Hodgkiss, A. and GilliganLee, K.A. and Thomas, Michael S.C. and Tolmie, A.K. and Farran, E.K. (2021) The developmental trajectories of spatial skills in middle childhood. British Journal of Developmental Psychology , ISSN 0261-510X. (In Press)

British Journal of Developmental Psychology (2021)
2021 The Authors. British Journal of Developmental Psychology published by
John Wiley & Sons Ltd
on behalf of British Psychological Society.

www.wileyonlinelibrary.com

The developmental trajectories of spatial skills in middle childhood

Alex Hodgkiss*[1], Katie A. Gilligan-Lee[2], Michael S.C. Thomas[3], Andrew K. Tolmie[4] and Emily K. Farran[5]

[1]Department of Education, University of Oxford, UK
[2]School of Psychology, University of Surrey, Guildford, UK
[3]Department of Psychological Sciences, Birkbeck, University of London, UK
[4]Department of Psychology and Human Development, UCL Institute of Education, UCL, London, UK
[5]School of Psychology, University of Surrey, Guildford, UK

The multidimensional structure of spatial ability remains a debated issue. However, the developmental trajectories of spatial skills have yet to be investigated as a source of evidence within this debate. We tested the intrinsic versus extrinsic and static versus dynamic dimensions of the Uttal et al. (2013, *Psychol. Bull.*, *139*, 352) typology in relation to spatial development. Participants ($N = 184$) aged 6–11 completed spatial tasks chosen to measure these spatial dimensions. The results indicated that the developmental trajectories of intrinsic versus extrinsic skills differed significantly. Intrinsic skills improved more between 6 and 8 years, and 7 and 8 years, than extrinsic skills. Extrinsic skills increased more between 8 and 10 years than intrinsic skills. The trajectories of static versus dynamic skills did not differ significantly. The findings support the intrinsic versus extrinsic, but not the static versus dynamic dimension, of the Uttal et al. (2013, *Psychol. Bull.*, *139*, 352) typology.

Statement of contribution

What is already known on this subject?

- The dimensional structure of spatial ability is a debated issue.
- The Uttal et al. (2013) model proposes that spatial thinking is comprised of two dimensions.
- There is a scarcity of developmental findings assessing the validity of these dimensions.

What does this study add?

- The developmental trajectories of intrinsic versus extrinsic skills differ significantly.
- The developmental trajectories of static versus dynamic skills do not differ significantly.
- Spatial skill developmental trajectories support the intrinsic versus extrinsic dimension only.

**Correspondence should be addressed to Alex Hodgkiss, Department of Education, University of Oxford, 15 Norham Gardens, Oxford OX2 6PY, UK (email: alex.hodgkiss@education.ox.ac.uk).*

DOI:10.1111/bjdp.12380

Background

Spatial cognition is the processing, representation, comparison, and transformation of spatial information. Spatial cognition was first distinguished from general intelligence in the 1930s, and since this time, attempts at defining a typology for spatial thinking have led to the emergence of many contrasting typologies (for example, Linn & Petersen, 1985). While these models all assume that spatial cognition is composed of several dimensions, the precise nature of these remains hotly debated. To date, evidence for different spatial dimensions has been mostly derived from psychometric analyses, as well as findings from experimental cognitive psychology and cognitive neuroscience. However, the similarities and differences between spatial skill developmental trajectories in childhood are currently an untapped source of evidence for separable spatial dimensions. The aim of the current study was therefore to investigate the developmental trajectories of different spatial dimensions through middle childhood, as a means of testing the proposed dimensions of a prominent spatial typology.

Typology of spatial thinking

Among the spatial typologies proposed, one theory-driven approach has gained significant support. Uttal et al. (2013), and Newcombe and Shipley (2015), proposed a classification of spatial thinking which distinguishes skills as being intrinsic versus extrinsic along one dimension, and static versus dynamic, along the other (see Figure 1). Intrinsic skills are within-object, that is, pertaining to the size and orientation of an object, its parts, and their relationships, and extrinsic skills are between-object, that is, relating to the relationship between objects, and between objects and their frames of reference. Dynamic skills involve movement or transformation, for example, imagined rotation or folding, whereas static skills do not, and involve object representation only.

Previous evidence supports the intrinsic versus extrinsic dimension. From an evolutionary perspective, it has been proposed that humans have two primary spatial functions, tool use, and navigation, each with a distinct evolutionary and neural basis (Newcombe, 2018). Tool use involves intrinsic spatial relations, whereas navigation involves the extrinsic coding of relations between objects (e.g., the car is next to the tree) and between objects and wider frames of reference (e.g., the car is on the west side of the mountain). Behavioural evidence also supports the intrinsic–extrinsic distinction in adults. Hegarty and Waller (2004) performed a confirmatory factor analysis (CFA) of three intrinsic tasks (mental rotation) and three extrinsic tasks (perspective taking measures

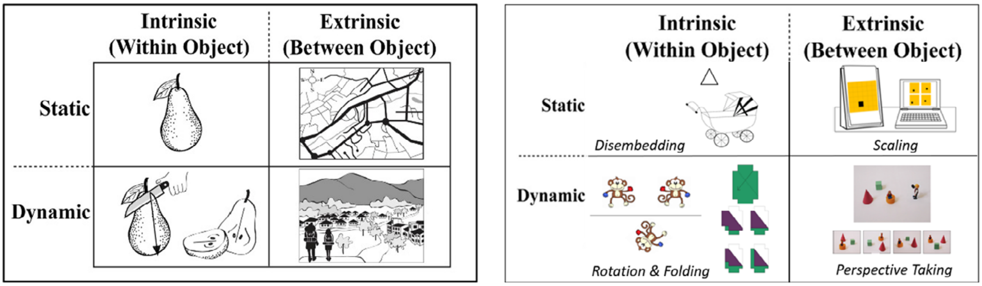

Figure 1. On the left, the Uttal et al. (2013) framework of spatial thinking (source: Newcombe, 2018). On the right, the tasks included in this study to assess each of the Uttal et al. (2013) spatial dimensions.

requiring visualisation of a scene from different vantage points). It was revealed that a two-factor model, that is, intrinsic versus extrinsic, fitted the data better than a one-factor model, that is, all tasks measuring a single spatial skill (Hegarty & Waller, 2004). There is also evidence that real-world navigation is more closely related to spatial perspective taking than mental rotation (Kozhevnikov, Motes, Rasch, & Blajenkova, 2006), as is self-reported sense of direction (Kozhevnikov & Hegarty, 2001).

