Learning and Instruction 75 (2021) 101515

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

Developmental trajectories of children's spatial skills: Influencing variables and associations with later mathematical thinking

Wenke Möhring [a,*], Andrew D. Ribner [b], Robin Segerer [a], Melissa E. Libertus [b,c], Tobias Kahl [a], Larissa Maria Troesch [a,d], Alexander Grob [a]

[a] University of Basel, Department of Psychology, Switzerland
[b] University of Pittsburgh, Learning Research and Development Center, USA
[c] University of Pittsburgh, Department of Psychology, USA
[d] Institute for Research, Development and Evaluation, Bern University of Teacher Education, Switzerland

ARTICLE INFO

Keywords:
Spatial skills
Mathematical thinking
Growth
Developmental trajectory
Children
Cognitive development

ABSTRACT

Few studies have examined the long-term relations between children's early spatial skills and their later mathematical abilities. In the current study, we investigated children's developmental trajectories of spatial skills across four waves from age 3–7 years and their association with children's later mathematical understanding. We assessed children's development in a large, heterogeneous sample of children ($N = 586$) from diverse cultural backgrounds and mostly low-income homes. Spatial and mathematical skills were measured using standardized assessments. Children's starting points and rate of growth in spatial skills were investigated using latent growth curve models. We explored the influence of various covariates on spatial skill development and found that socioeconomic status, language skills, and sex, but not migration background predicted children's spatial development. Furthermore, our findings showed that children's initial spatial skills—but not their rate of growth—predicted later mathematical understanding, indicating that early spatial reasoning may play a crucial role for learning mathematics.

1. Introduction

Early mathematical understanding is predictive of success in later school years and post-secondary education (Davis-Kean et al., 2021; Duncan et al., 2007; Jordan et al., 2009; Sadler & Tai, 2007), and seems to be a building block for careers in science, technology, engineering, and mathematics (STEM, Hinojosa et al., 2016). Unfortunately, however, a large number of children do not meet the standards in mathematics education. For example, in a recent nationally representative assessment with 22,423 Swiss children from 11th grade (Konsortium & Hrsg, 2019), only 62% of the sample met the basic standards in mathematics, with large regional differences ranging from 44 to 83%. This finding is corroborated by similar patterns found in the U.S. (National Assessment of Educational Progress, 2019) and other countries in Western Europe. Given this large variability in mathematical understanding and the strong relation between early and later mathematical achievement (Duncan et al., 2007), it is crucial to increase our understanding of individual variables that allow successful mathematical learning during the early school years. One candidate skill that may play an important role for children's mathematical learning are spatial skills which are defined as children's abilities to think about "the location of objects, their shapes, their relations to each other, and the paths they take as they move" (Newcombe, 2010, p. 30).

A growing literature has described relations between children's spatial and mathematical skills (Hawes & Ansari, 2020; Mix & Cheng, 2012; Newcombe et al., 2018; Xie et al., 2020). The majority of these studies measured children's spatial abilities at one time point and assessed relations with concurrent or later mathematical ability. Though this is a valid design which has yielded valuable findings, this approach gives only a static snapshot of children's development, precluding the possibility to understand the importance of the developmental *path*. Such a trajectory may differ drastically among children, and individual differences therein may have implications for developing mathematical skills. For example, children with high initial levels of spatial skills may increase their spatial abilities to a larger extent than children with lower starting points, thus building on their initial, more sophisticated spatial skills. Several studies investigating growth trajectories of math skills found such amplifying, cumulative patterns (i.e., a "Matthew effect";

* Corresponding author. University of Basel, Department of Psychology, Missionsstrasse 62, 4055, Basel, Switzerland.
E-mail address: wenke.moehring@unibas.ch (W. Möhring).

https://doi.org/10.1016/j.learninstruc.2021.101515
Received 6 November 2020; Received in revised form 6 June 2021; Accepted 19 June 2021
Available online 25 June 2021

Aunola et al., 2004; Hong & You, 2012; Jordan et al., 2006). Given extant findings that spatial skills underlie mathematical skill development, such an amplifying pattern in spatial skills may have a positive impact on later mathematical skills. Alternatively, it is possible that children with lower starting points in spatial skills may develop spatial skills at a faster rate and catch up with their peers having higher starting points. Such steeper learning curves may have a long-term advantage for later mathematical skills irrespective of children's starting points. In the present study, we assessed children's developmental trajectories of early spatial skills. We were interested in explaining interindividual differences in both the initial spatial skill level and children's rate of growth using individual and family factors. Importantly, we examined associations of this spatial development with later mathematical abilities.

To date, the minimal research on this topic is contradictory and limited by the use of relatively small, homogenous samples and short time spans that restrict our understanding of the long-term relations between children's spatial and math abilities. Zhang and Lin (2017) found that in Chinese preschoolers, the starting point in spatial perception and growth therein across a 2-year-period ($N = 106$) predicted mathematical performance at the end of preschool. In contrast, Carr et al. (2017) investigated spatial development across a 2-year-period when American children were in second through fourth grade ($N = 304$). The authors showed that children's starting points in spatial skills—but not their rate of growth—predicted mathematical abilities. These different findings may reflect that gains in spatial development vary in their importance on mathematical ability at different phases of children's life, and that developmental trajectories early in children's emerging academic skills might be particularly consequential.

In the current study, we leveraged data from a cohort-sequential investigation of the development of children's language, cognitive, motor, and socio-emotional abilities from nearly 600 children age 3–7 years in Switzerland. Children were tested with four assessment sessions approximately 15 months apart. Thus, in contrast to previous research (Carr et al., 2017; Zhang & Lin, 2017), the current study consists of a larger sample covering a broad age range. Additionally, the majority of children in our sample had a migration background providing a unique opportunity to understand how children with diverse cultural backgrounds develop critical academic skills. While global migration trends of the world population have only increased from 2.9 to 3.4% from 1990 to 2017, some countries experienced more immigration than others. For example, Switzerland saw an increase of immigration by 42% in the same time period (United Nations, 2019). Using this sample, we investigated a) the development of spatial skills and characteristics that predict children's initial spatial skill level at the first assessment time point (i.e., their intercepts) and growth therein, b) whether children's intercept and growth in spatial skills were related to mathematical abilities, and c) identified different groups of growth trajectories and their relations to mathematical abilities.

First, we assessed the developmental path of children's spatial skills and the influence of various covariates. Based on previous research showing the influence of verbal ability on spatial and mathematical reasoning (cf. Purpura et al., 2017; Zhang et al., 2014), we expected children with higher German language skills to show higher starting points and steeper developmental trajectories in spatial skills. Furthermore, given that socioeconomic status (SES) and sex affect children's spatial skills (Levine et al., 2016; but see, 2005; Lachance & Mazzocco, 2006), we expected boys and children from high-SES backgrounds to show higher starting points and steeper developmental trajectories. Finally, given the characteristics of our sample, we were interested whether immigration was related to growth in spatial and math skills over time. There are reasons to believe that immigration may add to explained variance in children's development of spatial as well as mathematical skills. Immigrant children are less likely to be learning the language in which schooling occurs (in our study: German) as a home language. However, when being in a German-speaking, educational setting they may develop language skills more quickly than their peers, allowing them to catch up over time (Becker et al., 2013). Even though such effects of language might be reflected in our language skill variable, children from families with migration background may face unique challenges in their new residing country because they are less familiar with the educational system (Verhoeven, 2011). Additionally, parents may have different views on school success and might be differently involved in their child's education (Antony-Newman, 2019; Garcia Coll et al., 2002). These reasons might suggest that immigration adds to explained variance in children's spatial and mathematical skills beyond effects of language (and SES).

