Please cite the paper:

Szczygieł, M., & Hohol, M. (2026). Individual and Environmental Predictors of Math Anxiety in Primary School Children. *Journal of Experimental Child Psychology*.
https://doi.org/10.1016/j.jecp.2026.106711

Individual and Environmental Predictors of Math Anxiety in Primary School Children

Monika Szczygieł[1*], Mateusz Hohol[2]

[1] Institute of Psychology, Faculty of Philosophy, Jagiellonian University in Krakow, Poland
ORCID: 0000-0001-6544-0734

[2] Mathematical Cognition and Learning Lab, Copernicus Center for Interdisciplinary Studies, Jagiellonian University in Krakow, Poland
ORCID: 0000-0003-0422-5488

[*]Corresponding author address: Ingardena 6, 30-060 Krakow, Poland
Email: monika.szczygiel@uj.edu.pl

Abstract

Previous meta-analyses have indicated that math anxiety (MA) is weakly to moderately and negatively related to math achievement, yet relatively few studies have simultaneously examined a range of predictors of MA and the proportion of variance in MA explained by these predictors. We investigated individual and environmental predictors of MA using multilevel regression models based on data from 166 fourth-grade primary school children in Poland. We tested four models, sequentially introducing sociodemographic factors (gender, age), math domain-specific variables (math achievement, intellectual helplessness in math), domain-general affective factors (general and test anxiety), domain-general cognitive variables (fluid intelligence, working memory), and family-related factors (mothers' home numeracy activities, MA, and education), while controlling for school as a fixed effect and including class as a random effect. Test anxiety was a stable predictor of MA across successive regression models. Intellectual helplessness in math and fluid intelligence were associated with MA in some models, but these predictors were not stable across successive models. The fixed effects, including school, explained approximately 58–63% of the variance in MA, while differences between classes accounted for only 1.4–2.4% of the variance. These findings underscore the need to consider both individual and environmental factors when developing interventions aimed at reducing MA and fostering a supportive learning environment.

Keywords

math anxiety, test anxiety, mathematical education, home numeracy, multilevel regression

Introduction

Math proficiency is essential both for individuals who encounter numbers in daily tasks such as calculating tips, evaluating discounts, and assessing financial risks, and for societies, as mathematical skills are linked to economic growth and participation in STEM careers (OECD, 2013; Reyna et al., 2009). Because math proficiency begins to develop during the early stages of education, it is important to identify factors that may hinder math learning in primary school children. One of them is math anxiety (MA), defined as "a feeling of tension and anxiety that interferes with the manipulation of numbers and the solving of mathematical problems in a wide variety of ordinary life and academic situations" (Richardson & Suinn, 1972, p. 551). Academic situations include both learning math and being tested on it (Dowker et al., 2016). Although MA can manifest as either a trait or a state, and both forms have been linked to math achievement (Dowker, 2026), the present study focuses on trait MA. This focus reflects our interest in children's relatively stable tendency to experience MA and its association with individual and environmental characteristics, rather than momentary affective responses to a particular mathematical task. Meta-analyses have consistently shown that trait MA is weakly to moderately and negatively associated with math achievement and that this association emerges as early as the first years of primary school (Barroso et al., 2021; Namkung et al., 2019; Pelegrina et al., 2026).

Although considerable progress has been made in understanding both the relationship between MA and math achievement (Lau et al., 2024; Pelegrina et al., 2026) and the predictors of math achievement (Lavelle-Hill et al., 2024), the predictors of MA remain understudied. Numerous studies have identified individual correlates of MA, but relatively few have examined multiple predictors simultaneously, and even fewer have considered both individual and environmental predictors (e.g., del Río et al., 2025; Lau et al., 2022, 2024; Leppma & Darrah, 2024; Jameson, 2014; Meece et al., 1990; Szczygieł et al., 2024, 2025; Vanbinst et al.,

2020). Consequently, relatively little is known about which variables predict MA when several theoretically relevant factors are considered simultaneously and about the proportion of variance in MA jointly accounted for by these predictors. Addressing these gaps is important for planning effective mathematics education, as it can help identify which factors should be targeted to prevent or reduce MA.

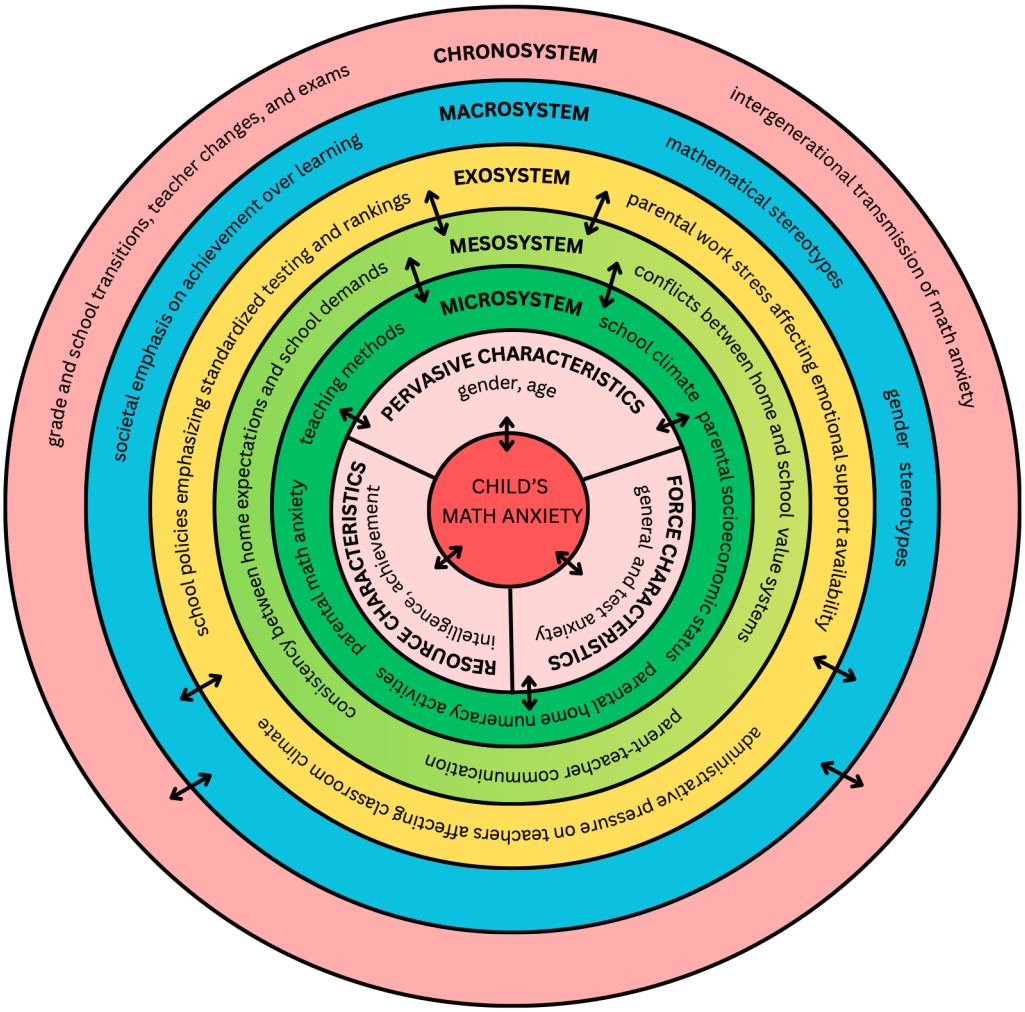
Predictors of MA

We organize variables described in the literature as correlates or predictors of MA within an explicit theoretical framework, providing a basis for their selection. Our framework is based on Luttenberger et al.'s (2018) distinction between individual and environmental factors and Bronfenbrenner's ecological model (Bronfenbrenner & Morris, 2007), recently applied to MA by Retanal and Maloney (2026).

Individual factors concern characteristics of the child, whereas environmental factors include the family, school, and broader social context (Luttenberger et al., 2018). Within Bronfenbrenner's model, individual characteristics can be classified into three categories: pervasive (relatively stable attributes of the individual), force (emotional and motivational tendencies shaping responses to situations), and resource (cognitive and experiential capacities available to the individual). Environmental factors are organized within nested ecological systems, including the microsystem (immediate environment), mesosystem (relations between immediate environments), exosystem (broader institutional contexts), and macrosystem (societal and cultural influences).

Figure 1 presents the proposed framework and the variables selected for the present study, which are discussed in detail below. For each variable, we indicate whether previous research has identified it as a correlate or as a predictor of MA and, where such information is available, report how much variance in MA was explained either by the variable itself or jointly with other predictors.

Figure 1. Model of the determinants of MA based on theoretical frameworks of Luttenberger et al. (2018) and Bronfenbrenner and Morris (2007).



Individual Factors

Pervasive Characteristics: Sociodemographic

Pervasive characteristics represent relatively stable attributes of the child that shape how they experience learning situations. In the present study, this category comprises two sociodemographic factors: gender and age. Gender may be relevant because girls may be more vulnerable to math-related stereotypes (Maloney et al., 2013), whereas age may capture developmental differences in children's experience with mathematical demands and their ability to regulate emotional responses to challenging or uncertain situations (Brocki & Bohlin, 2004; Wesarg-Menzel et al., 2023).

Several studies have indicated a gender gap in MA across educational levels (Eidlin-Levy et al., 2023; OECD, 2013; Szczygieł et al., 2024), with girls and women typically reporting higher MA than boys and men. However, other studies have failed to observe this difference (Ganley & McGraw, 2016; Namkung et al., 2023), indicating that the evidence for gender differences in MA is not entirely consistent. Studies that considered gender together with other relevant variables suggest that its association with MA may be partly explained by broader affective characteristics, although a small independent effect of gender may remain. Szczygieł (2020a) found that the relationship between gender and MA in children in Grades 1 and 2 was fully mediated by general anxiety. However, in a later longitudinal study of primary school children, gender remained a small but significant predictor of MA at the end of Grade 1 after previous MA and several cognitive and affective predictors were considered simultaneously (Szczygieł et al., 2024). In a large international sample of Grade 4 students, gender remained a significant predictor of MA when a broad range of student-, family-, and school-related characteristics were considered simultaneously (Lau et al., 2022). A similar pattern has been observed in older samples. Among secondary school students, Szczygieł et al. (2025) found that the effect of female gender on MA was weak and decreased after general and test anxiety were added to the model, but remained significant. In young adults, gender also showed a small effect, whereas in another adult sample it was no longer significant after domain-specific, domain-general affective, and domain-general cognitive variables were considered simultaneously. Leppma and Darrah (2024) similarly found a small gender effect after accounting for self-efficacy, mindfulness, self-compassion, and field of study. Taken together, these findings suggest that gender is unlikely to be a strong independent predictor of MA, but its association with MA is sufficiently consistent to warrant consideration.

Previous research suggests that MA emerges in early childhood and may increase across development (Dowker et al., 2016). Within the same grade, age may nevertheless capture

modest developmental differences in emotion regulation and coping abilities (Brocki & Bohlin, 2004; Wesarg-Menzel et al., 2023). Evidence on age as an independent predictor of MA is limited. Szczygieł et al. (2025) found no age effect in one adult sample, a weak negative association in secondary-school students, and a weak positive association in another adult sample, suggesting that the role of age may vary across developmental and/or educational contexts. Given the scarcity of evidence for an age effect in primary school children, age should still be considered, at least as a control variable for developmental differences between children.

Force Characteristics: Domain-General Affective

Force characteristics reflect children's general emotional responses to learning and testing situations. In the present study, they are represented by domain-general affective tendencies, specifically general anxiety and test anxiety. These forms of anxiety reflect a broader vulnerability that may increase children's susceptibility to anxiety in math-related contexts (Ashcraft & Kirk, 2001; Mandler & Sarason, 1952). General anxiety refers to a dispositional tendency to experience worry across a wide range of situations (Spielberger, 1983), whereas test anxiety refers to anxiety experienced in evaluative contexts (Cassady, 2023). Both constructs are conceptualized as relatively stable traits and are often considered as precursors of domain-specific forms of anxiety (Hirshfeld-Becker et al., 2008).

Both general and test anxiety are positively associated with MA, with particularly strong associations observed for test anxiety (Caviola et al., 2022; Szczygieł et al., 2025). In primary school children, general anxiety has been identified as an important predictor of MA. Szczygieł et al. (2024) found that it predicted MA at school entry when sociodemographic, domain-general, and domain-specific variables were considered simultaneously. However, test anxiety was not included in that study. Evidence from models including both forms of anxiety is more limited. Among secondary school students, Szczygieł et al. (2025) found that both general and test anxiety remained significant after gender, age, and domain-specific predictors were taken

into account, with test anxiety emerging as the strongest predictor. Their addition increased the variance explained in MA from 46% to 59%, weakened the effects of gender and age, and made math performance no longer significant. In contrast, in adults, neither general nor test anxiety remained significant once sociodemographic and domain-specific variables were considered, whereas age, mathematical resilience, intellectual helplessness in math, and math performance remained significant.

