi.org/10.3389/fnhum.2013.00645>
- Pearson, A., Ropar, D., & de C Hamilton, A. F. (2013). A review of visual perspective taking in autism spectrum disorder. *Frontiers in Human Neuroscience*, 7, Article 652. <https://doi.org/10.3389/fnhum.2013.00652>
- Pitt, B., & Casasanto, D. (2018). Spatializing emotion: No evidence for a domain-general magnitude system. *Cognitive Science*, 42(7), 2150–2180. <https://doi.org/10.1111/cogs.12568>
- Plummer, J. D., Bower, C. A., & Liben, L. S. (2016). The role of perspective taking in how children connect reference frames when explaining astronomical phenomena. *International Journal of Science Education*, 38(3), 345–365. <https://doi.org/10.1080/09500693.2016.1140921>
- Putko, A., & Zlotogórska-Suwińska, A. (2019). Visual perspective-taking: Are theory of mind and executive function involved? *Polish Psychological Bulletin*, 50(2), 184–192. <https://doi.org/10.24425/ppb.2019.129451>
- Quednow, B. B. (2017). Social cognition and interaction in stimulant use disorders. *Current Opinion in Behavioral Sciences*, 13, 55–62. <https://doi.org/10.1016/j.cobeha.2016.10.001>
- Quesque, F., & Rossetti, Y. (2020). What do theory-of-mind tasks actually measure? Theory and practice. *Perspectives on Psychological Science*, 15(2), 384–396. <https://doi.org/10.1177/1745691619896607>
- Reniers, R. L. E. P., Corcoran, R., Drake, R., Shryane, N. M., & Völlm, B. A. (2011). The QCAE: A Questionnaire of Cognitive and Affective Empathy. *Journal of Personality Assessment*, 93(1), 84–95. <https://doi.org/10.1080/00223891.2010.528484>
- Richter, F., Steinmair, D., & Löffler-Stastka, H. (2021). Construct validity of the Mentalization Scale (MentS) within a mixed psychiatric sample. *Frontiers in Psychology*, 12, Article 608214. <https://doi.org/10.3389/fpsyg.2021.608214>
- Saxe, R., & Kanwisher, N. (2005). People thinking about thinking people: The role of the temporo-parietal junction in “theory of mind”. In J. T. Cacioppo & G. G. Berntson (Eds.), *Social neuroscience: Key readings* (pp. 171–182). Psychology Press.
- Saxe, R., & Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia*, 43(10), 1391–1399. <https://doi.org/10.1016/j.neuropsychologia.2005.02.013>
- Schurz, M., Aichhorn, M., Martin, A., & Perner, J. (2013). Common brain areas engaged in false belief reasoning and visual perspective taking: A meta-analysis of functional brain imaging studies. *Frontiers in Human Neuroscience*, 7, Article 712. <https://doi.org/10.3389/fnhum.2013.00712>
- Shamay-Tsoory, S. G., Shur, S., Barcai-Goodman, L., Medlovich, S., Harari, H., & Levkovitz, Y. (2007). Dissociation of cognitive from affective components of theory of mind in schizophrenia. *Psychiatry Research*, 149(1–3), 11–23. <https://doi.org/10.1016/j.psychres.2005.10.018>
- Shelton, A. L., Clements-Stephens, A. M., Lam, W. Y., Pak, D. M., & Murray, A. J. (2012). Should social savvy equal good spatial skills? The interaction of social skills with spatial perspective taking. *Journal of Experimental Psychology: General*, 141(2), 199–205. <https://doi.org/10.1037/a0024617>
- Song, Y., Nie, T., Shi, W., Zhao, X., & Yang, Y. (2019). Empathy impairment in individuals with autism spectrum conditions from a multidimensional perspective: A meta-analysis. *Frontiers in Psychology*, 10, Article 1902. <https://doi.org/10.3389/fpsyg.2019.01902>
- Sprong, M., Schothorst, P., Vos, E., Hox, J., & van Engeland, H. (2007). Theory of mind in schizophrenia: Meta-analysis. *The British Journal of Psychiatry*, 191(1), 5–13. <https://doi.org/10.1192/bjp.bp.107.035899>
- Sulpizio, V., Committeri, G., Metta, E., Lambrey, S., Berthoz, A., & Galati, G. (2015). Visuospatial transformations and personality: Evidence of a relationship between visuospatial perspective taking and self-reported emotional empathy. *Experimental Brain Research*, 233(7), 2091–2102. <https://doi.org/10.1007/s00221-015-4280-2>
- Sun, Y., & Wang, H. (2014). Insight into others' minds: Spatio-temporal representations by intrinsic frame of reference. *Frontiers in Human Neuroscience*, 8, Article 58. <https://doi.org/10.3389/fnhum.2014.00058>
- Sunahara, C. S., Rosenfield, D., Alvi, T., Wallmark, Z., Lee, J., Fulford, D., & Tabak, B. A. (2022). Revisiting the association between self-reported empathy and behavioral assessments of social cognition. *Journal of Experimental Psychology: General*, 151(12), Article 3304. <https://doi.org/10.1037/xge0001226>
- Thakkar, K. N., Brugger, P., & Park, S. (2009). Exploring empathic space: Correlates of perspective transformation ability and biases in spatial attention. *PLOS ONE*, 4(6), Article e5864. <https://doi.org/10.1371/journal.pone.0005864>