Second, we examined whether growth in children's spatial skills across four assessment waves was related to growth in math abilities between the third and fourth waves, even after accounting for a number of control variables. Importantly, in contrast to previous studies (Carr et al., 2017; Zhang & Lin, 2017), we controlled for non-spatial cognitive ability (i.e., abstract-reasoning skills) at each of the four waves allowing us to draw conclusions specifically to *spatial* skill development. We chose children's abstract-reasoning skills as a control variable as they are closely related to fluid intelligence and executive functioning (Diamond, 2013), which in turn are involved in every spatial task performance and strongly associated with mathematical skills (Cragg & Gilmore, 2014). Given that spatial and mathematical skills are correlated with one another (Frick, 2019; Lauer & Lourenco, 2016; Verdine et al., 2014), and that spatial skills seem to (at least in part) underlie the development of mathematical skills as evidenced by transfer effects from spatial skill training (Gilligan et al., 2019), we anticipated children's rate of growth in spatial skills to predict growth in math abilities.

Third, to better understand inter- and intra-individual characteristics related to changes in spatial skills and their association with math over time, we augmented findings by using a person-oriented approach to explore whether different kinds of growth trajectories in spatial skills could be identified. A recent approach identified two developmental profiles in children's spatial development (i.e., children with high vs. low spatial skills, Carr et al., 2017). The likelihood of showing a particular developmental profile was explained by children's sex, SES, and verbal working memory. Importantly, class-membership predicted children's mathematical achievement, with children with high spatial skills showing higher mathematics competency as compared to children with low spatial skills. Building on this latter finding, we tested whether having a more advantageous pattern of growth (i.e., higher starting point and/or faster rate of growth) relates to greater growth in math abilities, and expected associations between class-membership and children's mathematical performance.

2. Methods

2.1. Participants

Data from the research project Zweitsprache (english translation: Second Language) was used. This project started in 2009 at the University of Basel with the goal to track children's second language trajectories across early and middle childhood. Consequently, the majority of children came from a diverse set of backgrounds and represent more than 60 countries of origin. The current study was approved by the local Ethics Committee. Parents gave written informed consent and children agreed verbally to participate. Children received a toy and parents received feedback about their child's test results after each wave.

At the first wave, the sample consisted of 586 children (50.2% female) with the majority of children speaking at least two different languages (85.2%) and not being born in Switzerland (63.8%). Mothers of children with a migration background had lived in Switzerland on average for 8.94 years ($SD = 6.65$); fathers had lived in Switzerland for 11.69 years ($SD = 8.27$). A small number of children were monolingual and spoke German as their native language ($n = 86$). The entire sample consisted of four consecutive birth cohorts: born 2005/2006 ($n = 84$); 2006/2007 ($n = 156$); 2007/2008 ($n = 186$); and 2008/2009 ($n = 160$).

Each cohort was assessed at four waves, beginning approximately 15 months prior to kindergarten entry (in Switzerland, children attend mandatory kindergarten for two years beginning at the age of four years). This first assessment took place when children were approximately 3.5 years old (M_{age} = 42.05 months, SD = 4.17). The next waves were at the start of kindergarten (M_{age} = 57.69 months, SD = 3.78), end of kindergarten (M_{age} = 74.23 months, SD = 3.84), and at the end of the first year in primary school (M_{age} = 87.48 months, SD = 3.80). Despite every effort to keep attrition to a minimum, the sample decreased over time from Wave 1 = 586, to Wave 2 = 429, to Wave 3 = 375, to Wave 4 = 325. Full Information Maximum Likelihood (FIML) was used to account for missing data, as FIML is more reliable than other methods including listwise deletion and mean imputation for dealing with data missing at random in structural equation modeling frameworks (Cham et al., 2017; Enders, 2001). Results of two prior Monte Carlo simulation studies suggest that our analyses are sufficiently powered to detect even small effects on linear and quadratic growth parameters in a latent growth context (Diallo et al., 2014; Fan & Fan, 2005). They found that a sample of at least 250 participants can adequately detect a small effect on linear (Fan & Fan, 2005) and quadratic growth (Diallo et al., 2014) with power of .80 given four measurement time points, as is the case for the current analyses.

Information about each indicator of socioeconomic status (SES)—described in greater detail below—were available for 74–82% of all families. Less than 2% of parents did not finish school; 17% of mothers and 15% of fathers had finished mandatory school as their highest educational degree; 17% of mothers and 19% of fathers had completed a vocational training; 13% of mothers and 11% of fathers had completed an academic high-school-level qualification; 23% of mothers and 33% of fathers had a college/university degree. Families had a median yearly income of approximately 66′000 Swiss Francs (SD = 27′300) which was below the average yearly income of 112′400 Swiss francs for Swiss families with children at the time (Swiss Federal Statistical Office, 2012).

2.2. Measures

2.2.1. Spatial skills

Children's spatial skills were assessed at each of the four waves using the subtest "Mosaics" from the Snijders-Oomen Nonverbal Intelligence Test 2.5–7 Revised (SON-R 2.5–7; Tellegen et al., 2007). This measure showed acceptable reliability (e.g., split-half reliability indices ranging from 0.78 to 0.93 for 3- to 6-year-old children; Renner et al., 2009) and concurrent validity in previous studies (Jenkinson et al., 1996; Moore et al., 1998). Children received colored squares and a frame and were asked to copy a pattern in this frame. Therefore, spatial skills were operationalized by using children's assembly skills in line with previous studies (Jirout & Newcombe, 2015; Kahl et al., 2019; Kyttäla et al., 2003). Children were presented with 15 items of increasing difficulty. The experimenter solved the first three items in her own frame and showed each step to the child using gestures and explained the task verbally. The next three items were untimed and children were asked to copy simple patterns using three-to-five red squares. Beginning with the 7th item, there was a time limit of 2.5 min to copy each pattern and children were presented with red, yellow, and red/yellow squares and asked to copy a set of patterns. Children's assembled configurations were scored correctly if the child was able to put the squares at the correct place on her own (e.g., twisted, mirrored solutions counted as incorrect). The subtest was stopped when participants answered two consecutive items incorrectly or produced a total of three incorrect configurations. The number of correctly solved items served as dependent variable.

2.2.2. Mathematical achievement

The subtest "Logical-mathematical thinking" from the Intelligence and Development Scales (IDS, Grob et al., 2009) was used to measure mathematical thinking at Wave 3 and 4 and thus, when children were at the end of kindergarten and in 1st grade. This test has shown high differential and concurrent validity (Hagmann-von Arx et al., 2008). The subtest assessed a wide range of mathematical skills such as counting, an understanding of ordinality and magnitudes, knowledge about invariance, mental addition, and proportional reasoning. Children solved a maximum of 18 items. Items were untimed except for the last four items that had a time limit (90 s). The task was stopped when participants answered three consecutive items incorrectly. The number of correctly solved items served as dependent variable.

2.2.3. Covariates

A series of covariates were included in all analyses to ensure variance was due to the hypothesized predictors and not a result of general cognitive functioning, general maturation, or characteristics not central to our research questions. All models controlled for non-spatial cognitive skills, sex, age at time of assessment, German language skills at Wave 1, migration background, and SES. In case of non-spatial cognitive skills and German language skills, data were obtained from direct assessment; SES, sex, migration background, and age were obtained via the Ministry of Education or a parent questionnaire.