Thus, considering general and test anxiety simultaneously is important for determining whether each contributes independently to MA, whether their inclusion changes the predictive effects of other variables, and how much additional variance in MA they jointly explain.

Resource Characteristics: Math Domain-Specific

Resource characteristics refer to the individual capacities and abilities that children bring to mathematical learning (Bronfenbrenner & Morris, 2007). In the present study, these include both math domain-specific resources and domain-general cognitive abilities. A range of math-specific variables have been linked to MA, for example, magnitude processing (Maloney et al., 2011), math self-concept (Ahmed et al., 2012), math achievement (Field et al., 2019), and intellectual helplessness in math (Szczygieł et al., 2025). We focus on math achievement and intellectual helplessness in math, as potential domain-specific predictors of MA (Szczygieł et al., 2024, 2025). Math achievement refers to success in mathematical tasks, typically measured through grades or performance tests. Intellectual helplessness in math is a domain-specific form of learned helplessness in which individuals perceive their cognitive efforts as ineffective and mathematical outcomes as beyond their control (Sędek & McIntosh, 1998; Seligman, 1972). It may develop following repeated difficulties or failure and is associated with reduced motivation and persistence, as well as greater avoidance of mathematics (Gürefe & Bakalım, 2018; Hwang, 2019).

Meta-analyses consistently show a negative association between math achievement and MA (Barroso et al., 2021; Namkung et al., 2019), while longitudinal evidence supports a reciprocal relationship in which lower math achievement predicts subsequent MA and MA predicts subsequent math achievement (Pelegrina et al., 2026). When math achievement is considered alongside other potential predictors, however, its independent contribution to MA appears less consistent. In primary school children, Szczygieł et al. (2024) found that math achievement at the end of Grade 1 predicted MA at the end of Grade 2 even after earlier MA, sociodemographic, domain-general, and domain-specific variables were considered. Similarly, Sorvo et al. (2019) found that earlier arithmetic achievement predicted subsequent anxiety about failure in mathematics among children in Grades 2–4. In students in Grades 7–9, Meece et al. (1990) found that prior math grades did not directly predict MA; instead, their association with MA was mediated by students' perceptions of their mathematical ability. The full model explained about half of the variance in MA. Among secondary school students, math performance predicted MA when sociodemographic and domain-specific variables were considered, but this effect became nonsignificant after general and test anxiety were added (Szczygieł et al., 2025). In contrast, math performance remained a significant predictor in both adult samples: when considered with sociodemographic and domain-specific variables in one sample, and after the additional inclusion of affective and cognitive variables in the other (Szczygieł et al., 2025).

Evidence for intellectual helplessness in math is more consistent, although it comes primarily from adolescents and adults. In secondary school students, intellectual helplessness in math was the strongest predictor of MA before general and test anxiety were added and remained the second strongest, after test anxiety, once these affective factors were included. In adults, it remained one of the strongest predictors when sociodemographic, affective, and cognitive variables were considered simultaneously (Szczygieł et al., 2025). Across these

studies, the full models explained approximately 46–59% of the variance in MA, although this variance cannot be attributed to intellectual helplessness in math alone.

Thus, considering math achievement and intellectual helplessness in math together with other relevant predictors is important for establishing their independent contributions to MA and determining whether these contributions change once broader affective and cognitive characteristics are taken into account.

Resource Characteristics: Domain-General Cognitive

Resource characteristics also include domain-general cognitive abilities that support learning and problem-solving across domains. Following Szczygieł et al. (2024, 2025), we consider fluid intelligence and working memory as potential predictors of MA. Fluid intelligence refers to the capacity to reason and adapt effectively to novel situations (Kent, 2017), whereas working memory refers to the temporary storage and manipulation of task-relevant information (Baddeley, 1992).

Some studies have found that higher fluid intelligence and working memory are associated with lower MA (Pellizzoni et al., 2022; Schillinger et al., 2018; Živković et al., 2023); however, when other potential predictors are considered simultaneously, evidence for their independent contributions to MA is limited. In primary school children, Ramirez et al. (2013) found no association between children's MA and digit-span working memory after controlling for grade level, while in Szczygieł et al.'s (2024) study, neither fluid intelligence nor verbal or visuospatial working memory predicted MA when sociodemographic, domain-general affective, and math domain-specific variables were considered simultaneously. Similarly, in adults, fluid intelligence and working memory did not predict MA after sociodemographic, math domain-specific, and domain-general affective variables were taken into account, and adding domain-general cognitive variables did not increase the variance explained in MA (Szczygieł et al., 2025).

Although fluid intelligence and working memory may be weakly associated with MA, evidence for their independent predictive contribution remains limited, particularly in primary school children. Their inclusion therefore allows us to test whether domain-general cognitive abilities explain additional variance in MA beyond sociodemographic, math domain-specific, and domain-general affective factors.

Environmental Factors

Microsystem: Family

Within Bronfenbrenner's framework, the microsystem includes children's immediate environments in which they directly participate and engage in repeated interactions with significant others, particularly family members, teachers, and peers (Bronfenbrenner & Morris, 2007; Retanal & Maloney, 2026). In the family environment, parents' MA and education, as well as home numeracy activities occurring in parent-child interactions, may shape children's early experiences with math and their emotional responses to it (del Río et al., 2025; Retanal & Maloney, 2026). Parental MA refers to anxiety experienced by parents in math-related situations (Richardson & Suinn, 1972), whereas home numeracy activities refer to math-related interactions and activities that parents engage in with their children (LeFevre et al., 2009). Parental education refers to the parent's level of formal educational attainment.

Although previous studies tested the relationship between parental characteristics and children's MA, the findings were inconsistent. Jameson (2014) found no correlation between parental MA and second-grade children's MA, whereas more frequent home math activities were correlated with higher MA of children, and associations between parental involvement and children's MA appear to depend on the form of involvement. Jerković and Milovanović (2020) found that fathers' helping itself was not correlated with second- to fourth-grade children's MA, whereas fathers' positive experiences with math and their expectations were associated with lower children's MA. Fathers' perceptions of difficulties in their children's

mathematics learning were associated with higher levels of children's MA. Vanbinst et al. (2020) found that sixth-grade children's MA was positively correlated with mothers' MA, but not fathers' MA, and negatively correlated with the educational level of both parents.

When family characteristics have been examined as predictors of MA alongside other variables, the findings have also been mixed. In second grade children, Jameson (2014) found that the frequency of home math activities positively predicted MA when child and family characteristics were considered together, although adding environmental variables did not significantly increase the variance explained in MA. In children in Grades 1–3, Szczygieł (2020b) found some predictive effects of parental MA, but these varied depending on the parent, child gender, and grade level. In Grade 6, Vanbinst et al. (2020) found that when mothers' and fathers' MA and education were considered simultaneously, parental education was no longer significant and only mothers' MA approached significance. Together, these parental variables explained 7.9% of the variance in children's MA. In Grade 1, del Río et al. (2025) found that advanced home math activities predicted lower MA, whereas parents' negative attitudes toward mathematics predicted higher MA when parental characteristics were considered simultaneously. In a large international sample of Grade 4 students, Lau et al. (2022) found that home mathematics activities, parental involvement in mathematics homework, parental attitudes toward mathematics, parental education, and home socioeconomic status were all significantly associated with MA when a broad set of individual, family, and school characteristics were considered simultaneously. Longitudinal findings for parental MA are also inconsistent: parental MA assessed when children were age 3 did not directly predict children's MA at age 8 (Tomasetto et al., 2025), whereas more recent longitudinal evidence from children in Grades 3–4 indicated that parental MA predicted subsequent MA partly through children's math learning involvement (Zhu et al., 2026).

Taken together, parental MA, home numeracy activities, and parental education or related socioeconomic characteristics have been associated with children's MA, but their independent predictive contributions remain inconsistent. Considering these family characteristics simultaneously is therefore important for determining whether each explains variation in children's MA beyond individual-level predictors and beyond the other family characteristics.

Microsystem: School

In the school microsystem, children's experiences with teachers and peers, and features of the educational environment such as instructional practices and classroom or school climate, may likewise be associated with children's MA (O'Hara et al., 2022; Wang et al., 2025). Supportive teacher-student relationships, positive classroom climate, and peer support tend to be related to lower MA, whereas competitive, highly evaluative, or punitive classroom environments tend to be associated with higher MA (O'Hara et al., 2022). Teachers' own characteristics may also be important. Teachers' MA and lower confidence in teaching mathematics have been associated with higher MA among students in some studies, although these relationships are not consistently observed across samples (McLean et al., 2023).

Longitudinal studies provide stronger evidence that features of the school environment may predict MA. In primary-school children, perceived teacher support has been found to predict subsequent MA or changes in MA over time, with some studies also indicating a role of peer support (Li et al., 2021, 2023; Zhuo et al., 2025). In Grade 4 students, Lau et al. (2022) simultaneously examined student-, family-, classroom-, and teacher-level characteristics. Students' perceptions of teacher competence were the strongest individual-level predictor of MA, whereas at the educational-environment level, greater teacher confidence in teaching mathematics was associated with lower MA and more frequent mathematics homework with higher MA. Their multilevel model also showed significant residual variation in MA at the

educational-environment level, indicating that students sharing the same educational context were more similar in MA than students from different contexts. Based on the reported variance components, approximately 3.6% of the residual variance in MA occurred at the educational-environment level after the included predictors were taken into account.

These findings suggest that MA is not solely an individual characteristic but may also reflect experiences shared within classrooms and schools. However, relatively few studies have estimated how much variation in MA occurs between educational contexts, particularly in primary school children and after individual and family-related predictors are taken into account.

Current Study

Novelty of the Study

Research examining MA as an outcome remains relatively limited, and previous studies have differed substantially in the predictors considered. Meece et al. (1990) focused primarily on math achievement, ability perceptions, expectancies, and value. Leppma and Darrah (2024) examined individual characteristics such as self-efficacy, mindfulness, and self-compassion. Szczygieł et al. (2024, 2025) considered several sociodemographic, affective, cognitive, and math domain-specific predictors. Studies combining individual and family factors have generally considered a relatively limited set of family-related predictors, such as parental MA, home mathematics activities, or parental attitudes (e.g., del Río et al., 2025; Jameson, 2014; Vanbinst et al., 2020). To our knowledge, no previous study has examined a broad set of individual and family-related predictors in successive models to determine which predictors remain significant, how strong their effects are, and whether these effects remain stable as additional theoretically relevant variables are introduced. Lau et al. (2022) provide an important exception by considering individual, family, and educational-context characteristics

simultaneously in Grade 4 students, although these predictors were not examined across successive theoretically organized models.

Previous studies have also differed considerably in the proportion of variance in MA accounted for by their models, ranging from 7.9% in a model that included parental characteristics in primary-school children (Vanbinst et al., 2020), through approximately 27% in a model of individual characteristics in undergraduate students (Leppma & Darrah, 2024) and approximately 42% in a model combining individual and environmental predictors in second-grade children (Jameson, 2014), to approximately half of the variance (46–59%) in broader models of individual predictors in adolescents and adults (Meece et al., 1990; Szczygieł et al., 2025). However, these estimates come from different developmental stages and substantially different sets of predictors, making it difficult to determine what proportion of variance in MA can be accounted for when sociodemographic, math domain-specific, domain-general affective, domain-general cognitive, and family-related factors are considered within the same study. Moreover, relatively little is known about how much variation in MA is associated with differences between schools and classes once individual and family-related factors are taken into account. The present study addresses these issues by examining both the strength and stability of individual predictors across successive models, the proportion of variance accounted for by these models, and variation in MA associated with school and class membership.

Existing studies on predictors of MA have focused mainly on children at the beginning of formal schooling (del Río et al., 2025; Szczygieł et al., 2024), adolescents (Meece et al., 1990), and older students or adults (Leppma & Darrah, 2024; Szczygieł et al., 2025). Although Lau et al. (2022) examined Grade 4 students across multiple countries, Grade 4 does not represent the same educational transition in all school systems. Relatively little is therefore known about predictors of MA specifically at the time when children move from generalist to

subject-specific mathematics instruction. In Poland, this transition occurs in Grade 4, when children move from a single-teacher model to mathematics taught by a specialist. Mathematics becomes a separate subject, and children encounter increasingly abstract content and more formal evaluation (Hamre & Pianta, 2001; OECD, 2024; Symonds & Galton, 2014; Zeidner, 1998). These changes may make mathematical difficulties and evaluation more salient and, consequently, increase vulnerability to MA. Thus, Grade 4 represents a particularly relevant stage for examining its predictors in the Polish educational context.

Research Questions and Hypotheses

Based on previous correlational findings and studies examining predictors of MA in children, adolescents, and adults (e.g., del Río et al., 2025; Lau et al., 2022, 2024; Leppma & Darrah, 2024; Jameson, 2014; Meece et al., 1990; Szczygieł et al., 2024, 2025; Vanbinst et al., 2020), we formulated the following research questions and, where research evidence was sufficient, we also specified hypotheses for particular predictors and sets of predictors.