- Thakkar, K. N., & Park, S. (2010). Empathy, schizotypy, and visuospatial transformations. *Cognitive Neuropsychiatry*, 15(5), 477–500. <https://doi.org/10.1080/13546801003711350>
- Tian, M., Luo, T., Ding, J., Wang, X., & Cheung, H. (2021). Spatial ability and theory of mind: A mediating role of visual perspective taking. *Child Development*, 92(4), 1590–1604. <https://doi.org/10.1111/cdev.13546>
- Tomei, A., Besson, J., & Grivel, J. (2017). Linking empathy to visuospatial perspective-taking in gambling addiction. *Psychiatry Research*, 250, 177–184. <https://doi.org/10.1016/j.psychres.2016.12.061>
- Viana, K. M. P., Zambrana, I. M., Karevold, E. B., & Pons, F. (2016). Beyond conceptual knowledge: The impact of children's theory-of-mind on dyadic spatial tasks. *Frontiers in Psychology*, 7, Article 1635. <https://doi.org/10.3389/fpsyg.2016.01635>
- Wade, M., Prime, H., Jenkins, J. M., Yeates, K. O., Williams, T., & Lee, K. (2018). On the relation between theory of mind and executive functioning: A developmental cognitive neuroscience perspective. *Psychonomic Bulletin & Review*, 25(6), 2119–2140. <https://doi.org/10.3758/s13423-018-1459-0>
- Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, 101(4), 817–835. <https://doi.org/10.1037/a0016127>
- Warnell, K. R., & Redcay, E. (2019). Minimal coherence among varied theory of mind measures in childhood and adulthood. *Cognition*, 191, Article 103997. <https://doi.org/10.1016/j.cognition.2019.06.009>
- White, S., Hill, E., Happé, F., & Frith, U. (2009). Revisiting the strange stories: Revealing mentalizing impairments in autism. *Child Development*, 80(4), 1097–1117. <https://doi.org/10.1111/j.1467-8624.2009.01319.x>
- Wilkins, L. K., Girard, T. A., King, J., King, M. J., Herdman, K. A., Christensen, B. K., & King, J. (2013). Spatial-memory deficit in schizophrenia spectrum disorders under viewpoint-independent demands in the virtual courtyard task. *Journal of Clinical and Experimental Neuropsychology*, 35(10), 1082–1093. <https://doi.org/10.1080/13803395.2013.857389>
- Yirmiya, N., Erel, O., Shaked, M., & Solomonica-Levi, D. (1998). Meta-analyses comparing theory of mind abilities of individuals with autism, individuals with mental retardation, and normally developing individuals. *Psychological Bulletin*, 124, 283–307. <https://doi.org/10.1037/0033-2909.124.3.283>
- Zobel, I., Werden, D., Linster, H., Dykierck, P., Drieling, T., Berger, M., & Schramm, E. (2010). Theory of mind deficits in chronically depressed patients. *Depression and Anxiety*, 27(9), 821–828. <https://doi.org/10.1002/da.20713>

(Appendix follows)

## Appendix

## Supplementary Materials, Including Summary of Prior Behavioral Studies With Neurotypical Adults (Tables A1–A2) and Supplementary Analyses From the Present Study (Tables A3–A5 and Figures A1–A2)