2.2.3.1. Non-spatial cognitive skills. Children's non-spatial cognitive skills were assessed at each of the four waves using the subtest "Categories" from the Snijders-Oomen Nonverbal Intelligence Test 2.5–7 Revised (SON-R 2.5–7; Tellegen et al., 2007), which measures children's abstract-reasoning skills. This measure showed acceptable reliability (e.g., split-half reliability indices ranging from 0.76 to 0.89 for 3- to 6-year-old children; Renner et al., 2009) and concurrent validity in previous studies (Jenkinson et al., 1996; Moore et al., 1998). Children were presented with 15 items of increasing difficulty. For the first seven items, children were presented with four or six cards showing pictures of objects (e.g., different dolls and teddy bears) and asked to sort these cards into two categories. The experimenter helped with the first two cards within the first five items and explained the task verbally and by using gestures. Items 6 and 7 were solved without help by the experimenter. Beginning with the 8th item, children were presented with three cards showing objects that have something in common (e.g., pictures showing dogs). Children were asked to choose two pictures from a set of five alternatives that share the same commonality. Children's answers were scored as correct if the child ordered each card correctly. The subtest was stopped when participants produced a total of three incorrect answers. The number of correct answers served as dependent variable.

2.2.3.2. Language skills. German language competence was assessed using the SETK-2 (Sprachentwicklungstest, Grimm, 2000), which measures receptive and expressive language skills and is normed for 2-year-old children. Given that most children with migration background had just begun to learn German, this test for younger children was chosen to avoid floor effects. However, even children without a migration background showed considerable variation in performance. Children solved four subtests: producing words, understanding words, producing sentences, and understanding sentences. Receptive language skills were tested by presenting them with a set of four pictures and asking them to choose the picture corresponding to an orally presented word or sentence. Expressive language skills were assessed by asking children to label or describe pictures of objects or activities. Grimm (2000) reported acceptable-to-high reliabilities for each of those subtests (Cronbach's αs ranging from 0.56 to 0.95). A latent factor score was created from the four subtests. The four variables contributed nearly equal variance to the resulting factor ($\lambda_{\text{Understanding words}}$ = 0.899; $\lambda_{\text{Understanding sentences}}$ = 0.904; $\lambda_{\text{Producing words}}$ = .880; $\lambda_{\text{Producing sentences}}$ = .796). The estimated factor fit the data well $\chi^2(1) = 3.542$, p = .0599, RMSEA < 0.001 90% CI [0.000,.146], CFI = 0.998, and was adequately reliable, α = 0.868.

2.2.3.3. Socioeconomic status. A latent factor score was created from three indicators of SES: Educational attainment of both parents and family household income which were assessed via parent questionnaires. Educational attainment was measured using the European categories for countries with a dual educational system (1 = no school education, 2 = compulsory school, 3 = vocational training, 4 = high school, 5 = college or university). Parents were also asked to rate their monthly family income using the following categories: 1 = less than 1′000 Swiss Francs, 2 = 1′001–2′000, 3 = 2′001–3′000, 4 = 3′001–4′000, 5 = 4′001–5′000, 6 = 5′001–6′000, 7 = 6′001–7′000, 8 = 7′001–8′000, 9 = 8′001–9′000, 10 = 9′001–10′000, 11 = more than 10′000 Swiss Francs. All variables contributed nearly equal variance to the resulting factor ($ƛ_{\text{Maternal Education}} = 0.736$; $ƛ_{\text{Paternal Education}} = .824$; $ƛ_{\text{Income}} = 0.693$). Given the latent factor was composed of three variables, fit indices were unavailable and thus traditional reliability statistics are a better indicator of model fit; the factor was adequately reliable, α = 0.749.

2.3. Procedure

Children were recruited by the help of the Ministry of Education of the Canton of Basel-Stadt. Eighteen months before kindergarten start, each family in town with a pre-kindergarten child was sent a questionnaire with the goal to assess their language skills. Together with this questionnaire, families received an invitation to participate in the present study. Families who agreed were contacted to set up an appointment.

Testing took place mainly at children's homes with some testing sessions taking place at the University. Testing sessions were led by trained research assistants and took approximately 1.5 h. Each assessment session began with a 10-min play session between the experimenter and the child to familiarize the child with the test situation. Immediately afterwards, the experimenter tested each child's German language skills, their spatial and non-spatial cognitive abilities, and mathematical achievement among other tasks that are not included in the present research project. A description of these other measures can be found elsewhere (e.g., Grob et al., 2014; Troesch et al., 2021). Given that children completed a large test battery, counterbalancing was not feasible and children were tested in the above-mentioned order. If parents' and children's German language skills were not sufficient to understand the instructions, experimenters were accompanied by an interpreter from the Ministry of Education who explained the procedure, instructions, and helped to complete the questionnaire.

2.4. Analysis plan

To address our first research question regarding the predictors of the development of spatial skills, we conducted a series of latent growth curve models. We first estimated unconditional latent growth mixture models to assess the average growth pattern of spatial skills across the sample. To determine the number of latent growth factors that best fit the data, we compared model fit of a growth model with intercept and fixed linear slope parameters to models of increasing complexity (e.g., a model with free linear slope parameters). Model comparison was assessed using the Satorra-Bentler Scaled Chi-Square Difference Test. We used the baseline cutoff criteria outlined by Hu and Bentler (1999): We expected a well-fitting model to have a Comparative Fit Index (CFI) of 0.95, and Root Mean Squared Error of Approximation (RMSEA) of 0.08. For each of the four waves, the linear slope time points were set with equal distance between them such that each of the four time points had a value of 0, 1, 2, and 3 respectively. In order to enable conclusions about children's developmental trajectories in *spatial* skills, we controlled for non-spatial cognitive skills at each time point. To this end, each time point of spatial skills was regressed on non-spatial cognitive skills from the same time point as a time-varying covariate (i.e., Wave 1 spatial skills was regressed on Wave 1 non-spatial cognitive skills; Wave 2 spatial skills were regressed on Wave 2 non-spatial cognitive skills etc.). To account for general maturation effects, each time point of spatial skills was additionally regressed on age at time of assessment. Growth parameters (intercept and slope parameters) were then regressed on the predictor variables (e.g., language skills, SES).

To address our second research question as to whether growth in spatial skills predicted gain in math abilities across one year, growth curves from research question 1 were retained. A latent change score representing growth in math skills from Wave 3 to Wave 4 was estimated. That latent change score was then regressed on growth parameters and demographic covariates. Growth parameters and covariates were allowed to correlate.