1. Which individual and family-level variables, including sociodemographic factors (gender, age), math domain-specific factors (math achievement, intellectual helplessness in math), domain-general affective factors (general anxiety, test anxiety), domain-general cognitive factors (fluid intelligence, working memory), and family-related factors (parental education, home numeracy activities, parental MA), predict children's MA when other theoretically relevant variables are taken into account, and which predictors show the strongest and most stable associations across successive models?
2. What proportion of variance in MA is accounted for by the predictors included in the successive models and by differences between schools and classes?

We formulated the following hypotheses for each set of predictors. First, we tested a model with basic sociodemographic and domain-specific variables. This decision was guided by our theoretical framework, consistent with the observation that most research on MA focuses

on gender differences, age-related changes, and the role of mathematical achievement and math-specific affective-motivational factors. Based on previous research, we hypothesized that girls would report slightly higher MA than boys (Eidlin-Levy et al., 2023; OECD, 2013; Szczygieł et al., 2024). Age was included as a control variable because previous evidence regarding its independent association with MA is limited and inconsistent (Brocki & Bohlin, 2004; Wesarg-Menzel et al., 2023). We further hypothesized that lower math achievement would be associated with higher MA (Barroso et al., 2021; Namkung et al., 2019) and that intellectual helplessness in math would be positively related to children's MA (Szczygieł et al., 2025).

In the second model, we added domain-general affective factors and hypothesized that both general and test anxiety would positively predict MA, with test anxiety being a stronger predictor (Carey et al., 2017; Szczygieł et al., 2025). Our hypothesis stems from studies showing that MA is rooted in more general anxieties, particularly at the beginning of schooling (Szczygieł & Pieronkiewicz, 2022). Given the results of previous studies, we expected that sociodemographic and mathematics-specific variables may be important in predicting MA but may appear weaker when general and test anxiety are controlled for, as these broader forms of anxiety share substantial variance with MA and may partly explain observed gender differences in MA. Indeed, previous research has shown that MA is more strongly associated with general and test anxiety than with math achievement (Carey et al., 2017; Szczygieł et al., 2024, 2025). Next, we introduced cognitive abilities and hypothesized that higher fluid intelligence and verbal working memory would be associated with lower MA, although these associations were expected to be relatively weak (Pellizzoni et al., 2022; Schillinger et al., 2018; Živković et al., 2023). We expected that adding these variables to the model may weaken the strength of predictors from the previous levels, because general cognitive abilities are related to both mathematical achievement and affective traits (Szczygieł et al., 2024, 2025).

Finally, we tested family-related variables, hypothesizing that higher parental MA would be positively associated with children's MA, while more frequent home numeracy activities and higher parental education would be associated with lower children's MA (Maloney et al., 2015; Soni & Kumari, 2017; Szczygieł, 2020b; Vukovic et al., 2013). Overall, we aimed to explore whether individual factors would remain robust predictors of MA even after accounting for family-level variables. The inclusion of parent-level variables was intended to test whether individual variables were still important for predicting MA when family-related environmental factors were also taken into account.

We also examined the proportion of variance in MA accounted for by successive sets of predictors. Introducing predictors block by block, in a theoretically motivated order, allowed us to observe whether the association of each factor with MA remained stable, weakened, or disappeared once conceptually related variables were controlled for, and thereby to identify which predictors showed the strongest and most stable associations with MA when considered jointly. In addition to child- and parent-level factors, we also examined the proportion of variance in MA associated with differences between schools and classes.

Method

Sample

All participants were fourth-grade Polish students who took part in the study at the end of the 2023/2024 school year. Principals of selected schools were invited to participate and provided with the study's details. Once a school agreed, parental consent was obtained, and only children with permission were included. We recruited 171 students from 22 classes in nine primary schools across Lesser Poland (seven schools from urban area) and Silesian Voivodeships (one school from urban and one from suburban area) (see Supplementary Materials Table A for the detailed sample structure). Parental reports of learning disorders and health conditions were collected to characterize the sample and adapt testing procedures. Three

children diagnosed with dyscalculia were excluded. Among the remaining participants, parental surveys indicated dyslexia in 10 children, autism spectrum disorder in 6, ADHD in 3, intellectual disability in 1, visual impairment in 7, hearing impairment in 1, and chronic disease in 2. Polish was the native language for 137 children; 29 were bilingual (Ukrainian: 22; Russian: 5; English: 1; German: 1), and 2 children were excluded for insufficient Polish proficiency, leaving 166 participants (80 boys, 86 girls). The mean age at first measurement (n = 164; data missing for two children) was 131.54 months (SD = 4.31, range = 117–143). In total, 132 mothers and 28 fathers completed the questionnaires. All the materials described below, for both children and parents, were provided in the Polish version.

Measurements

Children

Modified Abbreviated Math Anxiety Scale for Elementary Students (mAMAS-E; Szczygieł, 2019). The mAMAS-E measures learning and testing anxiety (e.g., "Listening to another student explain a math formula"; "Thinking about an upcoming math test one day in advance"). It consists of nine items in which children indicate whether they feel anxious on a three-point Likert scale (0 = no, 1 = a little, 2 = yes). Mean scores reflect overall MA, with higher results indicating a higher MA. The mAMAS-E had α = .74.

Mathematical Test for 4th/5th Grade Children (MATH; Karpińska & Radtke, 2019). The MATH test assesses mathematical achievement expected by the end of fourth grade, in accordance with the Polish core curriculum. It covers symbolic arithmetic operations, fractions, geometric calculations, clock reading, measurement, unit conversions, and arithmetic word problems. Children completed ten open-ended tasks, with possible scores ranging from 0 to 25 points. A higher sum score indicates greater mathematical achievement. The mean difficulty level, calculated as total points divided by the maximum score, was .56, indicating moderate difficulty. The reliability of the MATH test was α = .72.

Intellectual Helplessness in Mathematics Scale (IHMS; Szczygieł et al., 2025; for the complete version see Sędek & McIntosh, 1998). We used a shortened version of the IHMS consisting of 10 items (e.g., "When called on to answer a math question, I forget almost everything I've learned"), which together form a total score. Participants responded on a five-point Likert scale (1 = never; 5 = always). A higher mean score indicates greater intellectual helplessness in math. The reliability of the IHMS was $\alpha = .89$.

State and Trait Anxiety Inventory (STAI-6-item; Szczygieł et al., 2025; for the original version see Marteau & Bekker, 1992). The six-item STAI was used to assess children's general trait anxiety. Students responded to statements about how they usually feel (e.g., "I am tense") using a four-point scale (1 = not at all typical of me, 2 = somewhat typical of me, 3 = quite typical of me, 4 = very typical of me). Higher mean scores indicate greater general trait anxiety. The scale's reliability was $\alpha = .67$.

Cognitive Test Anxiety Scale (CTAS-7; Szczygieł et al., 2025; for the complete version see Cassady, 2023). The seven-item CTAS was used to assess children's test anxiety. Children responded to statements about how they typically feel in testing situations (e.g., "When I take tests, I get so nervous that I make mistakes") using a four-point scale (1 = not at all typical of me, 2 = somewhat typical of me, 3 = quite typical of me, 4 = very typical of me). Higher mean scores indicate greater test anxiety. The reliability of the CTAS was $\alpha = .79$.

Raven's Standard Progressive Matrices (RAVEN; Jaworowska & Szustrowa, 2000; for the original version see Raven & Court, 2000). The RAVEN measures fluid intelligence through matrices where participants identify the missing element based on logical relationships. To reduce testing time while maintaining appropriate difficulty, we selected parts B, C, and D, omitting part A as too easy and part E as too difficult for most children. Each series included 12 items, for a sum score of 0–36. Average item difficulty (total points divided by maximum possible) was .66, indicating moderate level. Reliability was $\alpha = .85$.

WISC-R Digit Span Backward (WISC-R; Matczak et al., 1991; for the original version see Wechsler, 1974). To assess verbal working memory, children repeated sequences of digits in reverse order, starting with two digits and increasing to eight if successful. Each span level included four trials, and testing stopped after two consecutive mistakes. Scores ranged from 0 to 28, with higher sum scores reflecting better verbal working memory. Split-half reliability, Spearman-Brown corrected, was $r = .89$.

Parents

Sociodemographic Survey

Parental Education (EDU-M; EDU-F). Parental education was assessed using a demographic questionnaire completed by one parent. Information about both the mother's and father's education could be provided, regardless of which parent completed the questionnaire. Respondents indicated the highest level completed by the mother and father separately, selecting from 12 predefined categories, ranging from primary education to full professor. For analyses, a binary variable distinguished between lower (up to post-secondary) and higher education (from BSc./Eng.). *Child Learning and Developmental Difficulties.* Information on children's learning and developmental difficulties was collected via a parent-report questionnaire. A parent indicated whether the child had been formally diagnosed with any neurodevelopmental or health-related condition. *Child Polish Language Proficiency*. One parent reported their child's language proficiency: native Polish, bilingual (specifying the second language), or non-Polish speaking. *Child Sociodemographic Data*. A parent reported the child's date of birth and gender.

Home Numeracy Activities Scale (HNAS; Girard et al., 2021; LeFevre et al., 2009). The HNAS measures how frequently parents engage in math-related interactions with their children at home, spanning informal, formal-basic, and formal-advanced activities. The scale can be used either by analyzing each type of activity separately or by combining them into a

single indicator. Because home numeracy practices are context-sensitive and the original items targeted preschoolers, we adapted the scale for fourth grade students. First, we increased task difficulty to align with the skills expected of fourth graders in the Polish core curriculum. Second, we added new items, particularly in the formal-advanced category, to better capture math-related support parents might provide at this stage. The adapted scale included 44 activities: 17 informal (e.g., "Counting money while shopping, playing number games"), 14 formal-basic (e.g., "Counting objects, adding numbers, multiplying single-digit numbers"), and 13 formal-advanced (e.g., "Converting units, performing written calculations, working with fractions"). One parent rated how often they engaged in each activity with their child over the past month on a five-point Likert scale (1 = never/very rarely to 5 = very often). The HNAS also included an additional response option to clarify cases where parents selected "1" but we did not analyze these responses because only a minority of parents completed this supplementary part of the questionnaire. The mean score across all 44 items served as an overall indicator of parental involvement in math learning, with higher scores reflecting greater home numeracy. The reliability was $\alpha = .97$ for mothers and $\alpha = .98$ for fathers.

Math Anxiety Questionnaire for Adults (MAQA; Szczygieł, 2022). The MAQA assesses adults' MA related to solving math problems (e.g., "Calculation of 12% less than 215"). MAQA consists of 19 items rated on a four-point scale (1 = I definitely do not feel anxiety; 4 = I definitely feel anxiety). Higher mean scores indicate greater MA. The reliability was $\alpha = .95$ for mothers and $\alpha = .98$ for fathers.

Procedure

The study was conducted in accordance with the principles of the Declaration of Helsinki, and its protocol was approved by the Research Ethics Committee at the Faculty of Philosophy, Jagiellonian University (approval number: 221.0042.28_2024). Before data collection with children, parents completed consent, sociodemographic survey, HNAS, and

MAQA. Recruitment of parents and children for the study took place during school parent-teacher meetings (usually only one parent, mainly mothers, completed the questionnaires, which affected data collection at the parental level). The study was conducted in primary schools by four trained research assistants across two sessions (with an interval of approximately 3 weeks between measurements), during which children completed IHMS, MATH, STAI-6, CTAS-7, RAVEN, WISC-R, mAMAS-E. All tasks, except for the MATH, were administered individually in a quiet room without teachers. The MATH task was administered in small groups, resembling a typical classroom test setting, and was also conducted by the assistants. For IHMS, STAI-6, CTAS-7, and mAMAS-E, instructions and questions were read aloud to the children to minimize the influence of reading ability on their responses. Children answered verbally using a printed response-scale for reference. The RAVEN and WISC-R were administered per their manuals: children either pointed to answers in the booklet or repeated digits after the assistant. All responses were recorded. For the MATH task, children received printed problems and solved them independently during a 45-minute lesson.

Data Analysis

Data Availability

All data, analysis code, and research materials are available on the Open Science Framework at https://osf.io/am4yd. Data were handled, inspected, and analyzed using the R language in *RStudio* (Posit Team, 2025).

Missing Data

The proportion of missing data differed across groups of variables due to the study design, wherein either mothers or fathers completed the parent questionnaire. Missingness was very low for child-reported and performance-based variables (0–4.82%), moderate for mother-reported variables (20–21%), and substantially higher for father-reported variables (84.34%).

Missing data were handled using listwise deletion in the main analyses. The pattern of missing data for these analyses was examined using Little's MCAR test, which was non-significant, $\chi^2(76) = 89.00$, $p = .146$, indicating that the data were consistent with the assumption of missing completely at random. The main analyses (see Table 3) were conducted without imputing missing data. However, to examine the robustness of the results across the four models, two sensitivity analyses were conducted. First, all four models were re-estimated in the complete-case sample, including only participants with complete data on all variables in Model 4 (see Supplementary Materials Table B). Second, missing values in child and maternal variables were handled using multiple imputation by chained equations (MICE), with 20 imputations and 20 iterations (see Supplementary Materials Table C).