Table A1

Summary of Behavioral Studies With Neurotypical Adults on the Association of Spatial and Affective PT

| References                                  | N   | Spatial PT measure                                    | Spatial PT measure: agent | Affective PT measure | Significant result                                                                                                                                                                | Control task/condition                                               |
|---------------------------------------------|-----|-------------------------------------------------------|---------------------------|----------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------|
| Thakkar et al. (2009)                       | 40  | OBT; accuracy and RT                                  | Yes                       | IRI                  | Female ([RT back – RT front] × IRI, EC): $r = .70, p = .0008$                                                                                                                     | Line Bisection × IRI, $r = .55, p = .0002$                           |
| Thakkar and Park (2010)                     | 32  | OBT; accuracy and RT                                  | Yes                       | IRI                  | Female (Accuracy Front × IRI, PT): $r = .63, p = .008$                                                                                                                            | Letter Rotation × IRI, $r = .57, p = .03$                            |
| Mohr et al. (2010)                          | 100 | OBT; accuracy and RT                                  | Yes                       | EQ (reverse-scored)  | Female (RT Front × EQ): $r = -.29, p = .04$                                                                                                                                       | NA                                                                   |
| Gronholm et al. (2012)                      | 85  | OBT; accuracy and RT                                  | Yes                       | EQ (reverse-scored)  | PT strategy users ( $n = 45$ ; RT × EQ, total): $r = .413, p = .005$ ; PT strategy users ( $n = 45$ ; RT × EQ, ER): $r = .339, p = .023$                                          | NA                                                                   |
| T. Erle and Topolinski (2015; Experiment 1) | 113 | TFT; accuracy                                         | No                        | IRI                  | All (Accuracy × IRI, PT): $r = .20, p = .033$                                                                                                                                     | NA                                                                   |
| T. Erle and Topolinski (2015; Experiment 2) | 120 | TFT; accuracy                                         | No                        | IRI                  | All (IRI, PT): $r = .20, p = .026$ ; all (Accuracy × IRI, PT, removing effect of arithmetic): $r = .20, p = .031$                                                                 | Arithmetic <i>ns</i> with IRI; Arithmetic × TFT, $r = .34, p < .001$ |
| T. Erle and Topolinski (2015; Experiment 3) | 169 | TFT; accuracy                                         | No                        | IRI                  | All (Accuracy × IRI, PT): $r = .22, p = .004$ ; PT strategy users ( $n = 86$ ; Accuracy × IRI, PT): $r = .29, p = .008$                                                           | NA                                                                   |
| T. Erle and Topolinski (2015; Experiment 4) | 163 | Table PT task (Level 2); RT                           | Yes                       | IRI                  | All (Accuracy × IRI, PT): $r = -.20, p = .013$                                                                                                                                    | NA                                                                   |
| Sulpizio et al. (2015)                      | 148 | Spatial memory PT task (Level 2); RT                  | Yes                       | IRI and BEES         | All (Nonavatar × BEES): $r = -.19, p < .05$ ; all (Nonavatar × IRI, FS): $r = -.17, p < .05$ ; all (Avatar 135° × BEES): $r = -.25, p < .01$                                      | Memory without spatial PT <i>ns</i> with IRI and BEES                |
| Tomei et al. (2017)                         | 76  | Table PT task (Level 2); no. of correct per minute    | Yes                       | IRI                  | All (Correct/Min × IRI, PT): $\beta = 0.20, p = .04$ ; all (correct/Min × IRI, PD): $\beta = -.023, p = .03$                                                                      | NA                                                                   |
| Chiu and Yeh (2018)                         | 158 | Spatial task-switching; switch cost                   | No                        | RMET and MET         | All (Switch Cost × MET, Emotional): $\beta = -.032, p < .01$                                                                                                                      | NA                                                                   |
| Martin et al. (2019)                        | 146 | Street PT task (Level 1 and Level 2); accuracy and RT | Yes                       | RMET and SEC         | Younger adults (allocentric CE accuracy, Level 1 × SEC): $r = -.209, BF_{10} = 3.120$ ; all (allocentric CE accuracy, Level 2 × SEC): $r = -.244, BF_{10} = 9.904$                | Spatial PT with traffic light <i>ns</i> with SEC                     |
| Nielsen et al. (2015; Experiment 1)         | 32  | Dot task (Level 1); RT                                | Yes                       | IRI                  | All (Avatar × IRI, PT): $r = .32, p = .035$ ; all (Avatar × IRI, EC): $r = .34, p = .029$                                                                                         | Spatial PT with arrow <i>ns</i> with IRI                             |
| Nielsen et al. (2015; Experiment 2)         | 28  | Dot task (Level 1); gaze duration                     | Yes                       | IRI                  | All (Avatar × IRI, PT): $r = .37, p = .027$ ; all (Avatar × IRI, EC): $r = .32, p = .051$                                                                                         | Dot task without avatar <i>ns</i> with IRI                           |
| Mattan et al. (2016)                        | 43  | Dot task (Level 1); accuracy and RT                   | Yes                       | IRI                  | All (Self-Avatar CE RT × IRI, total): $r = -.410, p = .008$ ; all (Self-Avatar CE RT × IRI, PT): $r = -.455, p = .003$ ; all (Self-Avatar CE RT × IRI, EC): $r = -.471, p = .002$ | NA                                                                   |