Finally, to address our third research question as to the common heterogeneous patterns of growth in spatial skills, we conducted latent growth mixture models. Latent growth mixture models estimate growth curves within each class and capture individual variation around these growth curves by estimating the growth parameter variances within each class (Muthén & Muthén, 2000). As a baseline model, we retained unconditional growth curves from research question 1. We ascertained the appropriate number of classes through a process of class enumeration without covariates and with freely estimated growth parameters. Models with different numbers of classes were compared using a series of information criteria and likelihood-based tests. Vuong-Lo-Mendell-Rubin and Lo-Mendell-Rubin adjusted likelihood ratio tests for *N* versus *N–1* classes were used to decide the appropriate number of classes (Nylund et al., 2007; Tofighi & Enders, 2008; Wang & Bodner, 2007). We retained the model for which likelihood ratio tests expressed no significant difference from *N–1* classes. After ascertaining the number of classes, we tested predictors of membership in differing growth trajectories (e.g., whether language skills would predict membership in one particular class). Finally, we tested membership in growth trajectories as a predictor of change in math skills over and above covariates. To this end, we used the three-step (or three-step ML) method (Asparouhov & Muthén, 2014; Vermunt, 2010). The three-step method—in contrast with the one-step method, the pseudo-class method—first estimates the latent classes independent of auxiliary variables and covariates. Then, this method uses the posterior distribution to estimate a most likely class membership variable, and finally regresses this variable on predictor variables (Asparouhov & Muthén, 2014). Because the three-step method does not allow for estimation of a latent predictor or distal outcome, estimated latent factors scores were exported and subsequently used as observed variables for models. All models were estimated using Mplus 8 (Muthén & Muthén, 2017).

3. Results

3.1. Descriptive statistics

Descriptive statistics and correlations among all variables included in the analyses are presented in Table 1. Spatial and non-spatial cognitive skills were relatively stable across administrations ($rs = 0.41–0.59$, $ps < .001$). Spatial and non-spatial cognitive skills were also moderately correlated with one another at each wave ($rs = 0.40–0.59$, $ps < .001$). Spatial and non-spatial cognitive skills at all waves were weakly to moderately correlated with math skills at Wave 3 ($rs = 0.18–0.48$, $ps < .01$) and math skills at Wave 4 ($rs = 0.25–0.49$, $ps < .001$).

3.2. Which variables predict growth in spatial skills?

3.2.1. Unconditional latent growth curve models

Unconditional latent growth models were estimated for spatial skills to examine the average pattern of growth for all children. Raw scores were used to estimate within-person growth. Four competing models for spatial skills with increasing complexity were tested. The first model had a fixed linear term wherein there was no variance in linear growth across the sample (i.e., all participants developed skills at the same rate); the

Table 1
Descriptive statistics and correlations of all study variables.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	Age Months W1	–																	
2	Age Months W2	.91***	–																
3	Age Months W3	.89***	.96***	–															
4	Age Months W4	.90***	.95***	.94***	–														
5	Spatial Skills W1	.43***	.40***	.36***	.42***	–													
6	Spatial Skills W2	.24***	.27***	.30***	.28***	.53***	–												
7	Spatial Skills W3	.17**	.16**	.17**	.20***	.48***	.58***	–											
8	Spatial Skills W4	.15**	.18**	.17**	.18**	.42***	.54***	.59***	–										
9	Non-Spatial Skills W1	.38***	.37***	.38***	.36***	.59***	.50***	.43***	.41***	–									
10	Non-Spatial Skills W2	.18***	.22***	.22***	.19**	.43***	.55***	.44***	.36***	.47***	–								
11	Non-Spatial Skills W3	.14**	.18**	.17**	.18**	.31***	.40***	.40***	.36***	.35***	.46***	–							
12	Non-Spatial Skills W4	.12*	.14*	.17**	.18**	.29***	.36***	.30***	.41***	.31***	.35***	.41***	–						
13	Mathematics W3	.10*	.15**	.14**	.16**	.37***	.40***	.48***	.44***	.38***	.38***	.28***	.18**	–					
14	Mathematics W4	.14**	.21***	.19**	.19**	.37***	.49***	.48***	.47***	.40***	.38***	.31***	.25***	.64***	–				
15	SES Factor	-.06	-.08	-.05	-.03	.35***	.45***	.49***	.50***	.34***	.38***	.26***	.26***	.51***	.55***	–			
16	Language Factor	.17***	.16**	.16**	.19**	.32***	.43***	.38***	.31***	.37***	.37***	.27***	.21***	.48***	.46***	.50***	–		
17	Migration Background	.11**	.06	.07	.01	-.03	-.09	-.21***	-.10	-.07	-.09	-.09	-.04	-.25***	-.24***	-.19***	-.39***	–	
18	Child Female	.00	.01	-.03	-.03	.10*	.02	-.04	-.10	.10*	.08	.01	.07	-.13*	-.24***	-.05	.02	-.03	–
	N	586	429	375	325	574	426	375	325	563	421	366	325	374	324	586	586	586	586
	Mean (*SD*)	42.05 (4.17)	57.69 (3.78)	74.23 (3.84)	87.48 (3.80)	4.25 (2.67)	8.73 (2.05)	11.71 (2.09)	13.16 (1.82)	5.06 (2.43)	8.77 (2.52)	11.41 (2.09)	12.84 (1.56)	6.15 (1.85)	8.24 (2.26)	0.00 (1.10)	0.00 (2.62)	64%;	50%;
	Range	34–52	50–65	60–82	79–95	0–12	0–14	6–15	9–15	0–13	0–15	0–15	7–15	1–13	3–15	−2.34–2.14	−2.93–5.32	*n* = 374	*n* = 294

Note: *$p < .05$; **$p < .01$; ***$p < .001$; W1—Wave 1, W2—Wave 2, W3—Wave 3, W4—Wave 4; *SD*—Standard Deviation. Migration background: 0 = no; 1 = yes.

second model involved a free linear term (i.e., all participants developed spatial skills, but the rate of growth was allowed to vary between participants); the third model had a fixed quadratic term, and the last model involved a free quadratic term. Fit statistics for each estimated model are presented in Table 2.

Given the four estimated growth models for spatial skills, the model with the fixed quadratic term fit the data best while being the most parsimonious. The model with the free linear term fit better than the model with the fixed linear term (Satorra-Bentler Chi-Squared Difference χ^2 (2) = 22.883, $p < .0001$). Building upon that, the model with the fixed quadratic term fit better than the free linear model (Satorra-Bentler Chi-Squared Difference χ^2 (1) = 220.169, $p < .0001$), and a model with a free quadratic term fit no better than the model with the fixed quadratic term (Satorra-Bentler Chi-Squared Difference χ^2 (3) = 0.5805, $p = .9009$). The retained model with the fixed quadratic term fit the data well, χ^2 (4) = 1.008, $p = .9086$, RMSEA < 0.001 90% CI [0.000,.025], CFI = 1.00. For the resulting model, means were significant for all parameters (Intercept: $\mu = 4.21, p < .001$; Linear Slope: $\mu = 5.20, p < .001$; Quadratic Slope: $\mu = -0.76, p < .001$); variances were significant for the freely estimated parameters (Intercept: $\sigma = 3.39, p < .001$; Linear Slope: $\sigma = 0.16, p < .001$); the variance around the Quadratic term was fixed at zero.

These findings show that, on average, children at the first time point scored approximately 4 points on the test of spatial skills and developed skills at an average of 5 points per assessment period over time. However, the extent to which scores increased slowed over time (as indicated by the negative quadratic slope). Therefore, there was less of an increase between assessment time points 3 and 4 than there was between assessment time points 2 and 3. Importantly, there were substantial individual differences in the starting point of spatial skills (i.e., the intercept), and some individual differences in the rate of growth as evidenced by significant variance around the mean of the intercept and linear slope terms. Children's starting point was correlated with linear slope ($r = -0.59, p < .001$), such that children who started with higher levels of spatial skills showed a slower growth rate than those who started with lower levels of spatial skills.