Descriptive Statistics and Correlation Matrix

Pairwise deletion was applied when calculating zero-order Pearson correlations. Because only one parent (either mother or father) completed the questionnaire per family, correlations for home numeracy activities and MA between mothers and fathers are not available in Table 2.

Multilevel Linear Regression Models

Categorical variables were coded as factors. All continuous variables were *z*-scored. Each regression coefficient represents the expected change in MA (in standard deviation units), associated with a one standard deviation increase in the predictor, controlling for the remaining variables.

Given the clustered structure of the data, we first estimated a series of unconditional models to determine the appropriate analytic specification. Specifically, we compared null models including random intercepts for (a) classes nested within schools, (b) schools only, (c) classes only, (d) classes with school entered as a fixed effect, and (e) crossed random intercepts for classes and schools. These models were used to examine the distribution of variance in MA

across levels and to calculate intraclass correlation coefficients (ICCs). Then, four nested mixed-effects models were estimated to examine the incremental contribution of distinct sets of predictors. This sequential specification was adopted not because we expected each block to produce a large increase in explained variance, but because it allowed us to assess the stability of individual predictors as additional, potentially confounding, variables were introduced. Model 1 included gender, age, math achievement, and intellectual helplessness in math; Model 2 added general anxiety and test anxiety; Model 3 added fluid intelligence and verbal working memory; Model 4 added mother's home numeracy, MA, and education. Only mother-level variables were included in Model 4 because substantially more mothers than fathers completed the survey. Although education was requested from both parents, questions on home numeracy and MA pertained only to the respondent. Since including father-level predictors would have substantially reduced the sample, father-level associations are shown only in the correlation matrix (Table 2).

All models were estimated using restricted maximum likelihood (REML). We computed marginal R^2 (variance explained by fixed effects) and conditional R^2 (variance explained by fixed and random effects). In addition, variance components were extracted from each model to calculate ICCs and to examine how successive models reduced class-level and residual variance relative to the unconditional model. Class variance reduction (CVR) and residual variance reduction (RVR) were calculated as the proportional decrease in variance components after adding predictors, relative to the unconditional model estimated on the full sample.

Diagnostic analyses indicated no substantial multicollinearity among predictors (all VIFs < 5). Tests of heteroscedasticity indicated homoscedastic residual variance across models (ps > .88). Residual distributions approximated normality, with minor deviations in the first two

models that disappeared after additional predictors were introduced. None of the models exhibited singularity problems.

To provide more robust inference given the relatively small number of classes, bootstrap standard errors and 95% confidence intervals were estimated for the mixed-effects models using parametric bootstrapping procedures, with 50,000 replications. These bootstrap intervals were used to characterize uncertainty around both fixed-effect estimates and variance components.

Sensitivity analyses were conducted by (a) re-estimating all models on the same complete-case subsample and (b) repeating the analyses using datasets with imputed values to examine whether missing data influenced the results. These additional analyses were conducted to evaluate the robustness and stability of the model estimates across alternative analytic specifications (see Supplementary Materials Tables B, C, and D).

Finally, because REML log-likelihoods are not comparable across models that differ in their fixed-effects structure, we additionally re-estimated all models using full maximum likelihood (ML) and compared the nested models using likelihood ratio tests and the Akaike and Bayesian information criteria (AIC and BIC), computed on the common complete-case sample (n = 111). Full details of these procedures and the corresponding results are reported in the Supplementary Materials (Table E).

Results

Descriptive Statistics and Correlation Matrix

Descriptive statistics are presented in Table 1. Zero-order correlations between all variables are presented in Table 2. MA was positively correlated with general anxiety (weak), test anxiety (moderate), intellectual helplessness in math (moderate), and mothers' MA (weak), and negatively correlated with fathers' education (weak), indicating higher levels of MA among children whose fathers had lower levels of education. Girls had higher MA than boys. MA was

not significantly associated with age, math achievement, fluid intelligence, verbal working memory, mothers' education, or home numeracy activities.

Table 1. *Descriptive statistics*

	N	*M*	*SD*	*Min*	*Max*	*Sk.*	*K.*
Gender	166	.52					
Age	164	131.54	4.31	117	143	0.09	0.05
mAMAS-E	166	0.74	0.44	0	1.75	0.19	-1.02
STAI-6	166	1.85	0.44	1	3.33	0.57	0.56
CTAS-7	166	2.15	0.65	1	4	0.34	-0.31
IHMS	159	2.10	0.75	1	4.3	0.59	-0.35
RAVEN	166	23.57	5.77	7	35	-0.64	0.09
WISC-R	165	10.43	3.25	4	23	0.68	1
MATH	158	14.10	5.24	0	24	-0.33	-0.46
MAQA-M	132	1.99	0.71	1	3.89	0.58	-0.38
MAQA-F	26	1.75	0.66	1	3.16	0.68	-0.89
HNAS-M	131	2.72	0.93	1	4.63	0.07	-0.90
HNAS-F	26	2.32	0.81	1.09	3.89	0.23	-1.17
EDU-M	147	0.73	0.44	0	1	-1.05	-0.86
EDU-F	141	0.61	0.49	0	1	-0.46	-1.82

Age – the mean age at first measurement in months, mAMAS-E – MA, STAI-6 – general anxiety, CTAS-7 – test anxiety, IHMS – intellectual helplessness in mathematics, RAVEN – fluid intelligence, WISC-R – verbal working memory, MATH – math achievement, MAQA-M – mothers' MA, MAQA-F – fathers' MA, HNAS-M – mothers' home numeracy activities , HNAS-F – fathers' home numeracy activities, EDU-M – mothers' education, EDU-F – fathers' education. Gender was coded 0 – boys, 1 – girls; EDU-F and EDU-M were coded 0 – lower and 1 – higher education.

Table 2. *Pearson's zero-order correlations*

	Gender	Age	mAMAS-E	STAI-6	CTAS-7	IHMS	RAVEN	WISC-R	MATH	MAQA-M	MAQA-F	HNAS-M	HNAS-F	EDU-M
Age	-.24**													
mAMAS-E	.18*	-.14												
STAI-6	.16*	-.21**	.31***											
CTAS-7	.25***	-.17*	.49***	.41***										
IHMS	.12	-.13	.38***	.37***	.59***									
RAVEN	.05	-.03	-.11	.02	-.16*	.00								
WISC-R	.03	-.06	-.02	-.10	-.13	-.11	.36***							
MATH	-.08	-.01	-.15	-.12	-.35***	-.38***	.20**	.31***						
MAQA-M	-.03	.07	.21*	.09	.16	.15	-.11	-.13	-.17					
MAQA-F	.13	.18	.11	.11	-.34	-.13	-.03	-.10	-.01	-				
HNAS-M	.07	-.02	.01	-.04	.13	-.11	-.12	-.14	-.06	-.03	-			
HNAS-F	.20	.13	.18	-.36	.01	.03	-.27	-.30	-.29	-	.14	-		
EDU-M	.01	-.15	-.12	.07	-.13	-.04	.17*	.16	.26**	-.28**	-.02	-.10	-.34	
EDU-F	.05	-.13	-.26**	.02	-.19*	-.08	.09	.13	.18*	-.30***	-.38	-.08	-.32	.47***

*$p < .05$, **$p < .01$, ***$p < .001$, $n = 26–166$, pairwise deletion. Age – the mean age at first measurement in months, mAMAS-E – MA, STAI-6 – general anxiety, CTAS-7 – test anxiety, IHMS – intellectual helplessness in mathematics, RAVEN – fluid intelligence, WISC-R – verbal working memory, MATH – math achievement, MAQA-M – mothers' MA, MAQA-F – fathers' MA, HNAS-M – mothers' home numeracy activities, HNAS-F – fathers' home numeracy activities, EDU-M – mothers' education, EDU-F – fathers' education. Gender was coded 0 – boys, 1 – girls; EDU-F and EDU-M were coded 0 – lower and 1 – higher education

Multilevel Linear Regression Models

To determine the appropriate modeling strategy, a series of unconditional multilevel models were estimated. First, a model with classes nested within schools indicated that a substantial proportion of variance in MA was attributable to clustering, with ICC = .338. When only schools were modeled as a random effect, the ICC remained similar (.333), suggesting meaningful between-school differences in MA. A model including only class-level random intercepts yielded a higher ICC (.447), indicating that nearly 45% of the variance in MA could be attributed to differences between classes when school-level variation was not explicitly modeled. However, when school was entered as a fixed effect while retaining class as a random intercept, the class-level variance dropped substantially (ICC = .013). This pattern suggests that most of the between-class variability reflected systematic differences between schools rather than independent class effects. Inspection of the distribution of MA across schools further indicated pronounced differences between schools, with School 1 showing substantially higher mean levels of MA than the remaining schools. Given the small number of schools and the observed between-school heterogeneity, school was included as a fixed categorical predictor in subsequent models, while a random intercept for class was retained to account for clustering of students within classes. Four nested mixed-effects models were estimated to examine predictors of MA (see Table 3). In all models, school was included as a fixed categorical predictor (with School 1 as the reference category) to control for systematic between-school differences in MA, and a random intercept for class was specified to account for clustering of students within classes.

Table 3

Multilevel regression models of MA Predictors.

DV: mAMAS-E	Model 1				Model 2				Model 3				Model 4			
Random effects																
	Variance (*SD*) [95% CI SE]				Variance (*SD*) [95% CI SE]				Variance (*SD*) [95% CI SE]				Variance (*SD*) [95% CI SE]			
Class Intercept	0.007 (0.09) [0.00, 0.27]				0.009 (0.10) [0.00, 0.27]				0.008 (0.09) [0.00, 0.27]				0.006 (0.07) [0.00, 0.30]			
Residual	0.43 (0.65) [0.57, 0.73]				0.39 (0.62) [0.54, 0.69]				0.38 (0.62) [0.54, 0.69]				0.40 (0.63) [0.53, 0.72]			
Fixed effects																
	Estimate	SE	[95% CI]	*p*	Estimate	SE	[95% CI]	*p*	Estimate	SE	[95% CI]	*p*	Estimate	SE	[95% CI]	*p*
Intercept	1.03	0.13	[0.77, 1.28]	< .001	1.02	0.12	[0.77, 1.26]	< .001	1.01	0.12	[0.76, 1.25]	< .001	1.02	0.19	[0.65, 1.40]	< .001
School 2	-1.43	0.22	[-1.86, -1.01]	.002	-1.45	0.22	[-1.88, -1.03]	< .001	-1.50	0.21	[-1.92, -1.08]	< .001	-1.28	0.25	[-1.77, -0.78]	.002
School 3	-1.16	0.21	[-1.57, -0.74]	< .001	-1.17	0.20	[-1.57, -0.76]	< .001	-1.22	0.20	[-1.62, -0.82]	< .001	-1.20	0.23	[-1.65, -0.76]	< .001
School 4	-1.46	0.28	[-2.02, -0.92]	< .001	-1.39	0.27	[-1.92, -0.86]	< .001	-1.36	0.27	[-1.88, -0.83]	< .001	-1.35	0.30	[-1.94, -0.75]	< .001
School 5	-1.23	0.21	[-1.64, -0.82]	< .001	-1.14	0.20	[-1.54, -0.74]	< .001	-1.12	0.20	[-1.52, -0.72]	< .001	-.95	0.23	[-1.39, -0.50]	< .001
School 6	-1.86	0.20	[-2.25, -1.46]	< .001	-1.71	0.20	[-2.10, -1.33]	< .001	-1.70	0.20	[-2.09, -1.31]	< .001	-1.77	0.25	[-2.27, -1.27]	< .001

School 7	-1.54	0.18	[-1.89, -1.19]	< .001	-1.42	0.18	[-1.77, -1.07]	< .001	-1.37	0.18	[-1.73, -1.01]	< .001
School 8	-1.61	0.36	[-2.32, -0.90]	< .001	-1.57	0.35	[-2.26, -0.89]	< .001	-1.57	0.35	[-2.25, -0.88]	< .001
Gender	0.03	0.11	[-0.20, 0.25]	.808	-.04	0.11	[-0.26, 0.17]	.686	-0.03	0.11	[-0.24, 0.18]	.782
Age	0.09	0.06	[-0.03, 0.21]	.139	0.10	0.06	[-0.02, 0.21]	.103	0.10	0.06	[-0.01, 0.22]	.083
MATH	-0.12	0.07	[-0.25, 0.02]	.086	-0.07	0.06	[-0.20, 0.06]	.278	-0.06	0.07	[-0.19, 0.07]	.383
IHMS	**0.27**	**0.06**	**[0.16, 0.39]**	**< .001**	0.12	0.07	[-0.01, 0.26]	.074	**0.14**	**0.07**	**[0.01, 0.27]**	**.047**
STAI-6					0.04	0.06	[-0.07, 0.16]	.449	0.06	0.06	[-0.05, 0.18]	.296
CTAS-7					**0.26**	**0.07**	**[0.13, 0.40]**	**< .001**	**0.24**	**0.07**	**[0.10, 0.38]**	**< .001**
RAVEN									**-0.12**	**0.06**	**[-0.24, -0.01]**	**.037**
WISC-R									0.05	0.06	[-0.07, 0.16]	.431
MAQA-M												
HNAS-M												
EDU-M												

n	155 children, 20 classes	155 children, 20 classes	154 children, 20 classes	111 children, 18 classes
ICC	.017	.024	.021	.014