In child populations, Mix, Hambrick, Satyam, Burgoyne, and Levine (2018) tested Uttal et al.'s (2013) classification of spatial skills using CFA in 6-, 9-, and 12-year-olds. At 6 and 9 years, a two-factor model distinguishing between intrinsic skills and extrinsic skills fitted the data better than a one-factor model. At 12 years, none of the more complex models fitted better than a one-factor model. However, a one-factor model did not fit the data well, suggesting that spatial cognition does not have a unitary structure for this age group. The lack of support for the more complex models (i.e., intrinsic vs. extrinsic) for the older children in Mix et al. (2018) may relate to the particular choice of spatial tasks for this age group within this study. In line with this, using different spatial tasks than Mix et al. (2018), Vander Heyden, Huizinga, Kan, and Jolles (2016), and Heil (2018) showed through CFA that a two-factor model distinguishing between intrinsic (mental rotation, mental folding) and extrinsic (navigating through a route after a change of perspective) skills fitted the data better than a one-factor model, for children aged 10.5 years.

In terms of brain-based evidence, data from adults suggest that intrinsic and extrinsic tasks are associated with activation of dissociable but overlapping neural systems. Intrinsic tasks (e.g., mental rotation) activate the right temporo-parietal cortices and visuospatial cortical areas, whereas extrinsic tasks (e.g., perspective taking) activate the left temporo-parietal cortices and motor areas (Wraga, Shephard, Church, Inati, & Kosslyn, 2005; Zacks, Vettel, & Michelon, 2003). FMRI research indicates that perspective taking shows similar patterns of brain activation to navigation, such that both activate the retrosplenial cortex and hippocampus (Lambrey, Doeller, Berthoz, & Burgess, 2012).

There is less convincing evidence for the static versus dynamic dimension. Assessing the static versus dynamic dimension is confounded by the fact that, for many tasks, static skills may be a necessary prerequisite to dynamic skills; that is, it is necessary to encode a shape before mentally transforming it. In adults, evidence for the static versus dynamic distinction comes from behavioural studies. For example, adults can be separated into object visualizers, who have significantly higher performance on static tasks than dynamic tasks, and spatial visualizers, who show the opposite pattern (Kozhevnikov, Kosslyn, & Shephard, 2005). However, in child-based studies, the aforementioned CFA study by Mix et al. (2018) found no support for a two-factor static-dynamic model. No other known studies explore the static versus dynamic dimension in childhood.

To summarise, there is convincing evidence of an intrinsic versus extrinsic distinction in spatial thinking. For static versus dynamic skills, there is little evidence that these dimensions psychometrically dissociate. However, if static skills (perceiving and encoding static images) are a prerequisite for dynamic skills, static skills may show earlier development and an earlier plateau, than dynamic skills. We hypothesize that the current study may therefore reveal age-based differences along the static and dynamic dimension of spatial thinking, which has not been previously evident within psychometric analyses. In the current study, spatial tasks were selected based on Uttal et al.'s (2013) theoretical framework of spatial cognition (see also Newcombe & Shipley, 2015).

The development of spatial skills

Whilst the developmental structure of the Uttal et al. (2013) model has yet to be directly tested, below we provide a short summary of the current literature on the development of individual spatial skills. Given that there is currently a larger evidence base for the intrinsic versus extrinsic dimension, we have organized the literature along these lines.

Intrinsic skills

The most simplistic form of intrinsic spatial representation requires the 'coding of spatial features of objects, including their size and the arrangement of their parts' (Newcombe & Shipley, 2015, p. 6), that is, encoding intrinsic-static representations. The Children's Embedded Figures Task (CEFT) measures intrinsic-static spatial thinking and requires individuals to identify a shape that is embedded in a more complicated image (Ekstrom, French, & Harman, 1976; Okamoto, Kotsopoulos, McGarvey, & Hallowell, 2015; Witkin, Otman, Raskin, & Karp, 1971). There is evidence that children successfully complete preschool versions of the CEFT by 3 years, and performance continues to improve on the main version of the CEFT between 3 and 5 years and also until 10 years (Busch, Watson, Brinkley, Howard, & Nelson, 1993; Witkin et al., 1971).

Other forms of intrinsic spatial thinking require visualizing and mentally transforming 2D and 3D objects (Newcombe & Shipley, 2015), that is, intrinsic-dynamic skills. These tasks can be further subdivided into rigid and non-rigid transformations. For rigid transformations, distances between points on an object are preserved during the transformation, for example, mental rotation. Data from violation of expectation paradigms indicate that early precursors of successful mental rotation emerge at 16 months (Frick & Wang, 2014). Results from studies using imagined rotations, only report above chance accuracy on mental rotation tasks from 5 years of age (Broadbent, Farran, & Tolmie, 2014; Frick, Hansen, & Newcombe, 2013; Marmor, 1975, 1977); Okamoto-Barth & Call, 2008). Crescentini, Fabbro, and Urgesi (2014) found that basic 2D rotation skills improve significantly between 7 and 8 years, with no significant improvement thereafter. However, for 3D rotation, performance continues to improve until 10 (Vander Heyden et al., 2016) or even 13 years (Johnson & Meade, 1987). Similar findings have been reported for other rigid transformation tasks, for example, Child Mental Transformation Task (Levine, Huttenlocher, Taylor, & Langrock, 1999). There is less research exploring the development of non-rigid transformations where the distance between points changes as the transformation occurs (Atit, Shipley, & Tikoff, 2013). There is evidence that by 5 years the majority of children demonstrate above chance performance on mental folding (imagining an object after it has been folded) which improves until 7–8[1] years (Harris, Hirsh-Pasek, & Newcombe, 2013).

Overall, the findings for the development of intrinsic-dynamic skills indicate that the precursors of successful intrinsic-dynamic spatial thinking are evident in infancy; 2D intrinsic-dynamic skills, measured using both rigid and non-rigid transformations, continue to develop until at least 7–8 years; and 3D rigid intrinsic-dynamic skills continue to develop through later childhood. In the current study, two intrinsic-dynamic tasks were included to reflect rigid and non-rigid transformations.

---

[1] *The oldest age group of children included in this study was 8 years.*

Extrinsic skills

Considering next extrinsic skills, extrinsic-static spatial thinking involves the coding of object locations in relation to other objects, spatial frameworks, or landmarks. Historically, these skills were assessed through horizontal and vertical invariance tasks. For example, performance on the Rod and Frame Test, which tests the ability to accurately code horizontal and vertical dimensions of a rod as defined by gravity, while ignoring the reference of a tilted frame, improves with age from 4 years until adulthood (Bagust, Docherty, Haynes, Telford, & Isableu, 2013; Haywood, Teeple, Givens, & Patterson, 1977; Newcombe & Shipley, 2015). Spatial scaling tasks, which measure the ability to successfully map encoded distances between different sized spaces, also assess extrinsic-static spatial skills. In localisation paradigms, participants are shown the location of a target and are asked to find the corresponding location on a scaled target space. Using a 2D localization task, Frick and Newcombe (2012) reported that children's scaling ability improves with age from 3 to 6 years. In more naturalistic environments, Vasilyeva and Huttenlocher (2004) found that 90% of 5-year-olds but only 60% of 4-year-olds could successfully place objects on a rectangular rug using a 2D map. The spatial scaling data presented in the current study are based on a previously published discrimination paradigm (Gilligan, Hodgkiss, Thomas, & Farran, 2018) where participants determined which one of four referent maps corresponded to a scaled version of a model map. It was found that scaling performance improved between 5 and 8 years, with no significant improvement thereafter.