3.2.2. *Conditional latent growth curve models*

To examine predictors of growth for spatial skills, we estimated conditional latent growth models. The retained model parameters from model comparisons among unconditional models above were regressed on child and family characteristics. Spatial skills at each time point were regressed on non-spatial cognitive skills from the same time point and age at time of assessment as time-varying covariates. Growth parameters were then regressed on predictor variables; the latent slope term(s) were regressed on the intercept term.

Intercept and linear slope were regressed on the SES factor, German language skills, sex, and an indicator variable for whether or not the child had a migration background. Results are shown in Models 1 and 2 of Table 3. These models indicate that children's German language skills and SES were related to the intercept such that children with higher language skills and children who came from higher-SES homes began with higher levels of spatial skills ($\beta = 0.20, p = .023$; $\beta = 0.47, p < .001$, respectively). Neither migration background nor sex were related to the intercept; however, there was a tendency for the relation between sex and linear slope such that boys developed spatial skills at a faster rate than girls ($\beta = -0.31, p = .058$). Therefore, it seems that language skills and household SES relate to very early spatial skills, and sex relates to the rate of spatial skill development over time.

Table 2
Model fit for tested models of spatial skill growth.

	χ^2 Test of Model Fit	RMSEA	CFI	Satorra-Bentler Scaled Chi-Square Difference p-value
Model 1: Fixed Linear Growth	χ^2 (7) = 370.182, $p <$ .0001	.299 90%CI [.273,.325]	0.172	N/A
Model 2: Free Linear Growth	χ^2 (5) = 328.279, $p <$ .0001	.333 90%CI [.303,.363]	0.263	<.0001
Model 3: Fixed Quadratic Growth	χ^2 (4) = 1.008, $p =$ .9086	<.001 90%CI [.000,.025]	1.000	<.0001
Model 4: Free Quadratic Growth	χ^2 (1) = 0.410, $p =$ .5221	<.001 90%CI [.000,.094]	1.000	0.9009

3.3. *Does growth in spatial skills predict math?*

Using the predictors from above (i.e., SES factor, German language skills, migration background, sex) and growth parameters of spatial skills from Waves 1 to 4, a model predicting development of math from Wave 3 to Wave 4 was estimated. Again, slope was regressed on intercept and both intercept and slope were regressed on predictors and covariates. Latent change in math was regressed on all covariates as well as linear slope and intercept parameters. Results are shown in Model 3 of Table 3. The intercept predicted math ($\beta = 0.29$, $p = .001$), but linear slope did not ($\beta = 0.02$, $p = .911$). None of the other variables were related to the development of math skills. Thus, this result suggests that only children's initial spatial skills (i.e., their starting points) were related to their later mathematical abilities whereas their rate of growth in spatial skills did not.

3.4. *Do differentiable growth trajectories predict math skills?*

As an alternative approach to latent growth curve models which describe a normative growth trajectory for all children, we also estimated latent growth mixture models—a person-centered approach which allows for the estimation of different growth trajectories for different groups of children. The same growth parameters as previously established for the normative growth trajectory were used (i.e., intercept, linear slope, and fixed quadratic term).

A two-class model emerged as the best fitting model—it fit better than did a one-class model, Lo-Mendell-Rubin Adjusted LRT = 127.168, $p = .0005$, and a three-class model did not fit any better, Lo-Mendell-Rubin Adjusted LRT = 51.965, $p = .5528$. A visual inspection of classes revealed Class 1 was characterized by a low starting point (Low Spatial Skills). Class 2 was characterized by a relatively higher starting point, but similar slope (High Spatial Skills). Membership in the two classes was relatively evenly split: $n = 318$ (54.2%) were in the Low Spatial Skills class, and $n = 268$ (45.7%) were in the High Spatial Skills class (cf. Fig. 1).

To ascertain predictors of class membership, the three-step approach was used with all predictors and covariates as auxiliary variables (Asparouhov & Muthén, 2014). Results are presented in Table 4. Students in the Low Spatial Skills class were more likely to be from families of lower SES (Odds Ratio [OR] = 0.29, $p < .001$) and scored lower on assessment of German language skills (OR = 0.80, $p = .002$). Class membership was then tested as a predictor of latent change scores in math. Over and above the same covariates used above (i.e., SES, migration background, age at Wave 1, child sex, and German language skills), those in the High Spatial Skills class had a greater change in mathematical abilities from Wave 3 to Wave 4 than those in the Low Spatial Skills class ($M_{\text{HighSpatial}} = 2.15$, $SE = 0.09$; $M_{\text{LowSpatial}} = 1.50$, $SE = 0.06$; χ^2 (1) = 31.49, $p < .001$).

4. Discussion

Using a large-scale, longitudinal, multisource investigation across four time points, we investigated the development of children's spatial skills from age 3–7 years and its relation to the developmental trajectory of children's math abilities. Our findings demonstrate that SES, German

Table 3
Linear Regressions Predicting Growth Parameters of Spatial Skills (Intercept and Slope) as well as Math Latent Change from Wave 3 to Wave 4. Significant Results are Presented in Bold.

	Intercept of Spatial Skills			Slope of Spatial Skills			Math Latent Change W3 to W4		
	Beta	*SE*	*p*	Beta	*SE*	*p*	Beta	*SE*	*p*
Family Socioeconomic Status	**0.47**	**0.09**	**0.000**	0.29	0.20	0.140	−0.09	0.13	0.478
Child German Language Skills	**0.20**	**0.09**	**0.023**	−0.02	0.16	0.909	−0.12	0.12	0.328
Child With Migration Background	0.06	0.07	0.374	−0.09	0.13	0.467	−0.05	0.07	0.437
Child Female	0.10	0.06	0.093	*−0.31*	*0.16*	*0.058*	−0.09	0.09	0.328
Intercept of Spatial Skills				−0.32	0.28	0.261	**0.29**	**0.09**	**0.001**
Slope of Spatial Skills							0.02	0.15	0.911

Note. Non-spatial cognitive skill and age at time of assessment were entered as time-varying covariates into the same models, thereby controlling for those characteristics.

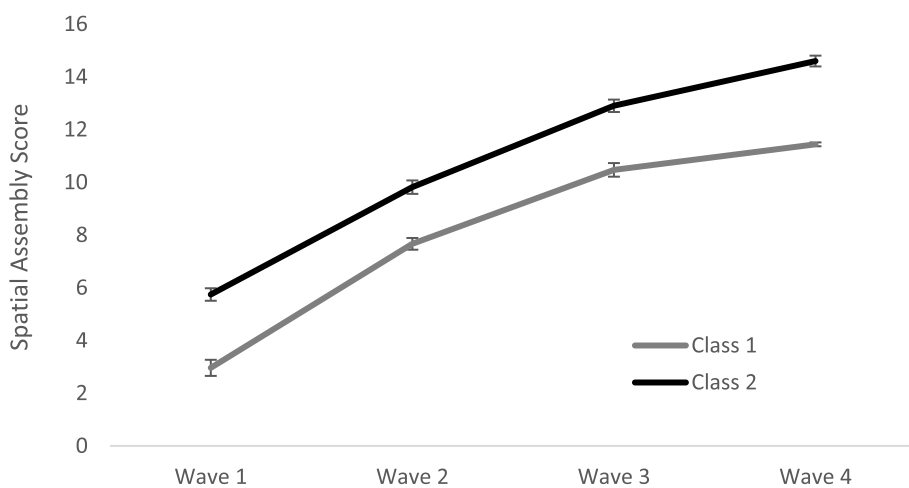

Fig. 1. Estimated means for children with low spatial skills (class 1) and high spatial skills (class 2).