School 7	-1.43	0.24	[-1.90, -0.96]	< .001								
School 8	-1.52	0.51	[-2.52, -0.52]	.004								
Gender	-0.09	0.14	[-0.37, 0.18]	.513								
Age	0.11	0.07	[-0.03, 0.24]	.132								
MATH	-0.10	0.09	[-0.29, 0.08]	.266								
IHMS	0.09	0.09	[-0.09, 0.27]	.334								
STAI-6	0.01	0.07	[-0.13, 0.16]	.839								
CTAS-7	**0.24**	**0.10**	**[0.06, 0.43]**	**.012**								
RAVEN	-0.11	0.07	[-0.25, 0.03]	.121								
WISC-R	0.11	0.07	[-0.03, 0.25]	.139								
MAQA-M	0.01	0.07	[-0.13, 0.16]	.835								
HNAS-M	0.04	0.07	[-0.09, 0.16]	.590								
EDU-M	0.01	0.16	[-0.30, 0.32]	.968								

CVR (%)	98.37	97.86	98.14	98.76
RVR (%)	22.83	30.29	30.94	27.99
Marginal *R*2	.583	.622	.628	.595
Conditional *R*2	.590	.631	.636	.601
REML	325.9	317.3	319.6	241.8

*$p < .05$, **$p < .01$, ***$p < .001$, listwise deletion within each model. DV – dependent variable, Age – the mean age at first measurement in months, mAMAS-E – MA, STAI-6 – general anxiety, CTAS-7 – test anxiety, IHMS – intellectual helplessness in mathematics, RAVEN – fluid intelligence, WISC-R – verbal working memory, MATH – math achievement, MAQA-M – mothers' MA, HNAS-M – mothers' home numeracy activities, EDU-M – mothers' education. Gender was coded 0 – boys, 1 – girls; EDU-M was coded 0 – lower and 1 – higher education. CVR – class variance reduction (%) and RVR – residual variance reduction (%) were calculated relative to the unconditional classroom model as the proportional decrease in variance components after adding predictors. They were calculated relative to the unconditional classroom model estimated on the full sample $n = 155$.

Predictors of MA Across Successive Models

In Model 1, a one standard deviation increase in intellectual helplessness in math was associated with a 0.27 *SD* increase in MA. Gender, age, and math achievement were not significant predictors. After adding general and test anxiety in Model 2, a one standard deviation increase in test anxiety was associated with a 0.26 *SD* increase in MA, whereas general anxiety was not significantly related to MA. The association between intellectual helplessness in math and MA weakened and was no longer statistically significant. In Model 3, which additionally included fluid intelligence and verbal working memory, test anxiety remained a significant positive predictor (a 0.24 *SD* increase in MA per 1 *SD* increase in test anxiety). Intellectual helplessness in math again became statistically significant, indicating that a one standard deviation increase was associated with a 0.14 *SD* increase in MA. Fluid intelligence showed a negative association with MA, such that a one standard deviation increase in fluid intelligence corresponded to a 0.12 *SD* decrease in MA. Verbal working memory was not significantly associated with MA. In Model 4, which added mothers' home numeracy activities, mothers' MA, and their education, none of the mother-level variables were significantly associated with children's MA. In this model, test anxiety remained the only significant predictor, indicating that a one standard deviation increase in test anxiety corresponded to a 0.24 *SD* increase in MA, whereas the effects of intellectual helplessness in math and fluid intelligence were no longer significant.

A comparison of predictors' significance across the main and sensitivity analyses (see Supplementary Materials Tables B, C, and D) showed a largely consistent pattern. Test anxiety remained the most stable predictor, being significantly positively associated with MA in Models 2–4 across all analyses. Intellectual helplessness in math was significant in Models 1 and 3 in the main analysis and in Models 1–4 in the imputed models, but only in Model 1 in the complete-case analysis. Fluid intelligence was significantly negatively associated with MA

in Model 3 in the main and imputed analyses but was not significant in the complete-case models. Gender, age, verbal working memory, and mother's variables (MA, home numeracy activities, and education) were not significantly associated with MA in any model specification.

Variance Accounted for by Predictors and School- and Class-Level Differences

Across models, the proportion of variance in MA explained by fixed effects ranged from 58% to 63%, while the total variance explained by both fixed and random effects ranged from 59% to 64%. The highest explanatory power was observed for Model 3, whereas Model 4 showed slightly lower R^2 values, likely due to the reduced sample size resulting from missing parent-reported data.

Formal model comparison conducted with full maximum likelihood (ML; Supplementary Materials, Table E) indicated that adding the affective variables (general and test anxiety) improved model fit, $\chi^2(2) = 8.15$, $p = .017$, whereas the subsequently added cognitive variables (fluid intelligence and verbal working memory) and maternal variables did not, $\chi^2(2) = 4.35$, $p = .114$, and $\chi^2(3) = 0.35$, $p = .949$, respectively. AIC favored Models 2–3, whereas BIC favored Models 1–2; both criteria provided little support for Model 4. The non-significance of the cognitive block also held on the larger child-complete sample ($n = 154$), $\chi^2(2) = 4.96$, $p = .084$, indicating that it was not an artifact of the smaller common sample.

Across all models, systematic differences between schools remained substantial, with School 1 consistently showing higher levels of MA than the remaining schools (ps $< .01$). Class-level variance was minimal (ICC = .014–.024). Relative to the unconditional class model, the inclusion of predictors reduced between-class variance by approximately 98%, indicating that most of the initial clustering reflected systematic differences between schools rather than independent class effects. In contrast, reductions in residual variance were moderate (23–31% across models).

Post-hoc Power Analysis

We conducted a post-hoc power analysis for each model as estimated (REML, with school fixed and a class random intercept), computing for every fixed effect its achieved power and the minimum effect detectable with 80% power; 95% CIs for the power estimates were obtained by parametric bootstrap (2,000 simulations). The omnibus blocks were well powered throughout: both the substantive predictors and the school factor were detected with power ≥ .99 in every model. Power at the level of individual predictors was more limited. The design could detect continuous partial slopes of approximately β = .16–.20 with 80% power in Models 1–3, rising to about β = .19–.27 in Model 4. Against these thresholds, only two associations were adequately powered at their observed magnitude: test anxiety, reliably detectable in Models 2 and 3 (power = .97 and .93) and near criterion in Model 4 (.72), and intellectual helplessness in math in Model 1 (power ≥ .99). The power to detect intellectual helplessness fell to .43–.51 once anxiety measures were added and to .16 in Model 4, paralleling the attenuation of its coefficient, while fluid intelligence was moderately powered in Model 3 (.55). For the remaining predictors (gender, age, math achievement, general anxiety, working memory, and the maternal variables), power did not exceed .41 (range .05–.41), indicating that their non-significant coefficients are inconclusive rather than evidence of negligible effects. Finally, because between-class variance was minimal (ICC = .014–.024), the test of the class random intercept had negligible power (≈ .06–.10).

Discussion

Although numerous individual and environmental correlates of MA have been identified (e.g., Barroso et al., 2021; Dowker et al., 2016; Dowker, 2021; Girard et al., 2021; LeFevre et al., 2009; Namkung et al., 2019; O'Hara et al., 2022), relatively few studies have examined a range of potential predictors simultaneously (e.g., del Río et al., 2025; Leppma & Darrah, 2024; Lau et al., 2022; Jameson, 2014; Meece et al., 1990; Szczygieł et al., 2024, 2025). Consequently, relatively little is known about which predictors make the strongest and

most stable contribution to MA when considered together, and how much variance is explained by different groups of predictors. Therefore, the present study addressed the following research questions: (1) which individual and environmental variables predict MA when other theoretically relevant variables are taken into account and which predictors show the strongest and most stable associations across successive models; (2) what proportion of variance in MA is accounted for by the predictors included in the successive models and by differences between schools and classes.

Answering the first research question, we found that test anxiety was the only stable predictor of MA across models. Intellectual helplessness in math was a relatively stable predictor, although its statistical significance varied across subsequent models, suggesting that its predictive contribution may be sensitive to missing data and sample size. Fluid intelligence appeared to be an even less stable predictor, as variation in confidence intervals and *p*-values across models led to divergent conclusions. Math achievement emerged as a significant predictor in only one sensitivity-analysis model, indicating that its predictive contribution was not stable across models. Other factors (gender, age, general anxiety, working memory, and mothers' MA, education level, and home numeracy activities) were not significant predictors in any model.

Answering the second research question, the models accounted for a substantial proportion of variance in MA, with the included individual and family-related predictors together with school membership explaining approximately 58–63% of the variance, and class-level differences increasing the total explained variance only slightly, to approximately 59–64%. When the successive models were compared on the same sample (Supplementary Materials, Table E), the model that included sociodemographic and math domain-specific predictors together with school membership explained 59.6% of the variance in MA. Adding domain-general affective predictors increased the explained variance to 62.5% and

significantly improved model fit. Adding domain-general cognitive predictors produced only a small further increase to 63.9% and did not significantly improve model fit. Finally, adding mother-related predictors resulted in almost no further increase in explained variance (64%) and did not improve model fit. Thus, the affective block was the only subsequent block that made a significant additional contribution, whereas the cognitive and mother-related variable blocks added little explanatory value.

Predictors of MA

Sociodemographic variables. Gender showed a small positive zero-order relationship to MA, indicating higher MA among girls, but it did not make a significant unique predictive contribution in any of the regression models. This pattern is consistent with previous evidence suggesting that gender differences in MA are relatively small (Eidlin-Levy et al., 2023; OECD, 2013) and may diminish when other relevant individual characteristics are considered simultaneously (Delage et al., 2022; Ganley & McGraw, 2016; Szczygieł, 2020a; Szczygieł et al., 2025). At the same time, it differs from studies in which gender remained a small but significant predictor after controlling for other variables (Lau et al., 2022; Leppma & Darrah, 2024; Szczygieł et al., 2024, 2025). Importantly, gender was already nonsignificant in Model 1, before general and test anxiety were introduced. One possible explanation is that the zero-order gender difference in MA partly reflected differences between schools, as school membership was already included in Model 1 and accounted for substantial variability in MA. However, the present analyses do not allow us to determine whether this reflects differences in how girls and boys experience mathematics across schools. Gender was also unrelated to math achievement, suggesting that the higher MA reported by girls was not simply accompanied by lower mathematics performance. Future studies should examine whether gender differences in MA vary across school contexts and whether school-level factors interact with gender in predicting MA.

Age was neither significantly correlated with MA nor a significant predictor in any model. This is consistent with previous findings among adults (Szczygieł et al., 2025). Given the relatively narrow age range of the sample, substantial age-related differences in MA may simply have been unlikely to emerge. Rather, MA is thought to develop gradually over the course of children's educational experiences, and age differences are therefore more likely to emerge across broader developmental periods than within a single grade (Dowker et al., 2016).

Math Domain-Specific Factors. Contrary to our expectations, math achievement was neither significantly correlated with MA nor a significant predictor of MA in any of the main models, despite extensive evidence of a negative correlation between the two (Barroso et al., 2021; Namkung et al., 2019; Pelegrina et al., 2026). However, although it did not reach statistical significance, the zero-order correlation observed in the present study was negative ($r = -.15$) and comparable in magnitude to associations reported for children in Grades 3–5 (Barroso et al., 2021; $r = -.20$, 95% CI [−.25, −.14]). Nevertheless, the result is consistent with previous research in a Polish sample showing that MA is only very weakly related to math achievement in children from the first and second grades (Szczygieł, 2020a). The lack of a unique predictive effect is also consistent with some previous findings in young children and adolescents. Krinzinger et al. (2009) found no significant longitudinal predictive relationships between MA and calculation ability among children followed from the end of Grade 1 to the middle of Grade 3. In students in Grades 7–9, prior math grades did not directly predict MA once students' perceptions of their mathematical ability were considered (Meece et al., 1990), whereas among secondary-school students, math performance initially predicted MA but became nonsignificant after general and test anxiety were added to the model (Szczygieł et al., 2025). By contrast, longitudinal evidence from younger primary-school children indicates that math achievement can predict subsequent MA even after earlier MA and other individual characteristics are controlled for (Szczygieł et al., 2024), and math performance remained a

significant predictor in adult samples (Szczygieł et al., 2025). Thus, the independent predictive role of math achievement appears to vary by developmental stage and by other variables included in the model. The relationship between MA and math achievement may also depend on the specific mathematical skills assessed (Wu et al., 2012).