A second component of extrinsic spatial thinking involves visualizing an environment from a different position (Uttal et al., 2013), that is, extrinsic-dynamic spatial skills. Perspective taking tasks require the ability to use an object-based (allocentric) reference frame, to represent a viewpoint that differs from one's own (Frick, Möhring, & Newcombe, 2014). Perspective taking is proposed to develop in two stages. During Level 1, a child understands that another person can see something different to themselves, but cannot imagine exactly what can be seen from a contrasting view point (Flavell, Everett, Croft, & Flavell, 1981; Newcombe & Huttenlocher, 2003). Level 1 skills have been reported in children from 24 months (Moll & Tomasello, 2007). Level 2 perspective taking is the ability to determine *exactly how* another person would perceive an object, or array of objects, from a different perspective. Level 2 perspective taking emerges from around 4–5 years and continues to develop until at least 8 years (Frick et al., 2014; Pillow & Flavell, 1986).

The literature reviewed above suggests that there may be subtle differences in the early developmental profiles of each spatial dimension. However, few studies compare performance on different measures of spatial thinking with the same sample of participants and there is limited evidence on the development of spatial skills beyond 8 years.

Gender differences in spatial task performance

Previous studies of gender differences in spatial skill in middle childhood report mixed findings that differ by age, by the spatial skills assessed and by the tasks used (Newcombe, 2020). In a meta-analysis, Lauer, Yhang, and Lourenco (2019) reported that small gender effects in mental rotation, that favour males, may be present from 6 years (Carr, Steiner, Kyser, & Biddlecomb, 2008). There is evidence that these gender differences may also be sensitive to development. For example, Neuburger, Jansen, Heil, and Quaiser-Pohl (2011) found small gender effects in mental rotation favouring boys at 10, but not at 8 years. Other intrinsic-dynamic tasks such as mental folding (Harris et al., 2013), mental transformation (Frick, 2019), and pattern construction (Gilligan, Flouri, & Farran, 2017) do not show a male advantage in middle childhood. Indeed, Gilligan et al. (2017) reported a small female advantage in pattern construction at 7 years. Prior research indicates that performance on the Children's Embedded Figures Task (Witkin et al., 1971) shows either no gender differences in children aged 5–10 (Morris, Farran, & Dumontheil, 2019), or a female advantage in children aged 8-to-9 years (Kaplan & Weisberg, 1987). For extrinsic tasks, no gender differences have been reported in spatial perspective taking (Frick, 2019; Frick et al., 2014) or spatial scaling (Frick, 2019) in primary school-aged children. Taken together, although some prior research highlights gender effects in spatial performance that may increase through development, the size of these effects is typically small in primary school-aged children (Nazareth, Huang, Voyer, & Newcombe, 2019; Newcombe, 2020). Furthermore, an increasing number of studies do not report gender-based effects. However, given that spatial task performance has sometimes been shown to differ by gender in middle childhood, it is important to account for this effect to provide an accurate analysis of age-based differences.

Current study

In the current study, we determine whether there are significant age-based differences in intrinsic versus extrinsic, and static versus dynamic spatial skills, respectively. The findings from this study will provide a novel source of evidence to assess the spatial dimensions of the Uttal et al. (2013) model. Individual tasks were included to assess each of Uttal et al.'s (2013) proposed spatial dimensions (Figure 1). Our analysis contrasts with existing studies to date which test the model using factor analysis. We have two main hypotheses. First, we predict significant age-based differences in intrinsic versus extrinsic spatial skills, with extrinsic skills continuing to develop into later childhood (significant age-based differences between older age groups), even after 2D intrinsic skills have been acquired[2] (no significant age-based differences after 8 years). Second, we predict significant age-based differences in static (significant age-based differences between younger age groups only) and dynamic (significant age-based differences between younger age groups as well as significant differences between older age groups) spatial skills. Finding a significant difference in the developmental trajectories of intrinsic versus extrinsic, and static versus dynamic skills, would support both dimensions of the Uttal et al. (2013) model.

Materials and methods

Participants

Participants were 185 children from a large, culturally diverse primary school in the UK. The eligibility for free school meals was 19%, slightly above the national average of 15% (Department for Education, 2019). Due to technical errors, 6 participants did not have a full set of scores available for analysis. Five of these participants were missing data for one task only, and to maximize statistical power, their missing scores were substituted using mean replacement, that is, replacing their missing score with the mean value on that task for their age group (missing data: two mental folding scores, two perspective taking scores, one CEFT score). The proportion of data replaced by mean scores was 0.005%. The sixth participant was missing data for several variables and was excluded. The final sample

---

[2] *3D rigid intrinsic-dynamic skills, which were not tested in this study, continue to develop into late childhood (Johnson & Meade, 1987); this prediction therefore applies only to 2D skills.*

Table 1. Demographic information of participants across age groups

Age group	*N*	% Male	Age years (mean ± *SD*)
6 years	30	53.33	6.00 ± 0.34
7 years	31	41.94	6.99 ± 0.29
8 years	32	56.25	8.03 ± 0.28
9 years	31	45.16	8.97 ± 0.32
10 years	31	51.61	9.95 ± 0.33
11 years	29	58.62	11.00 ± 0.30

therefore consisted of 184 participants across six age groups (see Table 1). A power analysis, conducted in GPower, indicated that a total sample size of 140 was needed to detect a medium to large effect (0.3) for the main developmental analyses (one-way ANOVA's with age as a between participant variable; power = 0.8, $\alpha = .05$; Table 1).

Procedure

This study was part of a larger investigation of the role of spatial thinking for mathematics and science. More information can be found in Gilligan et al. (2019) and Hodgkiss et al. (2018). Each participant completed five testing sessions in a set order. There was approximately 2 days between testing sessions. The Mental Rotation Task and the Mental Folding Task were completed in a group testing session (8 participants per group), lasting approximately 35 min. The order of these tasks was counterbalanced. Each group testing session was supervised by at least two researchers. The CEFT, the Perspective Taking Task, and the Scaling Task were completed in an individual testing session, lasting approximately 45 min. The order of these tasks was counterbalanced. In the remaining three sessions, participants completed mathematics and science tasks, not described here.

Measures

Children's Embedded Figures Task—CEFT

The Children's Embedded Figures Task (CEFT; Witkin et al., 1971) is a paper-based task that uses physical shapes. The task requires participants to locate a target shape within a more complex figure. The task was administered as per the administration guidelines (Witkin et al., 1971). The task contained a maximum of two blocks, presented in a fixed order. Participants were first shown the target shape, a 'tent' shape for block A, and a 'house' shape for block B. Before the practice and experimental trials, participants were familiarized to the shape through four discrimination trials, where they were required to identify the target shape from a selection of other shapes. Participants repeated the discrimination trials until two items were answered correctly in succession. After this, participants completed either two (block A) or one (block B) practice trials, where they located the target shape hidden within a more complex image. Participants outlined the shape with their finger to indicate their answer. Participants were required to successfully locate the target shape in each practice item, before progressing to the experimental trials. Following the practice trials, participants completed 11 experimental items in block A and 14 experimental items in block B. For experimental items, they again were required to locate the target shape hidden with a larger more complex image. Performance was measured as percentage accuracy. Cronbach's alpha for this task is high: .86 in children aged 6–11 years (Amador-Campos & Kirchner-Nebot, 1997).