Table 4
Predictors of Class Membership: Low Spatial Skills Relative to High Spatial Skills. Significant Results are Presented in Bold.

	Estimate	*SE*	*p* value	Odds Ratio
Family Socioeconomic Status	**−1.24**	**0.19**	**< 0.001**	**0.28**
Child German Language Skills	**−0.22**	**0.07**	**0.002**	**0.80**
Child With Migration Background	0.15	0.36	0.675	1.16
Child Female	0.02	0.32	0.961	1.02

language skills, and sex predict the development of spatial skills in our sample. This result is in line with findings from several studies indicating that SES (Levine et al., 2005, 2016), language skills (Zhang & Lin, 2014), and sex explain large parts of interindividual differences in children's spatial skills (Frick et al., 2014; Lauer et al., 2019). Given our focus on children's developmental trajectories, however, the current study extended previous work by showing that SES and language skills were particularly related to children's *initial* spatial skills whereas sex predicted children's *gains* in developing spatial skills. It was found that male children tended to develop their spatial skills more quickly as compared to females across the 3-year-period. Even though the mechanisms that underlie this faster growth remain speculative with the present data, it may be that boys hear more spatial language from their parents during this phase (Casasola et al., 2020; Pruden & Levine, 2017), engage more in spatial play (Jirout & Newcombe, 2015), or have begun to endorse stereotypical thinking about gender differences in spatial skills as conveyed by parents' or teachers' comments or behavior (for evidence in the math domain, cf. Beilock et al., 2010).

When looking at relations between children's starting points and growth in spatial skills, we found that children with lower starting points developed their spatial skills at a faster pace than children with higher starting points. Whereas this may indicate some ceiling effects for children with higher spatial skills, a look at Fig. 1 shows that these children did not necessarily approach ceiling even in the last measurement time point. From a practical standpoint, this may suggest that the mandatory kindergarten attendance and the formalized curriculum that children are exposed to in kindergarten support children's development of spatial skills, with children showing a lower initial level benefitting more from this educational experience than children with higher initial levels. Swiss kindergartens use a formalized curriculum which includes tasks aimed at improving children's spatial reasoning. Children learn about different geometric shapes and improve their spatial language by learning the correct usage of relational terms such as "left", "right", "below", and "above". Additionally, they are presented with arrays of multiple objects in different perspectives and are asked to re-build such arrangements either at the time of presentation or from memory.

Our finding of an inverse relation between children's initial level and their rate of growth may be interpreted that children with lower initial spatial skills catch up with their peers with higher initial spatial skills. However, results from our latent growth mixture models suggest this conclusion might be tempered. Findings indicated that children with lower initial spatial skills did not reach the same level of spatial skills as their peers with higher initial spatial skills at the end of our study. These latent growth mixture models also helped identify heterogeneity in children's developmental trajectories. Our analyses indicated that a 2-class model best fit the data. These two latent classes differed in their starting points of spatial skills but showed a similar slope, which is in line with recent results (Carr et al., 2017). Analyses on influential variables for class-membership indicated that students from higher-SES backgrounds and those with higher German language skills were more likely in the class with higher spatial skills (cf. Carr et al., 2017; Levine et al., 2005, 2016; Zhang & Lin, 2016). Similar to potential mechanisms for sex differences outlined above, it may be that parents from higher SES backgrounds use more spatial language and promote more spatial play (Jirout & Newcombe, 2015; Pruden & Levine, 2017).

These findings contribute to the limited knowledge base about children's spatial development and the predictors thereof, particularly given the dearth of studies investigating children's developmental trajectories of spatial skills. However, our study also aimed at clarifying whether the rate of growth in spatial skills—over and above the initial spatial skill level—holds added value in explaining mathematical abilities. Our results suggested that this was not the case. Children's initial spatial ability—but not their rate of change in spatial skills—predicted later mathematical abilities. Whereas the first result lends support to previous research (Hawes & Ansari, 2020; Mix & Cheng, 2012; Newcombe et al., 2018; Xie et al., 2020), it also extends this research by demonstrating that the pace of how quickly children develop spatial skills is unrelated to their later mathematical understanding. This result was unexpected and contrasts previous findings (Zhang & Lin, 2017). However, our results point to important implications such as the direct and indirect influence of spatial, language skills, and SES for children's mathematical achievement. These findings may help in creating successful interventions and determining effective timing in intervention deployment. An overwhelming amount of very early childrearing—both in care settings and in the context of parenting—is focused on the development of language skills, with far less attention given to mathematical or spatial skills (Engel et al., 2013). The present findings may help refine parenting programs which aim at compensating children's difficulties in spatial tasks.

Our study has several strengths and limitations. We consider it a strength that we have used instruments that were suitable and sensitive for the present age range and allowed for non-verbal testing. Moreover, we have controlled for non-spatial cognitive abilities at each wave enabling specific conclusions to children's developmental trajectories in *spatial* skills and their relation to mathematical abilities. Finally, our sample included many children with a migration background and from low-income homes. Considering that these children are most likely at risk for spatial and mathematical difficulties (Casey et al., 2011; Levine et al., 2005), it seems crucial to understand the factors which improve or hinder their development and academic success.

Several limitations warrant mention. Despite the strengths that arise from repeated measurement, a robust longitudinal design, and conservative within-person growth analyses, no causality can be inferred from these findings. A further limitation concerns using the same measures across the four waves. This decision enabled us to model developmental trajectories and increased our confidence that developmental change did not exist because of changing the instruments. However, at the same time, it may be that gains in spatial skills also reflect some practice effects. It can also be seen as a limitation that spatial skills were measured using a spatial assembly task. This decision reflects the shortage of age-appropriate measures at this young age and accords to the operationalization in other studies (Jirout & Newcombe, 2015; Kahl et al., 2019; Kyttälä et al., 2003). This approach is also in line with previous studies that predominantly used mental rotation tasks as a spatial measure (e.g., Carr et al., 2017). Mental rotation as well as the present spatial assembly task can be seen as intrinsic, dynamic spatial skills when referring to a typology from Newcombe and Shipley (2015, see also Uttal et al., 2013).[1] However, it is possible that results may differ for other spatial skills and in particular for spatial skills that are static and involve reasoning about extrinsic characteristics (e.g., considering relations between various objects in a large-scale environment). A similar criticism may refer to our math measure which was a composite covering students' knowledge about various mathematical topics and consequently, did not enable conclusions about relations to specific mathematical topics. Another limitation was that immigrant children were not tested in their native language which was not feasible given the variety of cultural backgrounds of our participants. With respect to our spatial and non-spatial cognitive measures, we tried to address this potential confound by choosing non-verbal tests. Furthermore, experimenters were accompanied by an interpreter who helped translating verbal instructions in several tests. Still, we cannot assess in which ways the scores of immigrant children may additionally reflect the extent of being challenged by the language or feeling uncomfortable in the testing situations.

Critically, our findings may be important for future studies and intervention work. First, we find that individual differences in children's initial spatial skills rather than the rate of change predict development of math skills, which indicates the long-term, robust association between early spatial skills and later mathematical achievement which may imply the need for early intervention. Second, we note groups for whom intervention might be particularly important: Future experimental studies using spatial training may particularly focus on children from low-SES backgrounds, females, and children with lower language skills, and assess potential transfer effects to their mathematical abilities. Finally, we posit the importance of increasing awareness in parents and teachers about the role of spatial skills for children's academic success. As early spatial skills predict the development of later math skills, it seems critical that parents, caregivers, and teachers learn more about the simple-to-execute but effective tools of how to improve children's spatial skills such as spatial language, spatial play, or gesture (Newcombe, 2010).