Taken together, these findings indicate that, in the present sample, math achievement was not directly associated with MA. However, this should not be interpreted as evidence that math achievement is unimportant for the development of MA, particularly given the cross-sectional design of the study. Notably, math achievement was moderately and negatively associated with test anxiety, raising the possibility of an indirect association between math achievement and MA through test anxiety, which would need to be tested in future longitudinal research. This interpretation aligns with evidence from younger children showing that math self-concept plays a stronger role than objective math performance in the development of MA (Jameson, 2014). Intellectual helplessness in math showed a moderate positive zero-order correlation with MA and was a significant predictor in Model 1, with a small to moderate effect. However, its predictive contribution weakened substantially after general and test anxiety were introduced and was not stable across subsequent models and sensitivity analyses. This is consistent with the substantial overlap between intellectual helplessness in math and test anxiety observed in the present study: intellectual helplessness in math was strongly correlated with test anxiety, suggesting that the two constructs share considerable variance. Conceptually, both constructs involve negative responses to academic challenges, although test anxiety reflects distress in evaluative situations, whereas intellectual helplessness in math reflects low perceived control and the expectation that further cognitive effort will be ineffective (Fincham et al., 1989; Gürefe & Bakalım, 2018). The present findings differ somewhat from previous regression results in older samples. Among secondary-school students, intellectual helplessness in math was the strongest predictor of MA, with a small to moderate effect, before

general and test anxiety were introduced. Once these affective variables were included, it remained the second strongest predictor, behind test anxiety, with a moderate effect (Szczygieł et al., 2025). In adults, intellectual helplessness in math likewise had a small to moderate effect and remained one of the strongest predictors of MA when sociodemographic, affective, and cognitive variables were considered simultaneously (Szczygieł et al., 2025). Thus, although intellectual helplessness in math was clearly associated with MA in the present sample, its unique predictive contribution appeared less stable than that reported in adolescents and adults. The substantial overlap between intellectual helplessness and test anxiety may therefore partly explain why the former contributed less unique variance once affective predictors were included. However, these comparisons should be interpreted cautiously because the studies differed in the predictors included in their models (e.g., mathematical resilience was included in Szczygieł et al., 2025), which limits conclusions about developmental differences in the predictive role of intellectual helplessness. Overall, math performance appears to be more cognitive in nature, whereas intellectual helplessness in math seems to have a stronger affective component.

Domain-General Affective Factors. Both general and test anxiety were weakly to moderately and positively correlated with MA, although the association was stronger for test anxiety. In the regression models, however, only test anxiety remained a small but significant predictor, whereas general anxiety no longer explained unique variance in MA. This pattern is partly consistent with previous findings. In secondary-school students, both general and test anxiety remained significant when considered simultaneously, but test anxiety was considerably stronger than general anxiety (Szczygieł et al., 2025). Importantly, general and test anxiety were also substantially correlated in that study, indicating that they share considerable variance while retaining some unique associations with MA. In younger primary-school children, general anxiety was a relatively strong predictor of MA at school entry

(Szczygieł et al., 2024); however, because test anxiety was not included, it is unclear how much of this association reflected variance unique to general anxiety rather than variance shared with more evaluation-specific anxiety. In the present study, the stronger association of test anxiety with MA, combined with the overlap between general and test anxiety, may explain why general anxiety contributed little unique variance once test anxiety was included. Shared mechanisms may include intrusive thoughts, heightened physiological arousal, and avoidance of evaluation (Cassady, 2023). Accordingly, once test anxiety is taken into account, broader general anxiety may contribute relatively little unique variance to MA.

Domain-General Cognitive Factors. Neither fluid intelligence nor verbal working memory was significantly correlated with MA. Fluid intelligence nevertheless emerged as a very weak but significant negative predictor of MA in Model 3, consistent with previous studies reporting small negative associations between cognitive abilities and MA (Schillinger et al., 2018; Živković et al., 2023). However, this effect was not stable across models: it was no longer significant in Model 4 and was also nonsignificant in the complete-case analysis, although it re-emerged in the imputed-data analysis. Thus, the present findings provide very limited evidence for an independent contribution of fluid intelligence to MA.

Working memory was not a significant predictor in any model, consistent with previous studies in primary-school children, adolescents, and adults (Szczygieł et al., 2024, 2025). This does not necessarily imply that working memory is irrelevant to MA-related processes. Rather, working memory may be particularly important when individuals perform mathematics tasks, as anxiety-related intrusive thoughts can consume cognitive resources needed for successful math performance. Consistent with this view, Ramirez et al. (2013) showed that working memory moderated the association between MA and math performance, with a stronger negative MA-performance relation among children with higher working memory capacity. Thus, working memory may influence the extent to which MA interferes with mathematical

performance rather than directly contributing to individual differences in MA itself. The present findings therefore suggest that domain-general cognitive abilities show little direct association with MA once domain-general affective and domain-specific factors are controlled.

Family Factors. Mothers' MA was weakly positively correlated with children's MA, whereas mothers' home numeracy activities and education were not significantly correlated with MA. None of the mothers' characteristics significantly predicted children's MA. This pattern is broadly consistent with mixed findings from previous studies. In primary school children, Jameson (2014) found that the frequency of home math activities positively predicted MA when child and family characteristics were considered together, although adding environmental variables did not significantly increase the variance explained in MA. Szczygieł (2020b) found predictive effects of parental MA, but these varied depending on the parent, the child's gender, and grade level. Vanbinst et al. (2020) found that parental education was no longer significant when mothers' and fathers' MA and education were considered simultaneously, and only mothers' MA approached significance. By contrast, del Río et al. (2025) found that advanced home math activities predicted lower MA and parents' negative attitudes toward mathematics predicted higher MA at the beginning of primary education, while Lau et al. (2022) reported significant associations of several family characteristics with MA in an international sample of primary school students even when individual, family, and school characteristics were considered simultaneously. Longitudinal evidence is similarly mixed: parental MA measured when children were in preschool did not predict children's MA in primary school (Tomasetto et al., 2025), whereas parental MA predicted subsequent MA partly through children's math learning involvement in primary school (Zhu et al., 2026). Previous studies on family-related predictors of MA have not always controlled for a broad set of child-level characteristics, which may contribute to differences in the estimated predictive role of family factors across studies. Associations between parental characteristics and

children's MA may partially reflect child-level differences, such as general and test anxiety, cognitive abilities, or experiences in math learning. When these characteristics are not considered simultaneously, apparent predictive effects of family variables may partly reflect variance shared with child-level characteristics that are more directly associated with MA. By including both individual and family predictors in the same models, the present study aimed to assess the unique contribution of family-related factors beyond child-level characteristics. Unfortunately, limited availability of parental data in our study was associated with a decrease in the number of participants in the model including mothers, thereby reducing statistical power. Therefore, our findings regarding family-related factors should be considered preliminary.

Variance in MA explained by Predictor Blocks and School- and Class-Level Differences

The previous sections focused on the unique contribution of MA predictors. However, considering predictors separately does not indicate whether broader theoretically defined groups of variables contribute meaningfully to explaining MA after taking into account previous sets of predictors. The comparison of successive models therefore provides a complementary perspective on the importance of sociodemographic and math domain-specific, affective, cognitive, and family-related factors.

Because the main models were estimated on somewhat different samples, differences in explained variance between them cannot be interpreted directly as incremental effects of the successive blocks. The formal comparison conducted on the same complete-case sample provides a more appropriate basis for this interpretation (see Supplementary Materials, Table E). Model 1, including sociodemographic and math domain-specific predictors together with school membership, accounted for 59.6% of the variance in MA. Adding general and test anxiety increased the explained variance to 62.5% and significantly improved model fit. The addition of fluid intelligence and verbal working memory produced only a small further

increase to 63.9%, and did not significantly improve model fit, whereas adding mothers' MA, home numeracy activities, and education resulted in almost no further increase, with 64% of the variance explained, and likewise did not improve model fit. Thus, the block-level findings complement the results for the individual predictors. The affective block provided the clearest additional explanatory contribution, consistent with test anxiety emerging as the most stable individual predictor, whereas the cognitive and family-related blocks contributed little additional explanatory value once the preceding variables had been considered. Importantly, this distinction also helps interpret the less stable effect of fluid intelligence: although intelligence reached significance in some model specifications, the cognitive block as a whole did not significantly improve the model. These findings should be interpreted in light of the specified order of the blocks. The incremental contribution of a given block reflects the variance it explained beyond variables entered earlier in this particular sequence and should not be interpreted as an absolute estimate of its importance. Moreover, adding predictors in successive blocks does not capture possible interactions between factors emphasized in ecological models (Bronfenbrenner & Morris, 2007).

The overall proportion of variance accounted for in the present study was relatively high and similar to that reported in previous studies examining a broad range of MA predictors. However, direct comparisons between studies require caution because they differed considerably in educational stage, sets of predictors, and statistical approaches. For example, parental characteristics accounted for 7.9% of the variance in MA in Grade 6 children in Vanbinst et al. (2020), whereas individual characteristics explained approximately 27% in undergraduate students in Leppma and Darrah (2024). Broader models have generally accounted for larger proportions of variance. In primary-school children, Szczygieł et al. (2024) explained approximately 29% of the variance in MA at school entry, 46% at the end of Grade 1, and 62% at the end of Grade 2, although the later models also included prior MA and, at the

final measurement, prior math achievement. In secondary-school students, Meece et al. (1990) accounted for approximately half of the variance in MA. Similarly, among secondary-school students, Szczygieł et al. (2025) explained 46% of the variance before general and test anxiety were introduced and 59% after these affective variables were added. Thus, the approximately 60% of variance accounted for in the present study is comparable to the upper range reported in previous broader models of MA. However, this relatively high proportion should not be interpreted as variance explained by the measured child- and family-level characteristics alone, because school membership was included as a fixed categorical effect in all models.

The variance decomposition indicates that school membership accounted for an important part of the structure of MA in the present sample. Before individual predictors were introduced, approximately 33% of the variance in MA was associated with differences between schools. When classes were considered without explicitly accounting for school membership, approximately 45% of the variance appeared to occur between classes. However, once school membership was included as a fixed effect, the remaining between-class variance decreased to only about 1%. Similarly, in the main regression models, only approximately 1.4–2.4% of the remaining variance was attributable to differences between classes. This pattern indicates that much of the variation that initially appeared to occur between classes reflected systematic differences between the schools to which these classes belonged rather than substantial independent differences between classes within the same schools. Accordingly, the approximately 98% reduction in between-class variance after the fixed effects were included should not be interpreted as an effect of school alone, because school membership and the individual predictors were modeled jointly.

The substantial differences between schools are consistent with previous observations that MA varies across educational contexts. Lau et al. (2022), for example, found that approximately 3.6% of the residual variance in MA remained at the educational-environment

level after a broad set of individual, family, classroom, and teacher characteristics had been considered. Longitudinal studies have also shown that perceived teacher support and other characteristics of the school environment may predict subsequent MA or changes in MA over time (Li et al., 2021, 2023; Zhuo et al., 2025). However, Lau et al.'s estimate is not directly comparable with the approximately 33% between-school variation observed in the present unconditional analyses. Their estimate concerned residual contextual variance after measured predictors had already been taken into account, whereas our estimate reflected variation associated with school membership before specific school characteristics were modeled.

The present findings therefore indicate that children's MA differed substantially between schools, but they do not identify the source of these differences. School membership was represented by categorical fixed effects rather than by direct measures of school characteristics. Consequently, the observed between-school variation may reflect differences in instructional practices, teacher characteristics, school climate, or peer norms, but it may also reflect differences in student characteristics and family backgrounds across schools or other unmeasured factors. This distinction is particularly important because one school showed substantially higher MA than the remaining schools and the total number of schools was small. Future studies including larger numbers of schools and direct measures of educational characteristics are needed to determine which features of the school environment account for between-school differences in MA.

Educational Implications

Our findings underscore the importance of the close association between test-related anxiety and MA in early schooling. However, because the present study was cross-sectional, the direction of this relationship cannot be determined. While test anxiety may be associated with heightened anxiety in mathematical contexts, the reverse interpretation is also plausible: children who experience elevated MA may become particularly concerned about evaluative

situations involving mathematics. Therefore, interventions aimed at reducing either test anxiety or MA could have broader benefits for students' emotional experiences of learning and assessment. Educational practices such as shifting the focus from summative to formative assessment, incorporating more low-stakes tests, and explicitly teaching coping and self-regulation strategies may help children better manage evaluative stress (Sammallahti et al., 2023). Although intellectual helplessness in math was not a consistent predictor across all models, its association with MA in some analyses suggests that children's beliefs about their ability to cope with mathematical challenges may also be important. Educational practices that provide opportunities for success and help students interpret difficulties as a normal part of learning may therefore contribute to reducing negative emotional experiences related to math.

The substantial variance attributable to school-level factors suggests that characteristics of the school environment may play an important role in either alleviating or exacerbating MA. Supportive environments where mistakes are normalized and framed as learning opportunities may reduce the emotional burden of assessment (Jameson, 2014). Although the present study did not identify the specific school-level mechanisms responsible for the observed differences between schools, the findings suggest that the school context deserves greater attention in future research.