Mental Rotation Task

This task was modified from Broadbent et al. (2014). Participants were required to identify which of two monkeys above a horizontal line displayed on screen, matched the target monkey below the horizontal line (see Figure 2). One monkey above the line was rotated by a fixed degree, relative to the target monkey. One monkey above the line was a mirror image of the target monkey. Participants indicated their answer using labelled keys on the computer keyboard. Participants completed four practice trials at 0°, that is, where the monkey below the horizontal line was not rotated. All participants passed the practice trials, that is, achieved 50% or higher on their first attempt. Participants next completed 36 experimental trials, 8 × 45° trials, 8 × 90° trials, 8 × 135° trials, and 8 × 180° trials and 4 trials at 0°. Equal numbers of clockwise and anticlockwise rotations were included. Performance was measured as percentage accuracy. Cronbach's alpha reliability score for this task was high, $\alpha = .88$ (Field, 2013; Figure 3).

Mental Folding Task

This task was taken from Harris et al. (2013). Participants were required to imagine folds made to a shape presented onscreen, without the physical representation of the fold itself (a physical piece of paper). Participants were shown a green shape on a computer screen (see Figure 2). The shape included a dotted line, which represented the folding line, and an arrow, which indicated the direction and distance of the required fold. Below this shape were four possible response options. Only one of these four response options (the target item) showed the outcome of the fold correctly. Participants completed two practice items, in which they were given a physical card to check their answer. If a child indicated an incorrect option, they were given one further attempt to answer the practice items. Most participants answered their practice items correctly on the first attempt and all participants answered correctly by the second attempt. Following the practice items, participants completed 14 experimental items. Performance was measured as percentage accuracy. Cronbach's alpha reliability score for this task was medium/high, $\alpha = .74$ (Field, 2013).

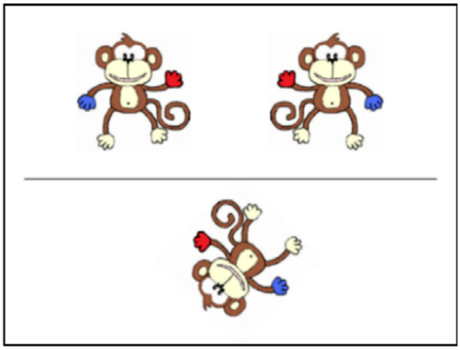

Figure 2. 135° anticlockwise mental rotation trial.

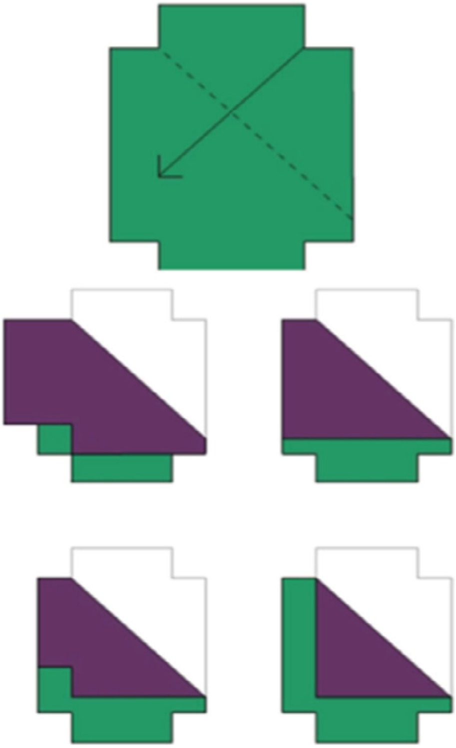

Figure 3. Sample trial from the Mental Folding Task.

Spatial scaling

In this task, children were required to locate the corresponding locations of hidden treasure (a black box) on two maps, when one was varied in size relative to the other (Gilligan et al., 2018). Participants were shown four treasure maps on a touch screen computer. To the left of the computer, children were presented with a printed treasure map (Figure 4). The child had to determine which of the four maps on the computer screen had the treasure positioned in the same place as the larger printed map. The three incorrect options were created uniformly for each trial. Trials differed by scaling factor. The printed maps were either unscaled with a ratio of 1:1 (7 cm × 7 cm), or, were scaled, to either a ratio of 1:2 (14 cm × 14 cm) or 1:4 (28 cm × 28 cm), relative to the maps on the computer (7 cm × 7 cm). The required level of visual acuity also differed across trials. At each scaling factor, the overall area of the maps, and by extension the scaling factor, did not change. However, half of the items were presented using a 6 × 6 square grid and therefore required gross level acuity, while the other half were presented using a 10 × 10 square grid and therefore required fine level acuity. Participants first completed two unscaled practice items. All participants passed the practice trials, that is, answered at least one of the practice items correctly on their first attempt and thus continued to the main trials. Participants completed 18 experimental trials including six items at each scaling factor. Performance was measured as percentage accuracy. Cronbach's alpha reliability score for this task was medium/high, $\alpha = .74$ (Field, 2013).

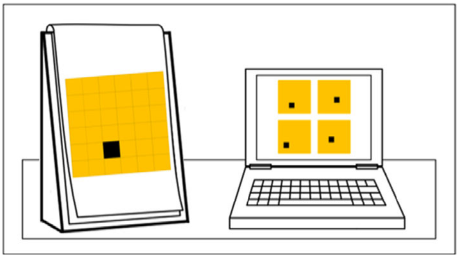

Figure 4. Spatial scaling trial at a scaling factor of 1:4.

Perspective taking task for children

In this task, participants visualised what photographs would look like when taken from cameras placed at different angles and positions relative to their viewpoint (Frick et al., 2014; Figure 5). Participants completed four practice questions with physical Playmobil characters (one holding a camera) in a specified arrangement on a table. For each practice trial, the participant was shown a photograph and was asked which of the two characters could have taken the photograph. Participants were able to check their answers by moving around the table to be positioned at the photographer's perspective. Feedback was given on each practice item. If a participant made an error on any of the practice items, they were allowed a maximum of one additional attempt. Few participants made errors on the first attempt and all participants passed on their second, if one was needed. Participants completed 18 experimental trials. The angular difference between the photographer's and the participant's perspective varied across experimental trials (0°, 90° or 180°). The number of objects in the layout (1, 2, or 3) also varied across experimental trials. Cronbach's alpha reliability score for this task was high, $\alpha = .84$ (Field, 2013).