Note. PT = perspective taking; OBT = Own Body Transformation Task; RT = reaction time; IRI = Interpersonal Reactivity Index; IRI, PT = perspective-taking subscale; IRI, FS = fantasy subscale; IRI, PD = personal distress subscale; IRI, EC = empathic concern subscale; EQ = Empathy Quotient; EQ, ER = emotional reactivity subscale; TFT = Tube Figures Test; BEES = Balanced Emotional Empathy Scale; RMET = Reading the Mind in the Eyes Task; MET = Multifaceted Empathy Test; SEC = Social-Emotional Cognition Task; CE = congruency effect; *ns* = nonsignificant; NA = not applicable; BF = Bayes factors.

(Appendix continues)

**Table A2***Summary of Behavioral Studies With Neurotypical Adults on the Association of Spatial and Cognitive PT*

| References                                    | N   | Spatial PT measure                                   | Spatial PT measure: agent | Cognitive PT measure          | Significant result                                                                                                                                                                                                                                               | Control task/condition                                                                          |
|-----------------------------------------------|-----|------------------------------------------------------|---------------------------|-------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|
| Shelton et al. (2012; main experiment)        | 72  | Three buildings task (Level 2); accuracy             | Yes                       | AQ                            | All (Agent $\times$ AQ, Social/Com.): $r = -.584, p < .05$ ; male (Agent $\times$ AQ, Social/Com.): $r = -.660, p < .05$ ; female (Agent $\times$ AQ, Social/Com.): $r = -.501, p < .05$                                                                         | Object rotation, <i>ns</i> with AQ; spatial PT without agent, <i>ns</i> with AQ                 |
| Shelton et al. (2012; replication)            | 70  | Three mountains task (Level 2); accuracy             | Yes                       | AQ                            | All (Agent $\times$ AQ, Social/Com.): $r = -.501, p < .05$                                                                                                                                                                                                       | NA                                                                                              |
| Brunyé et al. (2012)                          | 140 | Table PT task (Level 1 and Level 2); RT              | Yes                       | AQ                            | All (Level 2 Angle Effect $\times$ AQ, total): $\beta_{std} = 0.37, p < .01$ ; female (Level 2 Angle Effect $\times$ AQ, total): $\beta_{std} = 0.26, p < .05$ ; male (Level 2 Angle Effect $\times$ AQ, total): $\beta_{std} = 0.52, p < .01$                   | NA                                                                                              |
| Clements-Stephens et al. (2013; Experiment 1) | 72  | Three buildings task (Level 2); accuracy             | Yes                       | AQ; Social and Com. subscales | Agent trial presented first (Nonagent $\times$ AQ, Social/Com.): $r = -.50, p = .001$ ; agent trial presented first (Agent $\times$ AQ, Social/Com.): $r = -.46, p = .004$ ; agent trial presented second (Agent $\times$ AQ, Social/Com.): $r = -.45, p = .008$ | When agent trial presented second, accuracy for nonagent trials <i>ns</i> with AQ, Social/Com.  |
| Clements-Stephens et al. (2013; Experiment 2) | 72  | Three buildings task (Level 2); accuracy             | Yes                       | AQ; Social and Com. subscales | All (Face Agent $\times$ AQ, Social/Com.): $r = -.40, p = .001$                                                                                                                                                                                                  | Spatial PT with no-face agent <i>ns</i> with AQ, Social/Com.                                    |
| Clements-Stephens et al. (2013; Experiment 3) | 48  | Three buildings task (Level 2); accuracy             | Yes                       | AQ; Social and Com. subscales | All ("Triangle Aliens" $\times$ AQ, Social/Com.): $r = -.36, p = .013$                                                                                                                                                                                           | NA                                                                                              |
| Job et al. (2021)                             | 44  | Spatial orientation task (Level 2); error            | No                        | TSIS                          | NA                                                                                                                                                                                                                                                               | Object rotation <i>ns</i> with TSIS                                                             |
| Putko and Złotogórska-Suwińska (2019)         | 67  | Table PT task (Level 1 and Level 2); accuracy and RT | Yes                       | Strange Stories Task          | All (Level 2 RT $\times$ Strange Stories Accuracy): $r = -.36, p < .01$ ; all (Level 1 RT $\times$ Strange Stories Accuracy): $r = -.29, p < .05$                                                                                                                | Trail Making Task and Wisconsin Card Sorting Test <i>ns</i> with spatial PT or Strange Stories. |