Limitations and Future Directions

First, our study was correlational and cross-sectional, with a short interval between measurements. Therefore, the regression models should not be interpreted as evidence of causal ordering. We tested children in the fourth grade of primary school, a potentially stressful educational stage for many students in Poland, as it follows the transition from integrated education (with a generalist teacher) to subject-specific instruction. However, our study did not directly examine the transition from third to fourth grade, as students were tested only in fourth grade. In future studies, a longitudinal design following students from third to fourth grade

would be valuable for examining how this transition relates to changes in MA and its determinants. Moreover, the origins of MA are increasingly being investigated as early as the preschool years (Petronzi et al., 2019; Svraka et al., 2024); therefore, future longitudinal studies examining the factors contributing to MA should begin in preschool and follow children over several years.

Second, the analysis of parental MA predictors relied only on mothers' data, due to an insufficient amount of fathers' data stemming from convenience sampling. Future studies should include parallel data collection from both parents to determine their individual association with MA. It is also important to ensure sufficiently large samples in educational research. When analyzing cohorts, collecting data from a large number of students may be challenging; however, small samples can be particularly problematic in regression analyses that rely on listwise deletion, as this approach reduces statistical power and may weaken the robustness of the results, as illustrated by our sensitivity analyses.

Third, we included only a few schools and did not measure any variables at the school level, even though teaching methods and school characteristics may be important predictors of MA (Dowker et al., 2016; Dowker, 2021; Girard et al., 2021; Jameson, 2014; LeFevre et al., 2009; O'Hara et al., 2022). The significant school effect compared to the class effect may reflect the fact that a given school often has only one mathematics teacher, whose teaching style and related factors are shared across all students in that school. In future studies, it will therefore be necessary to control not only for the class level but also for clustering at the level of the mathematics teacher. Future research should include more schools with more classes and predictors of MA at both levels, to better understand how environmental factors relate to MA. Moreover, our study included only variables at the individual level (children's personal characteristics) and the microsystem level (family- and school-related factors). It can be expected that some of the variance in children's MA may be attributable to broader contextual

factors that cannot be easily examined or disentangled, as these influences are inherently interconnected.

Fourth, our sample size was at the lower limit generally considered acceptable for multilevel modeling. Nevertheless, the post hoc power analysis indicated that the study retained adequate sensitivity to detect effects of approximately $\beta = .16–.27$ with 80% power, depending on the model. Therefore, while effects of this magnitude were expected to be detectable, smaller effects may have remained undetected and our results should be interpreted with caution.

Finally, our approach to controlling for school effects may have reduced our ability to detect associations between children's MA and family-related predictors. Because school membership was represented by fixed-effect dummy variables, between-school variation in characteristics such as maternal education, maternal MA, and home numeracy activities may have been absorbed by the school coefficients. This possibility is particularly relevant given that parents of children attending the same school are likely to be more similar to one another in terms of educational and family-related characteristics than are parents from other schools. Consequently, the non-significant effects observed for these family variables should be interpreted with caution, as they may reflect either truly weak associations with children's MA or limited within-school variability for estimating such effects. Future research based on larger samples should employ multilevel models that explicitly disentangle within-school and between-school sources of variation, thereby providing a more precise assessment of associations between family characteristics and children's MA.

Conclusions

This study indicates that, among the variables examined, test anxiety is the key predictor of MA in primary school children, whereas evidence for independent contributions of other individual and family-related variables is less conclusive once this anxiety is

considered. Moreover, substantial between-school variability in MA suggests that interventions should not only target individual children but also address school practices. Overall, these findings underscore the need for integrated approaches that combine evidence-based techniques for reducing test anxiety (e.g., relaxation training, coping strategies) with systemic practices that promote supportive, low-stress school environments (e.g., using formative assessment and gradually increasing task difficulty).

References

Ahmed, W., Minnaert, A., Kuyper, H., & Van Der Werf, G. (2012). Reciprocal relationships between math self-concept and math anxiety. *Learning and Individual Differences*, *22*(3), 385–389. https://doi.org/10.1016/j.lindif.2011.12.004

Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General*, *130*(2), 224–237. https://doi.org/10.1037/0096-3445.130.2.224

Baddeley, A. (1992). Working memory. *Science*, 255(5044), 556–559. https://doi.org/10.1126/science.1736359

Barroso, C., Ganley, C. M., McGraw, A. L., Geer, E. A., Hart, S. A., & Daucourt, M. C. (2021). A meta-analysis of the relation between math anxiety and math achievement. *Psychological Bulletin*, *147*(2), 134–168. https://doi.org/10.1037/bul0000307

Brocki, K. C., & Bohlin, G. (2004). Executive functions in children aged 6 to 13: A dimensional and developmental study. *Developmental Neuropsychology*, *26*(2), 571–593. https://doi.org/10.1207/s15326942dn2602_3

Bronfenbrenner, U., & Morris, P. A. (2007). The bioecological model of human development. In R. M. Lerner (Ed.), *Handbook of child psychology: Vol. 1: Theoretical models of human development* (6th ed., pp. 793–828). Wiley. https://doi.org/10.1002/9780470147658.chpsy0114

Carey, E., Devine, A., Hill, F., & Szűcs, D. (2017). Differentiating anxiety forms and their role in academic performance from primary to secondary school. *PLoS ONE*, *12*(3), e0174418. https://doi.org/10.1371/journal.pone.0174418

Cassady, J. C. (2023). Cognitive Test Anxiety Scale. In C. U. Krägeloh, M. Alyami, & O. N. Medvedev (Eds), *International handbook of behavioral health assessment*. Springer. https://doi.org/10.1007/978-3-030-89738-3_51-1

Caviola, S., Toffalini, E., Giofrè, D., Ruiz, J. M., Szűcs, D., & Mammarella, I. C. (2022). Math performance and academic anxiety forms, from sociodemographic to cognitive aspects: A meta-analysis on 906,311 participants. *Educational Psychology Review*, *34*, 363–399. https://doi.org/10.1007/s10648-021-09618-5

Delage, V., Trudel, G., Retanal, F., & Maloney, E. A. (2022). Spatial anxiety and spatial ability: Mediators of gender differences in math anxiety. *Journal of Experimental Psychology: General*, *151*(4), 921–933. https://doi.org/10.1037/xge0000884

del Río, M. F., Morales, M. F., Susperreguy, M. I., & Espinoza, A. M. (2025). Predictors and outcomes associated with early math anxiety: Parental factors, working memory and math performance. *Cognitive Development*, *75*, 101613. https://doi.org/10.1016/j.cogdev.2025.101613

Dowker, A. (2021). Home numeracy and preschool children's mathematical development: Expanding home numeracy models to include parental attitudes and emotions. *Frontiers in Education*, *6*(575664). https://doi.org/10.3389/feduc.2021.575664

Dowker, A. (2026). Mathematics anxiety: Effects of age, gender and culture. *British Journal of Educational Psychology*, *96*(2), 896–921. https://doi.org/10.1111/bjep.70057

Dowker, A., Sarkar, A., & Looi, C. Y. (2016). Mathematics anxiety: What have we learned in 60 years? *Frontiers in Psychology*, *7*(508). https://doi.org/10.3389/fpsyg.2016.00508

Eidlin-Levy, H., Avraham, E., Fares, L., & Rubinsten, O. (2023). Math anxiety affects career choices during development. *International Journal of STEM Education*, *10*, 49. https://doi.org/10.1186/s40594-023-00441-8

Field, A. P., Evans, D., Bloniewski, T., & Kovas, Y. (2019). Predicting maths anxiety from mathematical achievement across the transition from primary to secondary education. *Royal Society Open Science*, *6*(11), 191459. https://doi.org/10.1098/rsos.191459

Fincham, F. D., Hokoda, A., & Sanders, R. (1989). Learned helplessness, test anxiety, and academic achievement: A longitudinal analysis. *Child Development*, *60*(1), 138–145. https://doi.org/10.2307/1131079

Ganley, C. M., & McGraw, A. L. (2016). The development and validation of a Revised Version of the Math Anxiety Scale for Young Children. *Frontiers in Psychology*, *7*, 1181. https://doi.org/10.3389/fpsyg.2016.01181

Girard, C., Bastelica, T., Léone, J., Epinat-Duclos, J., Longo, L., & Prado, J. (2021). The relation between home numeracy practices and a variety of math skills in elementary school children. *PLOS ONE*, *16*(9), e0255400. https://doi.org/10.1371/journal.pone.0255400

Gürefe, N., & Bakalım, O. (2018). Mathematics anxiety, perceived mathematics self-efficacy and learned helplessness in mathematics in faculty of education students. *International Online Journal of Educational Sciences*, *10*(3), 154–166. https://doi.org/10.15345/iojes.2018.03.010

Hamre, B. K., & Pianta, R. C. (2001). Early teacher–child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development*, *72*(2), 625–638. https://doi.org/10.1111/1467-8624.00301

Hirshfeld-Becker, D. R., Micco, J. A., Simoes, N. A., & Henin, A. (2008). High risk studies and developmental antecedents of anxiety disorders. *American Journal of Medical*

Genetics Part C: Seminars in Medical Genetics, *148C*(2), 99–117. https://doi.org/10.1002/ajmg.c.30170

Hwang, J. (2019). Relationships among locus of control, learned helpless, and mathematical literacy in PISA 2012: Focus on Korea and Finland. *Large-Scale Assessments in Education*, *7*(4). https://doi.org/10.1186/s40536-019-0072-7

Jameson, M. M. (2014). Contextual factors related to math anxiety in second-grade children. *Journal of Experimental Education*, *82*(4), 518–536. https://doi.org/10.1080/00220973.2013.813367

Jaworowska, A., & Szustrowa, T. (2000). *TMS-K - Test matryc Ravena w wersji standard - forma klasyczna*. PTP.

Jerković, I., & Milovanović, I. (2020). Medijatorski efekti matematičke anksioznosti u relaciji između očeve uključenosti u podučavanje i matematičkog postignuća učenika ranog osnovnoškolskog uzrasta. *Psihološka Istraživanja*, *23*(1), 43–61. https://doi.org/10.5937/psistra23-24633

Karpińska, E., & Radtke, B. M. (2019). *Zestaw sprawdzianów umiejętności szkolnych dla poradnictwa psychologiczno-pedagogicznego.* PTP.

Kent, P. (2017). Fluid intelligence: A brief history. *Applied Neuropsychology: Child*, *6*(3), 193–203. https://doi.org/10.1080/21622965.2017.1317480

Krinzinger, H., Kaufmann, L., & Willmes, K. (2009). Math anxiety and math ability in early primary school years. *Journal of Psychoeducational Assessment*, *27*, 206–225. https://doi.org/10.1177/0734282908330583

Lau, N. T. T., Ansari, D., & Sokolowski, H. M. (2024). Unraveling the interplay between math anxiety and math achievement. *Trends in Cognitive Sciences*, *28*(10), 937–947. https://doi.org/10.1016/j.tics.2024.07.006

Lau, N. T. T., Hawes, Z., Tremblay, P., & Ansari, D. (2022). Disentangling the individual and contextual effects of math anxiety: A global perspective. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(7), e2115855119. https://doi.org/10.1073/pnas.2115855119

Lavelle-Hill, R., Frenzel, A. C., Goetz, T., Lichtenfeld, S., Marsh, H. W., Pekrun, R., Sakaki, M., Smith, G., & Murayama, K. (2024). How the predictors of math achievement change over time: A longitudinal machine learning approach. *Journal of Educational Psychology*, *116*(8), 1383–1403. https://doi.org/10.1037/edu0000863

LeFevre, J.-A., Skwarchuk, S.-L., Smith-Chant, B. L., Fast, L., Kamawar, D., & Bisanz, J. (2009). Home numeracy experiences and children's math performance in the early school years. *Canadian Journal of Behavioural Science*, *41*(2), 55–66. https://doi.org/10.1037/a0014532

Leppma, M., & Darrah, M. (2024). Self-efficacy, mindfulness, and self-compassion as predictors of math anxiety in undergraduate students. *International Journal of Mathematical Education in Science and Technology*, *55*(4), 883–898. https://doi.org/10.1080/0020739x.2022.2054740

Li, H., Zhang, A., Zhang, M., Huang, B., Zhao, X., Gao, J., & Si, J. (2021). Concurrent and longitudinal associations between parental educational involvement, teacher support, and math anxiety: The role of math learning involvement in elementary school children. *Contemporary Educational Psychology*, *66*, 101984. https://doi.org/10.1016/j.cedpsych.2021.101984

Li, H., Zhang, M., Hou, S., Huang, B., Xu, C., Li, Z., & Si, J. (2023). Examining the dynamic links among perceived teacher support, mathematics learning engagement, and dimensions of mathematics anxiety in elementary school students: A Four-wave longitudinal study.