Analysis strategy

There were no significant floor or ceiling effects for any task, for any age group. We then created intrinsic, extrinsic, static, and dynamic composite variables by calculating the mean z-score for the tasks which were categorized as assessing that spatial dimension:

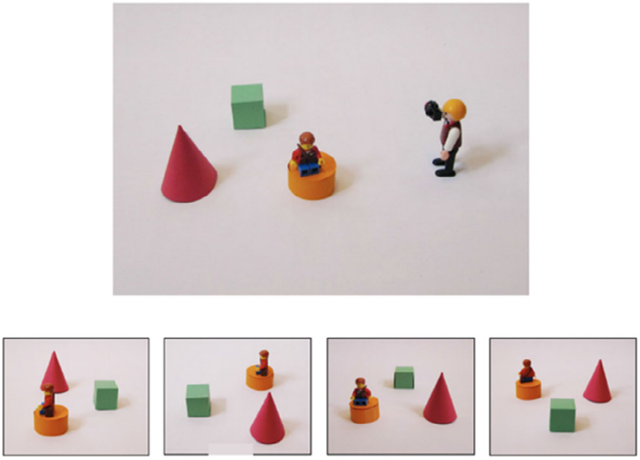

Figure 5. Perspective taking, 90°, three object trial.

intrinsic (CEFT, mental rotation, and mental folding); extrinsic (spatial scaling and perspective taking); static (CEFT and spatial scaling); and dynamic (mental rotation, mental folding and perspective taking; see Figure 1).

All variables were broadly normal based on visual inspection of Q–Q plots and boxplots. Gender effects were investigated through Bonferroni-corrected *t*-tests for each spatial dimension ($0.05/4 = .013$). We ran two mixed ANOVAs to investigate age-based differences in performance across spatial dimensions. Age group was the between-group factor and the intrinsic and extrinsic composite variables (or static and dynamic composite variables respectively) were the within-participant variables. An interaction between age and spatial dimension in either model would indicate that the trajectories of the spatial dimensions differed across age groups and would therefore support either or both dimensions of the Uttal et al. (2013) model.

Significant interactions were first explored by comparing the effect sizes of Tukey post-hoc tests. Two follow-up ANOVAs were also completed to statistically compare relatively younger and older children within two broader age categories. These ANOVAs only included children aged 6, 8, 9 and 11. The children were divided into two broad age categories: *Younger: 6- and 8-year-olds; Older: 9- and 11-year-olds*. The children were further categorized within these broad categories as being relatively younger or relatively older (e.g., within the broad category of younger, 6-year-olds were coded as relatively younger and 8-year-olds as relatively older). A significant interaction between broad and relative age grouping would indicate that age-based differences were greater within one of the broader age categories than the other.

Results

Analysis of gender effects

After applying a Bonferroni correction ($0.05/4 = .013$), there were no significant gender differences in spatial performance for any of the spatial dimensions ($p > .013$; $d < 0.360$ for all). In addition, no interactions between gender and age were found when gender was included within the main ANOVAs, reported below ($p > .05$; $\eta_p^2 < .005$ for all). Gender was therefore not included as a factor in subsequent analyses.

Age-based differences in intrinsic versus extrinsic spatial dimension

Descriptive statistics for each spatial task across each age group, including both raw scores and *z*-scores, can be found in the Supporting Information. For the ANOVA comparing the intrinsic versus extrinsic spatial dimension by age group, there was a main effect of age group, $F(5, 178) = 32.758, p < .001, \eta_p^2 = .479$, and a significant interaction between age group and spatial dimension, $F(5, 178) = 2.577$, $p = .028$, $\eta_p^2 = .068$ (Figure 6). Tukey post-hoc tests revealed subtle differences in the development of spatial skills across the intrinsic and extrinsic spatial dimension (see Table 2 for full results). For the youngest children, performance increased significantly between 6 and 8 years for both intrinsic ($p < .001$) and extrinsic skills ($p < .001$), although the effect was larger for the intrinsic ($d = 2.047$), than the extrinsic dimension ($d = 1.105$). Between 7 and 8 years, there was a significant improvement in performance for intrinsic ($p < .001$, $d = 1.393$) but not extrinsic skills ($p = .103, d = 0.655$). For older age groups, there was a significant increase in performance between 8 and 11 years for both intrinsic ($p = .048$; $d = 0.743$) and extrinsic skills ($p = .002$; $d = 0.988$). Between 9 and 11 years, there was a significant difference in performance for intrinsic ($p = .035, d = 0.777$) but not extrinsic skills ($p = .49$, $d = 0.457$). In contrast, there was a significant increase between 8 and 10 years for extrinsic skills ($p = .017$, $d = 0.820$) but not intrinsic skills ($p = .99$, $d = 0.165$). Overall, the interaction between age and spatial dimension was driven by a steep rate of early development for intrinsic skills and a slower and later development of extrinsic skills.

The additional ANOVAs exploring aged-based differences for younger versus older children revealed for intrinsic skills a significant interaction between broad and relative age groupings, $F(1, 118) = 13.8, p < .001, \eta_p^2 = .105$. The interaction was driven by the difference between the 6- and 8-year-olds (−1.20) being significantly greater than the 9- and 11-year-olds (−0.46). For extrinsic skills, there was no significant interaction between broad and relative age groupings, $F(1, 118) = 3.07, p = .082, \eta_p^2 = .025$.

Age-based differences in static versus dynamic spatial dimension

For the ANOVA of static versus dynamic skills by age group, there was a main effect of age group, $F(5, 178) = 34.190, p < .001, \eta_p^2 = .490$; the interaction between age group and spatial dimension was not statistically significant, $F < 1$ (Figure 7).

Discussion

This study is the first to use developmental trajectories as a source of evidence to assess the spatial dimensions of the Uttal et al. (2013) model. Subtle differences in the development of intrinsic versus extrinsic spatial skills were found. This was not the case for the static versus dynamic dimension. The findings in the current study complement prior psychometric data, for example, Mix et al. (2018) which reported a distinction between intrinsic versus extrinsic spatial skills, but not between static versus dynamic spatial skills.

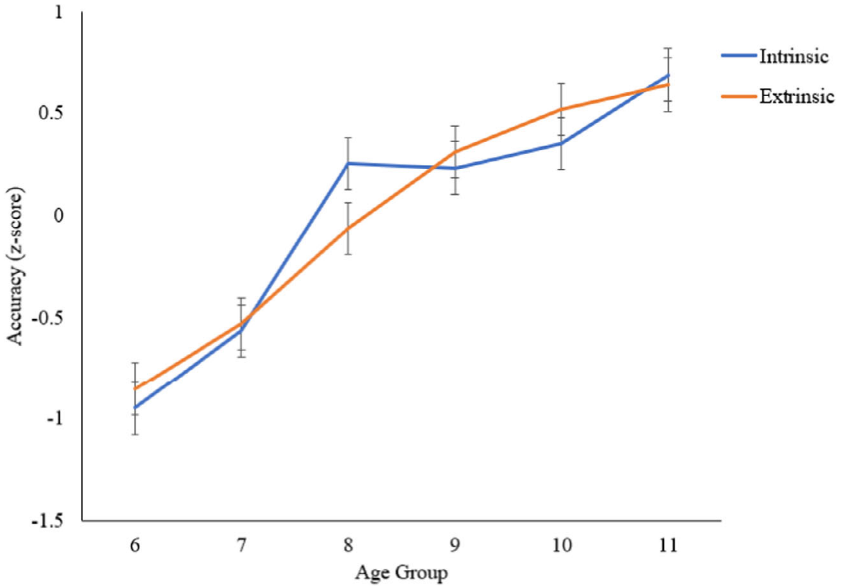

Figure 6. Intrinsic and extrinsic composite score accuracy by age group.