Note. PT = perspective taking; AQ = Autism Quotient; AQ, Social/Com. = averaged social subscale and communications subscale; RT = reaction time; TSIS = Tromsø Social Intelligence Scale; *ns* = nonsignificant; NA = not applicable.

**Table A3***Descriptive Statistics and Paired *t* Tests for Reaction Time by Angular Disparity on Perspective-Taking Task for Adults Trials*

| Angular disparity (°) | <i>n</i> | <i>M</i> | <i>SD</i> | Min. | Max.  | Comparison to 0° trial |                     |
|-----------------------|----------|----------|-----------|------|-------|------------------------|---------------------|
|                       |          |          |           |      |       | Mean difference        | Confidence interval |
| 0                     | 191      | 8.66     | 3.24      | 3.77 | 21.09 |                        |                     |
| 45                    | 188      | 10.80    | 4.14      | 4.29 | 34.95 | -2.20*                 | [-2.75, -1.65]      |
| 90                    | 191      | 10.99    | 3.93      | 5.62 | 30.09 | -2.24*                 | [-2.79, -1.70]      |
| 135                   | 189      | 11.74    | 5.34      | 4.44 | 41.76 | -3.11*                 | [-3.85, -2.37]      |
| 180                   | 183      | 11.82    | 6.19      | 3.98 | 42.06 | -3.29*                 | [-4.16, -2.41]      |

Note. The table shows the descriptive statistics and paired-samples *t* test results with 95% confidence intervals comparing the mean reaction time in seconds for the 0° angular disparity trials to the mean reaction time in seconds for each other trial type. Min. = minimum; Max. = maximum.

\* $p < .001$ .

(Appendix continues)

**Table A4***Descriptive Statistics for Task RTs*

| Construct         | Task RT (ms)    | <i>n</i> | <i>M</i>  | <i>SD</i> | Median   | Min.     | Max.      |
|-------------------|-----------------|----------|-----------|-----------|----------|----------|-----------|
| Spatial PT        | PTT-A           | 193      | 10,432.64 | 3,457.09  | 9,660.00 | 5,100.00 | 28,290.00 |
| Cognitive PT      | Strange Stories | 193      | 8,389.70  | 5,981.37  | 6,711.59 | 1,666.55 | 39,740.27 |
| Cognitive PT      | Director Task   | 173      | 2,947.86  | 593.91    | 2,868.00 | 137.00   | 5,040.00  |
| Affective PT      | Yoni            | 178      | 3,727.51  | 981.45    | 3,617.27 | 1,652.65 | 7,990.94  |
| General reasoning | MaRs            | 193      | 7,829.13  | 2,540.43  | 7,740.70 | 1,671.00 | 13,815.45 |

*Note.* PT = perspective taking; RT (ms) = reaction time in milliseconds; PTT-A = Perspective-Taking Task for Adults; MaRs = Matrix Reasoning; Min. = minimum; Max. = maximum.