Declarations of interest

None.

Author note

The present data have not been presented at a conference. Furthermore, the present research is neither available on a listserv nor on a website.

Author statement

Wenke Möhring: Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Andrew D. Ribner: Conceptualization, Methodology, Formal analysis, Writing – review & editing, Robin Segerer: Methodology, Writing – review & editing, Melissa E. Libertus: Conceptualization, Writing – review & editing, Tobias Kahl: Writing – review & editing, Larissa Maria Troesch: Investigation, Data curation, Writing – review & editing, Project administration, Alexander Grob: Resources, Writing – review & editing, Supervision, Funding acquisition

Acknowledgements

The current study was supported by grants awarded by the Jacobs Foundation Zurich and the Ministry of Education of the Canton of Basel-Stadt. The authors express their gratitude to the research assistants and participating families. The funding sources had no involvement in data collection, analysis, and interpretation of data, in writing, and in the decision to submit the article for publication.

References

Antony-Newman, M. (2019). Parental involvement of immigrant parents: A meta-synthesis. *Educational Review, 71*, 362–381.

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using M*plus*. *Structural Equation Modeling: A Multidisciplinary Journal, 21*, 329–341. https://doi.org/10.1080/10705511.2014.915181

[1] According to Uttal et al. (2013), intrinsic and dynamic spatial skills are defined as being able to "piece together objects into more complex configurations" (p. 355). Examples of measures refer to mental-rotation tests and the block design task (which is very similar to our spatial assembly task).

Aunola, K., Leskinen, E., Lerkkanen, M.-K., & Nurmi, J.-E. (2004). Developmental dynamics of math performance from preschool to grade 2. *Journal of Educational Psychology, 96*(4), 699–713. https://doi.org/10.1037/0022-0663.96.4.699

Becker, B., Klein, O., & Biedinger, N. (2013). The development of cognitive, language, and cultural skills from age 3 to 6: A comparison between children of Turkish origin and children of native-born German parents and the role of immigrant parents' acculturation to the receiving society. *American Educational Research Journal, 50*(3), 616–649.

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls, math achievement. *Proceedings of the National Academy of Sciences of the United States of America, 107*, 1860–1863. https://doi.org/10.1073/pnas.0910967107

Carr, M., Alexeev, N., Wang, L., Barned, N., Horan, E., & Reed, A. (2017). The development of spatial skills in elementary school students. *Child Development, 89*, 446–460. https://doi.org/10.1111/cdev.12753

Casasola, M., Wei, W. S., Suh, D. D., Donskoy, P., & Ransom, A. (2020). Children's exposure to spatial language promotes their spatial thinking. *Journal of Experimental Psychology: General, 149*, 1116–1136. https://doi.org/10.1037/xge0000699

Casey, B. M., Dearing, E., Vasilyeva, M., Ganley, C. M., & Tine, M. (2011). Spatial and numerical predictors of measurement performance: The moderating effects of community income and gender. *Journal of Educational Psychology, 103*, 296–311. https://doi.org/10.1037/a0022516

Cham, H., Reshetnyak, E., Rosenfeld, B., & Breitbart, W. (2017). Full information maximum likelihood estimation for latent variable interactions with incomplete indicators. *Multivariate Behavioral Research, 52*, 12–30. https://doi.org/10.1080/00273171.2016.1245600

Cragg, L., & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. *Trends in Neuroscience and Education, 3*, 63–68. https://doi.org/10.1016/j.tine.2013.12.001

Davis-Kean, P., Domina, T., Kuhfeld, M., Ellis, A., & Gershoff, E. T. (2021). It matters how You start: Early numeracy mastery predicts high school math course-taking and college attendance. February 11 https://doi.org/10.31234/osf.io/wdvth.

Diallo, T. M., Morin, A. J., & Parker, P. D. (2014). Statistical power of latent growth curve models to detect quadratic growth. *Behavior Research Methods, 46*, 357–371. https://doi.org/10.3758/s13428-013-0395-1

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*, 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428

Enders, C. K. (2001). The impact of nonnormality on full information maximum-likelihood estimation for structural equation models with missing data. *Psychological Methods, 6*, 352–370. https://doi.org/10.1037/1082-989X.6.4.352

Engel, M., Claessens, A., & Finch, M. A. (2013). Teaching students what they already know? The (mis) alignment between mathematics instructional content and student knowledge in kindergarten. *Educational Evaluation and Policy Analysis, 35*(2), 157–178. https://doi.org/10.3102/0162373712461850

Fan, X., & Fan, X. (2005). Power of latent growth modeling for detecting linear growth: Number of measurements and comparison with other analytic approaches. *The Journal of Experimental Education, 73*(2), 121–139. https://doi.org/10.3200/JEXE.73.2.121-139

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

Frick, A., Möhring, W., & Newcombe, N. S. (2014). Development of mental transformation abilities. *Trends in Cognitive Sciences, 18*, 536–542. https://doi.org/10.1016/j.tics. 2014.05.011

García Coll, C., Akiba, D., Palacios, N., Bailey, B., Silver, R., DiMartino, L., & Chin, C. (2002). Parental involvement in children's education: Lessons from three immigrant groups. *Parenting, 2*, 303–324. https://doi.org/10.1207/S15327922PAR0203_05

Gilligan, K. A., Thomas, M. S. C., & Farran, E. K. (2019). First demonstration of effective spatial training for near transfer to spatial performance and far transfer to a range of mathematics skills at 8 years. *Developmental Science*, 1–8. https://doi.org/10.1111/desc.12909

Grimm, H. (2000). *Sprachentwicklungstest für zweijährige Kinder (SETK-2)*. Hogrefe.

Grob, A., Keller, K., & Troesch, L. (2014). *ZWEITSPRACHE: Mit ausreichenden Deutschkenntnissen in den Kindergarten.* Universität Basel.

Grob, A., Meyer, C. S., & Hagmann-von Arx, P. (2009). *Intelligence and development scales (IDS)*. Huber.

Hagmann-von Arx, P., Meyer, C. S., & Grob, A. (2008). Assessing intellectual giftedness with the WISC-IV and the IDS. *Journal of Psychology, 216*, 173–180. https://doi.org/10.1027/0044-3409.216.3.172.

Hawes, Z., & Ansari, D. (2020). What explains the relationship between spatial and mathematical skills? A review of evidence from brain and behavior. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-019-01694-7

Hinojosa, T., Rapaport, A., Jaciw, A., LiCalsi, C., & Zacamy, J. (2016). *Exploring the foundations of the future STEM workforce: K–12 indicators of postsecondary STEM success (REL 2016–122)*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest.

Hong, S., & You, S. (2012). Understanding Latino children's heterogeneous academic growth trajectories: Latent growth mixture modeling approach. *The Journal of Educational Research, 105*, 235–244. https://doi.org/10.1080/00220671.2011.584921

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jenkinson, J., Roberts, S., Dennehy, S., & Tellegen, P. (1996). Validation of the Snijders–Oomen nonverbal intelligence test-revised 2 ½-7 for Australian children with disabilities. *Journal of Psychoeducational Assessment, 14*, 276–286. https://doi.org/10.1177/073428299601400307.