Table 2. Summary of Tukey post-hoc tests by age group

		Comparison		Intrinsic			
Extrinsic			Age groups		Mean difference	Tukey, *p*	Tukey, *d*
		Mean difference	Tukey, *p*	Tukey, *d*			
6	7	−0.38	.115	−0.654	−0.32	.496	−0.450
	8	−1.20	<.001	−2.047	−0.79	<.001	−1.105
	9	−1.18	<.001	−2.012	−1.17	<.001	−1.635
	10	−1.30	<.001	−2.212	−1.37	<.001	−1.925
	11	−1.64	<.001	−2.789	−1.49	<.001	−2.093
7	8	−0.82	<.001	−1.393	−0.47	.103	−0.655
	9	−0.80	<.001	−1.358	−0.85	<.001	−1.185
	10	−0.91	<.001	−1.558	−1.05	<.001	−1.475
	11	−1.25	<.001	−2.136	−1.17	<.001	−1.643
8	9	0.02	1.000	0.035	−0.38	.290	−0.530
	10	−0.10	.987	−0.165	−0.59	.017	−0.820
	11	−0.44	.048	−0.743	−0.70	.002	−0.988
9	10	−0.12	.970	−0.200	−0.21	.863	−0.290
	11	−0.46	.035	−0.777	−0.33	.487	−0.457
10	11	−0.34	.226	−0.578	−0.119	.987	−0.167

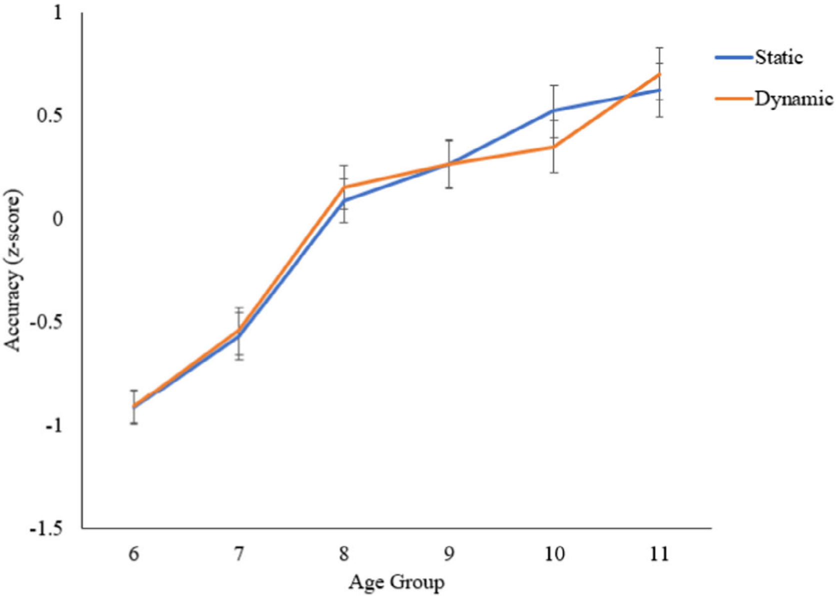

Figure 7. Static and dynamic composite score accuracy by age group.

Age-based differences were found for extrinsic compared to intrinsic spatial skills. Analysis of effect sizes suggested that the differences in trajectories were driven by a steep rate of early development for intrinsic skills and a slower and later development for extrinsic skills.

Performance increased more between 6 and 8, and 7 and 8 years, for intrinsic spatial skills, than for extrinsic spatial skills. In contrast, extrinsic spatial performance differed more between 8 and 10 years, compared with intrinsic spatial performance. There was no evidence that the trajectories of static versus dynamic skills differed significantly. This contrasted with our hypothesis that static skills (perceiving and encoding static images) may be a prerequisite for dynamic skills. Therefore, while static skills (e.g., disembedding) may show earlier development than dynamic skills (e.g., mental rotation) in early childhood, between 3 and 6 years (Busch et al., 1993; Witkin et al., 1971), there is no evidence from our study that this developmental dissociation continues into middle childhood. Alternatively, the static and dynamic tasks that were included in our study may not have been sensitive enough to highlight subtle developmental differences beyond 6 years.

The differing developmental patterns between intrinsic and extrinsic spatial skills lends support to the intrinsic versus extrinsic spatial dimension in Uttal et al.'s (2013) model of spatial thinking. The results reported here show that performance on tasks measuring the intrinsic versus extrinsic spatial dimension differ developmentally for intrinsic and extrinsic skills, thus suggesting that they are distinct constructs. The results align with the aforementioned CFA study by Mix et al. (2018) that found stronger evidence for the intrinsic versus extrinsic, compared to the static versus dynamic distinction of spatial thinking. These findings are also consistent with previous research that demonstrated that the developmental trajectories of mental rotation (intrinsic skill) and perspective taking (extrinsic skill) differed such that mental rotation developed earlier (from age 7 years) than perspective taking (from age 8 years; Crescentini et al., 2014). Here, we extend these findings and show earlier development of intrinsic compared to extrinsic skills using more comprehensive measures of the intrinsic versus extrinsic dimension.

Although significant age-based differences are reported in this study, the findings also highlight substantial individual differences in spatial performance (reflected by the range of z-scores, standard deviations, and standard errors) for both of Uttal et al.'s (2013) spatial dimensions. The roles of both development and individual differences must be considered in any discussion of spatial thinking. For example, one demographic factor that was explored in this study was gender. While some previous studies have outlined a male advantage in spatial task performance in childhood (Carr et al., 2008; Casey et al., 2008; De Lisi & Wolford, 2002; Lauer et al., 2019; Neuburger et al., 2011; Wiedenbauer & Jansen-Osmann, 2008), in this study no significant gender differences were found and all effect sizes reported for gender were small ($d < 0.360$). This finding is consistent with other recent studies that do not report a male advantage for spatial outcomes (Frick, 2019; Frick et al., 2014; Gilligan et al., 2017; Harris et al., 2013).

This study is not withstanding limitations. Owing to the cross-sectional design, it was not possible to compare individuals' performance across time. These findings could be strengthened by replication using a longitudinal design. The findings of this study provide a comparison of spatial performance between spatial dimensions, in this age range, for the first time. From a practical perspective, the results highlight developmental ages when spatial tasks may be particularly challenging and when scaffolding or spatial training, in specific skills, may be particularly beneficial.