**Table A5***Results of Pearson's Correlations Between All Measures Using Available Reaction Time Data*

| Measure                | Construct         | 1      | 2      | 3     | 4    | 5    | 6    | 7    | 8      | 9    | 10   | 11   | 12 |
|------------------------|-------------------|--------|--------|-------|------|------|------|------|--------|------|------|------|----|
| 1. SOT (error)         | Spatial PT        | —      |        |       |      |      |      |      |        |      |      |      |    |
| 2. PTT-A (RT)          | Spatial PT        | .28**  | —      |       |      |      |      |      |        |      |      |      |    |
| 3. OSIQ (self-report)  | Spatial PT        | -.27** | -.18   | —     |      |      |      |      |        |      |      |      |    |
| 4. Director (RT)       | Cognitive PT      | -.14   | .02    | .10   | —    |      |      |      |        |      |      |      |    |
| 5. Stories (RT)        | Cognitive PT      | -.01   | .03    | .08   | -.03 | —    |      |      |        |      |      |      |    |
| 6. MentS (self-report) | Cognitive PT      | .00    | .00    | .09   | -.05 | .03  | —    |      |        |      |      |      |    |
| 7. EAT (z-score)       | Affective PT      | -.27** | -.09   | .04*  | -.09 | -.22 | .00  | —    |        |      |      |      |    |
| 8. Yoni (RT)           | Affective PT      | .04    | .17    | .10** | .17  | .41  | -.03 | -.03 | —      |      |      |      |    |
| 9. QCAE (self-report)  | Affective PT      | .03    | .07    | .18   | .00  | .03  | .70* | -.01 | .03    | —    |      |      |    |
| 10. Orienting          | Attention         | .05    | .02    | -.01  | .07  | -.10 | .09  | .08  | .01    | .17  | —    |      |    |
| 11. Executing          | Attention         | -.03   | -.12   | .08   | -.01 | -.04 | .05  | .17  | .04    | -.04 | -.05 | —    |    |
| 12. MaRs               | General reasoning | -.11   | -.26** | .11   | -.03 | -.19 | .00  | .08  | -.29** | -.04 | -.03 | -.06 | —  |

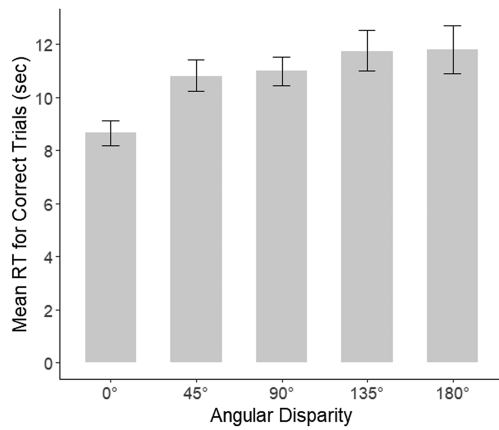
*Note.* False discovery rate corrections were applied to adjust for multiple comparisons. *n* = 178 for correlations with Yoni. *n* = 193 for all other correlations. MaRs = Matrix Reasoning; SOT = Spatial Orientation Test; PTT-A = Perspective-Taking Task for Adults; EAT = Empathic Accuracy Task; OSIQ = Object-Spatial Imagery Questionnaire; QCAE = Questionnaire of Cognitive and Affective Empathy; MentS = Mentalization Scale; PT = perspective taking; RT = reaction time.

\**p* < .05. \*\**p* < .001.

(Appendix continues)

**Figure A1**

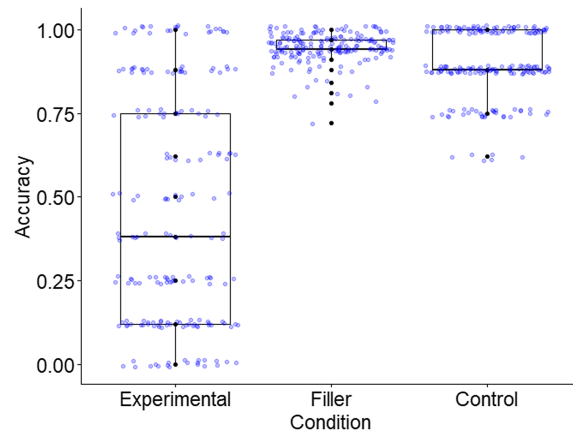
*Mean RT by Angular Disparity on Perspective-Taking Task for Adults Trials*



*Note.* Error bars represent  $\pm 1$  standard error of the mean. RT = reaction time.

**Figure A2**

*Mean Accuracy by Trial Type on Director Task*



*Note.* Blue dots are the jittered data points of individual participants' mean accuracy. Black dots are the nonjittered data points. Experimental trials required the participant to consider the director's perspective in order to achieve a correct answer. Control trials and filler trials did not require perspective taking for a correct response. See the online article for the color version of this figure.

Received September 21, 2024

Revision received September 11, 2025

Accepted October 29, 2025 ■