Jirout, J. J., & Newcombe, N. S. (2015). Building blocks for developing spatial skills: Evidence from a large, representative U.S. Sample. *Psychological Science, 26*, 302–310. https://doi.org/10.1177/0956797614563338

Jordan, N. C., Kaplan, D., Olah, L. N., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development, 77*, 153–175. https://doi.org/10.1111/j.1467-8624.2006.00862.x

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology, 45*, 850–867. https://doi.org/10.1037/a0014939

Kahl, T. P., Grob, A., Segerer, R., & Möhring, W. (2019). Executive functions and visual-spatial skills predict mathematical achievement: Asymmetrical associations across age. *Psychological Research*. https://doi.org/10.1007/s00426-019-01249-4. Advanced online publication.

Konsortium, Ü. G. K., & Hrsg. (2019). *Überprüfung der Grundkompetenzen. Nationaler Bericht der ÜGK 2016: Mathematik 11. Schuljahr.* EDK und SRED.

Kyttälä, M., Aunio, P., Lehto, J. E., Van Luit, J., & Hautamaki, J. (2003). Visuospatial working memory and early numeracy. *Educational and Child Psychology, 20*, 65–76.

Lachance, J. A., & Mazzocco, M. M. (2006). A longitudinal analysis of sex differences in math and spatial skills in primary school age children. *Learning and Individual Differences, 16*(3), 195–216. https://doi.org/10.1016/j.lindif.2005.12.001

Lauer, J. E., & Lourenco, S. F. (2016). Spatial processing in infancy predicts both spatial and mathematical aptitude in childhood. *Psychological Science, 27*, 1291–1298. https://doi.org/10.1177/0956797616655977

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., Foley, A., Lourenco, S., Ehrlich, S., & Ratliff, K. (2016). Sex differences in spatial cognition: Advancing the conversation. *Wiley Interdisciplinary Reviews: Cognitive Science, 7*, 127–155. https://doi.org/10.1002/wcs.1380

Levine, S. C., Vasilyeva, M., Lourenco, S. F., Newcombe, N. S., & Huttenlocher, J. (2005). Socioeconomic status modifies the sex difference in spatial skill. *Psychological Science, 16*, 841–845. https://doi.org/10.1111/j.1467-9280.2005.01623.x

Mix, K. S., & Cheng, Y. L. (2012). The relation between space and math: Developmental and educational implications. In J. B. Benson (Ed.), *Advances in child development and behavior* (Vol. 42, pp. 197–243). Elsevier Academic Press Inc.

Moore, C., O'Keefe, S. L., Lawhon, D., & Tellegen, P. (1998). Concurrent validity of the snijders–oomen nonverbal intelligence test 2 ½-7-revised with the wechsler preschool and primary scale of intelligence—revised. *Psychological Reports, 82*, 619–625. https://doi.org/10.2466/pr0.1998.82.2.619

Muthén, B., & Muthén, L. (2000). Integrating person-centered and variable-centered analysis: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research, 24*, 882–891. https://doi.org/10.1111/j.1530-0277.2000.tb02070.x

Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide* (8th ed.). Author.

National Assessment of Educational Progress. (2019). *Mathematics and reading assessments: Highlighted results at grades 4 and 8 for the nation, states, and districts*. National Center for Education Statistics, Institute of Education Sciences, U.S. Dept. of Education. NAEP.

Newcombe, N. S. (2010). Picture this: Increasing math and science learning by improving spatial thinking. *American Educator, 34*, 29–43.

Newcombe, N. S., Möhring, W., & Frick, A. (2018). How big is many? Development of spatial and numerical magnitude understanding. In A. Henik, & W. Fias (Eds.), *Heterogeneity of function in numerical cognition* (pp. 157–176). Elsevier Academic 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). Springer. https://doi.org/10.1007/978-94-017-9297-4_10.

Nylund, K. L., Asparouhov, T., & Muthen, B. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14*, 535–569. https://doi.org/10.1080/10705510701575396

Pruden, S. M., & Levine, S. C. (2017). Parents' spatial language mediates a sex difference in preschoolers' spatial language use. *Psychological Science, 28*, 1583–1596. https://doi.org/10.1177/0956797617711968

Purpura, D. J., Napoli, A. R., Wehrspann, E. A., & Gold, Z. S. (2017). Causal connections between mathematical language and mathematical knowledge: A dialogic reading intervention. *Journal of Research on Educational Effectiveness, 10*, 116–137. https://doi.org/10.1080/19345747.2016.1204639

Renner, G., Rausch, N., Krampen, G., & Irblich, D. (2009). Der SON-R 2.5-7 in der klinischen Praxis. Reliabilität, Validität und Erprobung einer Kurzform. *Kindheit und Entwicklung, 18*, 232–243.

Sadler, P. M., & Tai, R. H. (2007). The two high-school pillars supporting college science. *Science, 317*, 457–458.

Swiss Federal Statistical Office. (2012). löhne erwerbseinkommen – detaillierte daten [wages, income – detailed data]. http://www.bfs.admin.ch/bfs/portal/de/index/themen/0304/blank/data/00.html.

Tellegen, P. J., Laros, J. A., & Petermann, F. (2007). *Snijders-oomen non-verbaler intelligenztest*. Hogrefe.

Tofighi, D., & Enders, C. K. (2008). Identifying the correct number of classes in growth mixture models. In R. Hancock, & K. M. Samuelsen (Eds.), *Advances in latent variable mixture models* (pp. 317–341). Information Age Publishing Inc.

Troesch, L. M., Segerer, R., Claus-Pröstler, N., & Grob, A. (2021). Parental Acculturation Attitudes: Direct and Indirect Impacts on Children's Second Language Acquisition. Early Education and Development, 32:2, 272-290. doi: 10.1080/10409289.2020.1740640.

United Nations. (2019). *International migrant stock 2017*. Department of Economic and Social Affairs. https://www.un.org/en/development/desa/population/migration/data/estimates2/estimates19.asp.

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*(2), 352–402. https://doi.org/10.1037/a0028446

Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., Newcombe, N. S., Filipowicz, A. T., & Chang, A. (2014). Deconstructing building blocks: Preschoolers' spatial assembly performance relates to early mathematical skills. *Child Development, 85*, 1062–1076. https://doi.org/10.1111/cdev.12165

Verhoeven, M. (2011). Multiple embedded inequalities and cultural diversity in educational systems: A theoretical and empirical exploration. *European Educational Research Journal, 10*(2), 189–203.

Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political Analysis, 18*, 450–469. https://doi.org/10.1093/pan/mpq025

Wang, M., & Bodner, T. E. (2007). Growth mixture modeling: Identifying and predicting unobserved subpopulations with longitudinal data. *Organizational Research Methods, 10*, 635–656. https://doi.org/10.1177/1094428106289397

Xie, F., Zhang, L., Chen, X., & Xin, Z. (2020). Is spatial ability related to mathematical ability: A meta-analysis. *Educational Psychology Review, 32*, 113–155.

Zhang, X., Koponen, T., Rasanen, P., Aunola, K., Lerkkanen, M. K., & Nurmi, J. E. (2014). Linguistic and spatial skills predict early arithmetic development via counting sequence knowledge. *Child Development, 85*, 1091–1107. https://doi.org/10.1111/cdev.12173

Zhang, X., & Lin, D. (2017). Does growth rate in spatial ability matter in predicting early arithmetic competence? *Learning and Instruction, 49*, 232–241. https://doi.org/10.1016/j.learninstruc.2017.02.003.