To conclude, this is the first study to use developmental trajectories as a source of evidence to assess the Uttal et al. (2013) model. We showed that there were age-based differences in intrinsic and extrinsic spatial performance, such that intrinsic skills demonstrated particularly rapid early development in middle childhood (6–8 years) compared to extrinsic skills. In contrast, there were larger differences in extrinsic spatial performance between 8 and 10 years. There were no age-based differences in static versus dynamic spatial performance. The findings therefore lend support to the intrinsic versus extrinsic dimension of the Uttal et al. (2013) model only. By exploring spatial performance in older children, the results demonstrate for the first time that some spatial subdomains continue to develop at least until 11 years.

Acknowledgements

Funding for this research was provided by the Bloomsbury Colleges Ph.D Scholarship Programme and the Economic and Social Research Council (ESRC). The research was also supported by the National Centre for Curriculum and Assessment (NCAA) Ireland. We would like to thank the schools and children who took part in this research.

Conflicts of interest

All authors declare no conflict of interest.

Author contributions

Alex Hodgkiss (Conceptualization; Formal analysis; Investigation; Methodology; Project administration; Writing – original draft; Writing – review & editing). Katie A. Gilligan-Lee (Conceptualization; Formal analysis; Investigation; Methodology; Project administration; Writing – original draft; Writing – review & editing). Michael S. C. Thomas (Conceptualization; Supervision; Writing – review & editing). Andrew K Tolmie (Conceptualization; Supervision; Writing – review & editing). Emily K Farran (Conceptualization; Methodology; Supervision; Writing – review & editing).

Data availability statement

Data available on request from the authors.

References

Amador-Campos, J. A., & Kirchner-Nebot, T. (1997). Relations of scores on children's embedded figures test with age, item difficulty and internal consistency. *Perceptual and Motor Skills*, *85*, 675–682. https://doi.org/10.2466/pms.1997.85.2.675

Atit, K., Shipley, T. F., & Tikoff, B. (2013). Twisting space: Are rigid and non-rigid mental transformations separate spatial skills? *Cognitive Processing*, *14*(2), 163–173. https://doi.org/10.1007/s10339-013-0550-8

Bagust, J., Docherty, S., Haynes, W., Telford, R., & Isableu, B. (2013). Changes in rod and frame test scores recorded in schoolchildren during development – A longitudinal study. *PLoS One*, *8*(5), e65321. https://doi.org/10.1371/journal.pone.0065321

Broadbent, H. J., Farran, E. K., & Tolmie, A. (2014). Object-based mental rotation and visual perspective-taking in typical development and Williams syndrome. *Developmental Neuropsychology*, *39*(3), 205–225. https://doi.org/10.1080/87565641.2013.876027

Busch, J. C., Watson, J. A., Brinkley, V., Howard, J., & Nelson, C. (1993). Preschool embedded figures test performance of young children: Age and gender differences. *Perceptual and Motor Skills*, *77*, 491–496. https://doi.org/10.2466/pms.1993.77.2.491

Carr, M., Steiner, H. H., Kyser, B., & Biddlecomb, B. (2008). A comparison of predictors of early emerging gender differences in mathematics competency. *Learning and Individual Differences*, *18*(1), 61–75. https://doi.org/10.1016/j.lindif.2007.04.005

Casey, B. M., Andrews, N., Schindler, H., Kersh, J. E., Samper, A., & Copley, J. (2008). The development of spatial skills through interventions involving block building activities. *Cognition and Instruction*, *26*(3), 269–309. https://doi.org/10.1080/07370000802177177

Crescentini, C., Fabbro, F., & Urgesi, C. (2014). Mental spatial transformations of objects and bodies: Different developmental trajectories in children from 7 to 11 years of age. *Developmental Psychology*, *50*, 370–383. https://doi.org/10.1037/a0033627

De Lisi, R., & Wolford, J. L. (2002). Improving children's mental rotation accuracy with computer game playing. *Journal of Genetic Psychology*, *163*(3), 272–282. https://doi.org/10.1080/00221320209598683

Department for Education. (2019). *Schools, pupils and their characteristics: January 2019*.

Ekstrom, R. B., French, J. W., & Harman, H. H. (1976). *Manual for kit of* (Vol. 102, Issue August). Educational Testing Service Princeton, NJ.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. London: Sage.

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

Frick, A. (2019). Spatial transformation abilities and their relation to later mathematics performance. *Psychological Research*, *83*, 1465–1484. https://doi.org/10.1007/s00426-018-1008-5

Frick, A., Hansen, M. A., & Newcombe, N. S. (2013). Development of mental rotation in 3- to 5-year-old children. *Cognitive Development*, *28*, 386–399. https://doi.org/10.1016/j.cogdev.2013.06.002

Frick, A., Möhring, W., & Newcombe, N. S. (2014). Picturing perspectives: Development of perspective-taking abilities in 4- to 8-year-olds. *Frontiers in Psychology*, *5*, 1–7. https://doi.org/10.3389/fpsyg.2014.00386

Frick, A., & Newcombe, N. S. (2012). Getting the big picture: Development of spatial scaling abilities. *Cognitive Development*, *27*(3), 270–282. https://doi.org/10.1016/j.cogdev.2012.05.004

Frick, A., & Wang, S. H. (2014). Mental spatial transformations in 14- and 16-month-old infants: Effects of action and observational experience. *Child Development*, *85*(1), 278–293. https://doi.org/10.1111/cdev.12116

Gilligan, K. A., Flouri, E., & Farran, E. K. (2017). The contribution of spatial ability to mathematics achievement in middle childhood. *Journal of Experimental Child Psychology*, *163*, 107–125. https://doi.org/10.1016/j.jecp.2017.04.016

Gilligan, K. A., Hodgkiss, A., Thomas, M. S. C., & Farran, E. K. (2018). The use of discrimination scaling tasks: A novel perspective on the development of spatial scaling in children. *Cognitive Development*, *47*, 133–145. https://doi.org/10.1016/j.cogdev.2018.04.001

Gilligan, K. A., Hodgkiss, A., Thomas, M. S., & Farran, E. K. (2019). The developmental relations between spatial cognition and mathematics in primary school children. *Developmental science*, *22*(4), e12786.

Harris, J., Hirsh-Pasek, K., & Newcombe, N. S. (2013). Understanding spatial transformations: Similarities and differences between mental rotation and mental folding. *Cognitive Processing*, *14*(2), 105–115. https://doi.org/10.1007/s10339-013-0544-6

Haywood, K., Teeple, J., Givens, M., & Patterson, J. (1977). Young children's rod-and-frame test performance. *Perceptual and Motor Skills*, *45*(1), 163–169. https://doi.org/10.2466/pms.1977.45.1.163

Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, *32*(2), 175–191. https://doi.org/10.1016/j.intell.2003.12.001

Heil, C. (2018). A dissociation between two classes of spatial abilities in elementary school children. *German Conference on Spatial Cognition*, 228–243.

Hodgkiss, A., Gilligan, K. A., Tolmie, A. K., Thomas, M. S., & Farran, E. K. (2018). Spatial cognition and science achievement: The contribution of intrinsic and extrinsic spatial skills from 7 to 11 years. *British Journal of Educational Psychology*, *88*(4), 675–697.

Johnson, E. S., & Meade, A. C. (1987). Developmental patterns of spatial ability: An early sex difference. *Child Development*, *58*, 725. https://doi.org/10.2307/1130210

Kaplan, B. J., & Weisberg, F. B. (1987). Sex differences and practice effects on two visual-spatial tasks. *Perceptual and Motor Skills*, *64*(1), 139–142. https://doi.org/10.2466/pms.1987.64.1.139

Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory and Cognition*, *29*, 745–756. https://doi.org/10.3758/BF03200477

Kozhevnikov, M., Kosslyn, S., & Shephard, J. (2005). Spatial versus object visualizers: A new characterization of visual cognitive style. *Memory and Cognition*, *33*, 710–726. https://doi.org/10.3758/BF03195337

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*, 397–417. https://doi.org/10.1002/acp.1192

Lambrey, S., Doeller, C., Berthoz, A., & Burgess, N. (2012). Imagining being somewhere else: Neural basis of changing perspective in space. *Cerebral Cortex*, *22*(1), 166–174. https://doi.org/10.1093/cercor/bhr101

Lauer, J. E., Yhang, E., & Lourenco, S. F. (2019). The development of gender differences in spatial reasoning: A meta-analytic review. *Psychological Bulletin*, *145*, 537–565. https://doi.org/10.1037/bul0000191

Levine, S. C., Huttenlocher, J., Taylor, A., & Langrock, A. (1999). Early sex differences in spatial skill. *Developmental Psychology*, *35*, 940–949. https://doi.org/10.1037/0012-1649.35.4.940

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development*, *56*, 1479–1498. https://doi.org/10.1111/j.1467-8624.1985.tb00213.x

Marmor, G. S. (1975). Development of kinetic images: When does the child first represent movement in mental images? *Cognitive Psychology*, *7*, 548–559. https://doi.org/10.1016/0010-0285(75)90022-5

Marmor G (1977). Mental rotation and number conservation: Are they related?. *Developmental Psychology*, *13*, (4), 320–325. http://dx.doi.org/10.1037/0012-1649.13.4.320.

Mix, K. S., Hambrick, D. Z., Satyam, V. R., Burgoyne, A. P., & Levine, S. C. (2018). The latent structure of spatial skill: A test of the 2 × 2 typology. *Cognition*, *180*, 268–278. https://doi.org/10.1016/j.cognition.2018.07.012

Moll, H., & Tomasello, M. (2007). How 14- and 18-month-olds know what others have experienced. *Developmental Psychology*, *43*, 309–317. https://doi.org/10.1037/0012-1649.43.2.309

Morris, S., Farran, E. K., & Dumontheil, I. (2019). Field independence associates with mathematics and science performance in 5- to 10-year-olds after accounting for domain-general factors. *Mind, Brain, and Education*, *13*(4), 268–278. https://doi.org/10.1111/mbe.12214

Nazareth, A., Huang, X., Voyer, D., & Newcombe, N. (2019). A meta-analysis of sex differences in human navigation skills. *Psychonomic Bulletin and Review*, *26*, 1503–1528. https://doi.org/10.3758/s13423-019-01633-6

Neuburger, S., Jansen, P., Heil, M., & Quaiser-Pohl, C. (2011). Gender differences in pre-adolescents' mental-rotation performance: Do they depend on grade and stimulus type? *Personality and Individual Differences*, *50*, 1238–1242. https://doi.org/10.1016/j.paid.2011.02.017

Newcombe, N. S. (2018). Three kinds of spatial cognition. In J.T Wixted *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience. Volume 3. Language and Thought.* (pp. , *3*, 1–31). Hoboken, NJ: John Wiley & Sons. https://doi.org/10.1002/9781119170174.epcn315

Newcombe, N. S. (2020). The puzzle of spatial sex differences: Current status and prerequisites to solutions. *Child Development Perspectives*, *14*(4), 251–257. https://doi.org/10.1111/cdep.12389

Newcombe, N. S., & Huttenlocher, J. (2003). *Making space: The development of spatial representation and reasoning*. Cambridge, MA: MIT Press.

Newcombe, N. S., & Shipley, T. F. (2015). Thinking about spatial thinking: New typology, new assessments. In J. Gero (Ed.), *Studying visual and spatial reasoning for design creativity* (pp. 179–192). New York: NY: Springer. https://doi.org/10.1007/978-94-017-9297-4_10

Okamoto, Y., Kotsopoulos, D., McGarvey, L., & Hallowell, D. (2015). The development of spatial reasoning in young children. In B. Davis *Spatial reasoning in the early years* (pp. 25–38). Abingdon: Routledge.

Okamoto-Barth, S., & Call, J. (2008). Tracking and inferring spatial rotation by children and great apes. *Developmental Psychology*, *44*, 1396–1408. https://doi.org/10.1037/a0012594

Pillow, B. H., & Flavell, J. H. (1986). Young children's knowledge about visual perception: Projective size and shape. *Child Development*, *57*(1), 125–135. https://doi.org/10.1111/j.1467-8624.1986.tb00013.x

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, *139*, 352–402. https://doi.org/10.1037/a0028446

Vander Heyden, K. M., Huizinga, M., Kan, K. J., & Jolles, J. (2016). A developmental perspective on spatial reasoning: Dissociating object transformation from viewer transformation ability. *Cognitive Development*, *38*, 63–74. https://doi.org/10.1016/j.cogdev.2016.01.004

Vasilyeva, M., & Huttenlocher, J. (2004). Early development of scaling ability. *Developmental Psychology*, *40*, 682–690. https://doi.org/10.1037/0012-1649.40.5.682

Wiedenbauer, G., & Jansen-Osmann, P. (2008). Manual training of mental rotation in children. *Learning and Instruction*, *18*(1), 30–41. https://doi.org/10.1016/j.learninstruc.2006.09.009

Witkin, H. A., Otman, P. K., Raskin, E., & Karp, S. (1971). *A manual for the embedded figures test*. Palo Alto, CA: Consulting Psychologists Press.

Wraga, M., Shephard, J. M., Church, J. A., Inati, S., & Kosslyn, S. M. (2005). Imagined rotations of self versus objects: An fMRI study. *Neuropsychologia*, *43*, 1351–1361. https://doi.org/10.1016/j.neuropsychologia.2004.11.028

Zacks, J. M., Vettel, J. M., & Michelon, P. (2003). Imagined viewer and object rotations dissociated with event-related fMRI. *Journal of Cognitive Neuroscience*, *15*, 1002–1018. https://doi.org/10.1162/089892903770007399

Received 18 September 2020; revised version received 7 February 2021

Supporting Information

The following supporting information may be found in the online edition of the article:

Table S1. Descriptive statistics for each spatial task by age group (raw scores).
Table S2. Descriptive statistics for each spatial skill by age group (*z*-scores).