Contemporary Educational Psychology, *75*, 102211. https://doi.org/10.1016/j.cedpsych.2023.102211

Luttenberger, S., Wimmer, S., & Paechter, M. (2018). Spotlight on math anxiety. *Psychology Research and Behavior Management*, *11*, 311–322. https://doi.org/10.2147/prbm.s141421

Maloney, E. A., Ansari, D., & Fugelsang, J. A. (2011). Rapid communication: The effect of mathematics anxiety on the processing of numerical magnitude. *Quarterly Journal of Experimental Psychology*, *64*(1), 10–16. https://doi.org/10.1080/17470218.2010.533278

Maloney, E. A., Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2015). Intergenerational effects of parents' math anxiety on children's math achievement and anxiety. *Psychological Science*, *26*(9), 1480–1488. https://doi.org/10.1177/0956797615592630

Maloney, E. A., Schaeffer, M. W., & Beilock, S. L. (2013). Mathematics anxiety and stereotype threat: Shared mechanisms, negative consequences and promising interventions. *Research in Mathematics Education*, *15*(2), 115–128. https://doi.org/10.1080/14794802.2013.797744

Mandler, G., & Sarason, S. B. (1952). A study of anxiety and learning. *The Journal of Abnormal and Social Psychology*, *47*(2), 166–173. https://doi.org/10.1037/h0062855

Marteau, T. M., & Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology*, *31*(3), 301–306. https://doi.org/10.1111/j.2044-8260.1992.tb00997.x

Matczak, A., Piotrowska, A., & Ciarkowska, W. (1991). *WISC: Skala Inteligencji Wechslera dla Dzieci*. PTP.

McLean, L., Janssen, J., Espinoza, P., Lindstrom Johnson, S., & Jimenez, M. (2023). Associations between teacher and student mathematics, science, and literacy anxiety in

fourth grade. *Journal of Educational Psychology*, *115*(4), 539–551. https://doi.org/10.1037/edu0000790

Meece, J. L., Wigfield, A., & Eccles, J. S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of Educational Psychology*, *82*(1), 60–70. https://doi.org/10.1037/0022-0663.82.1.60

Namkung, J. M., Goodrich, J. M., & Lee, K. (2023). The factor structure of mathematics anxiety and its relation to gender and mathematics performance. *Psychology in the Schools*, *60*(11), 4740–4757. https://doi.org/10.1002/pits.23016

Namkung, J. M., Peng, P., & Lin, X. (2019). The relation between mathematics anxiety and mathematics performance among school-aged students: A meta-analysis. *Review of Educational Research*, *89*(3), 459–496. https://doi.org/10.3102/0034654319843494

OECD. (2013). Chapter 4: Mathematics anxiety. In *PISA 2012 Results. Volume III* (pp. 98–112). OECD.

OECD. (2024). *An evolution of mathematics curriculum: Where it was, where it stands and where it is going.* OECD Publishing. https://doi.org/10.1787/0ffd89d0-en

O'Hara, G., Kennedy, H., Naoufal, M., & Montreuil, T. (2022). The role of the classroom learning environment in students' mathematics anxiety: A scoping review. *British Journal of Educational Psychology*, *92*(4), 1458–1486. https://doi.org/10.1111/bjep.12510

Pelegrina, S., Toffalini, E., & Martín-Puga, M. E. (2026). Stability of math anxiety and its relation with math performance over time: A meta-analysis of longitudinal studies. *Journal of Educational Psychology*, *118*(5), 677–702. https://doi.org/10.1037/edu0001008

Pellizzoni, S., Cargnelutti, E., Cuder, A., & Passolunghi, M. C. (2022). The interplay between math anxiety and working memory on math performance: A longitudinal study. *Annals of the New York Academy of Sciences, 1510*(1), 132–144. https://doi.org/10.1111/nyas.14722

Petronzi, D., Staples, P., Sheffield, D., Hunt, T., & Fitton- Wilde, S. (2019). Further development of the Children's mathematics anxiety scale UK (CMAS- UK) for ages 4-7 years. *Educational Studies in Mathematics, 100*, 231–249. https://doi.org/10.1007/s10649-018-9860-1

Posit Team. (2025). *RStudio: Integrated Development Environment for R* [Computer software]. Posit Software, PBC. http://www.posit.co/

Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2013). Math anxiety, working memory, and math achievement in early elementary school. *Journal of Cognition and Development*, *14*(2), 187–202. https://doi.org/10.1080/15248372.2012.664593

Raven, J. C., & Court, J. H. (2000). *Standard progressive matrices*. Psychology Press.

Retanal, F., & Maloney, E. A. (2026). Individual, interpersonal and sociocultural contributors to maths anxiety in children. *Nature Reviews Psychology*, *5*(2), 93–104. https://doi.org/10.1038/s44159-025-00519-w

Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin*, *135*(6), 943–973. https://doi.org/10.1037/a0017327

Richardson, F. C., & Suinn, R. M. (1972). The Mathematics Anxiety Rating Scale: Psychometric data. *Journal of Counseling Psychology*, *19*(6), 551–554. https://doi.org/10.1037/h0033456

Sammallahti, E., Finell, J., Jonsson, B., & Korhonen, J. (2023). A meta-analysis of math anxiety interventions. *Journal of Numerical Cognition*, *9*(2), 346–362. https://doi.org/10.5964/jnc.8401

Schillinger, F. L., Vogel, S. E., Diedrich, J., & Grabner, R. H. (2018). Math anxiety, intelligence, and performance in mathematics: Insights from the German adaptation of the

Abbreviated Math Anxiety Scale (AMAS-G). *Learning and Individual Differences*, *61*, 109–119. https://doi.org/10.1016/j.lindif.2017.11.014

Sędek, G., & McIntosh, D. N. (1998). Intellectual helplessness: Domain specificity, teaching styles, and school achievement. In M. Kofta, G. Weary, & G. Sędek (Eds.), *Personal control in action: Cognitive and motivational mechanisms* (pp. 419–443). Springer. https://doi.org/10.1007/978-1-4757-2901-6_17

Seligman, M. E. P. (1972). Learned helplessness. *Annual Review of Medicine*, *23,* 407–412. https://doi.org/10.1146/annurev.me.23.020172.002203

Soni, A., & Kumari, S. (2017). The role of parental math anxiety and math attitude in their children's math achievement. *International Journal of Science and Mathematics Education*, *15*(2), 331–347. https://doi.org/10.1007/s10763-015-9687-5

Sorvo, R., Koponen, T., Viholainen, H., Aro, T., Räikkönen, E., Peura, P., Tolvanen, A., & Aro, M. (2019). Development of math anxiety and its longitudinal relationships with arithmetic achievement among primary school children. *Learning and Individual Differences, 69*, 173–181. https://doi.org/10.1016/j.lindif.2018.12.005

Spielberger, C. D. (1983). *Manual for the State-Trait Anxiety Inventory STAI (Form Y)*. Mind Garden.

Szczygieł, M. (2019). How to measure math anxiety in young children? Psychometric properties of the modified Abbreviated Math Anxiety Scale for Elementary Children (mAMAS-E). *Polish Psychological Bulletin*, *50*(4), 303–315. https://doi.org/10.24425/ppb.2019.131003

Szczygieł, M. (2020a). Gender, general anxiety, math anxiety and math achievement in early school-age children. *Issues in Educational Research*, *30*(3), 1126–1142.

Szczygieł, M. (2020b). When does math anxiety in parents and teachers predict math anxiety and math achievement in elementary school children? The role of gender and grade year.

Social Psychology of Education, *23*(4), 1023–1054. https://doi.org/10.1007/s11218-020-09570-2

Szczygieł, M. (2022). Not only reliability! The importance of the ecological validity of the Math Anxiety Questionnaire for Adults. *European Journal of Psychological Assessment*, *38*(2), 78–90. https://doi.org/10.1027/1015-5759/a000646

Szczygieł, M., Hunt, T. E., & Sarı, M. H. (2025). Domain-specific and domain-general predictors of math anxiety in adolescents and adults. *Contemporary Educational Psychology*, *82*, 102400. https://doi.org/10.1016/j.cedpsych.2025.102400

Szczygieł, M., & Pieronkiewicz, B. (2022). Exploring the nature of math anxiety in young children: Intensity, prevalence, reasons. *Mathematical Thinking and Learning*, *24*(3), 248–266. https://doi.org/10.1080/10986065.2021.1882363

Szczygieł, M., Szűcs, D., & Toffalini, E. (2024). Math anxiety and math achievement in primary school children: Longitudinal relationship and predictors. *Learning and Instruction*, *92*, 101906. https://doi.org/10.1016/j.learninstruc.2024.101906

Svraka, B., Álvarez, C., & Szücs, D. (2024). Anxiety predicts math achievement in kindergarten children. *Frontiers in Psychology, 15*, 1335952. https://doi.org/10.3389/fpsyg.2024.1335952

Symonds, J. E., & Galton, M. (2014). Moving to the next school at age 10–14 years: An international review of psychological development at school transition. *Review of Education, 2*(1), 1–27. https://doi.org/10.1002/rev3.3021

Tomasetto, C., Passolunghi, M. C., De Vita, C., Guardabassi, V., & Morsanyi, K. (2025). Parental mathematics anxiety is related to children's mathematical development in preschool and the first school years. *Journal of Experimental Child Psychology*, *252*, 106185. https://doi.org/10.1016/j.jecp.2024.106185

Vanbinst, K., Bellon, E., & Dowker, A. (2020). Mathematics anxiety: An intergenerational approach. *Frontiers in Psychology*, *11*(1648). https://doi.org/10.3389/fpsyg.2020.01648

Vukovic, R. K., Roberts, S. O., & Green Wright, L. (2013). From parental involvement to children's mathematical performance: The role of mathematics anxiety. *Early Education & Development*, *24*(4), 446–467. https://doi.org/10.1080/10409289.2012.693430

Wang, C., Li, X., & Lin, X. (2025). Peer support as a buffer: Reducing math anxiety through psychological resilience in left-behind rural students. *Humanities and Social Sciences Communications*, *12*(1), 715. https://doi.org/10.1057/s41599-025-05075-5

Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children-Revised*. The Psychological Corporation.

Wesarg-Menzel, C., Ebbes, R., Hensums, M., Wagemaker, E., Zaharieva, M. S., Staaks, J. P. C., van den Akker, A. L., Visser, I., Hoeve, M., Brummelman, E., Dekkers, T. J., Schuitema, J. A., Larsen, H., Colonnesi, C., Jansen, B. R. J., Overbeek, G., Huizenga, H. M., & Wiers, R. W. (2023). Development and socialization of self-regulation from infancy to adolescence: A meta-review differentiating between self-regulatory abilities, goals, and motivation. *Developmental Review*, *69*, 101090. https://doi.org/10.1016/j.dr.2023.101090

Wu, S. S., Barth, M., Amin, H., Malcarne, V., & Menon, V. (2012). Math anxiety in second and third graders and its relation to mathematics achievement. *Frontiers in Psychology, 3*, 162. https://doi.org/10.3389/fpsyg.2012.00162

Zeidner, M. (1998). *Test anxiety: The state of the art.* Kluwer Academic Publishers. https://doi.org/10.1007/b109548

Zhu, H., Wu, S., Guo, K., Xu, Y., Wang, Y., & Si, J. (2026). Intergenerational transmission of math anxiety in Chinese primary school children: A longitudinal study of math learning involvement. *Journal of Experimental Child Psychology*, *268,* 106526. https://doi.org/10.1016/j.jecp.2026.106526

Zhuo, X., Wang, Y., Xu, Y., Feng, H., Liu, C., Wang, Y., & Si, J. (2025). How students' math anxiety profiles change in primary school: The roles of teacher support, peer support and math attitudes. *British Journal of Educational Psychology*, *95*(3), 811–835. https://doi.org/10.1111/bjep.12758

Živković, M., Pellizzoni, S., Mammarella, I. C., & Passolunghi, M. C. (2023). The relationship betweens math anxiety and arithmetic reasoning: The mediating role of working memory and self-competence. *Current Psychology*, *42*(17), 14506–14516. https://doi.org/10.1007/s12144-022-02765-0

Funding: This study was funded by the Priority Research Area 'Society of the Future' of the Strategic Programme 'Excellence Initiative' at Jagiellonian University.

Data availability statement: The data and R analysis script that support the findings of this study are available on the Open Science Framework at https://osf.io/am4yd/ (DOI: 10.17605/OSF.IO/AM4YD)

Acknowledgements: We would like to thank Anna Domagała, Natalia Hetman, Wiktoria Rabińska, and Maja Wójcik for their assistance in data collection. MH's work on Math Anxiety is supported by the European Union under the Erasmus+ Cooperation partnerships in higher education (KA220-HED), Project title ErasMAs+ Enriching HEI training to help future primary school teachers tackle math anxiety in themselves and their pupils, Project number ID: 2025-1-PL01-KA220-HED-000360673, total amount: 400 000 EUR, funding period 2025-2028 [The views and opinions expressed in this article are those of the authors only and do not necessarily reflect the views of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.]

Supplementary Material

Structure of Students, Classes, Schools

The sample had a hierarchical structure, with students nested within 20 classes across 9 schools. The number of students per class ranged from 1 to 16.

Table A. *Sample structure by school and class*

School ID	Class ID	*N* Students
1	1	11
1	2	13
1	3	3
1	4	10
1	5	10
2	6	16
3	7	8
3	8	10
4	9	4
4	10	3
4	11	1
5	12	4
5	13	8
5	14	3
6	15	12
6	16	2
6	17	10
7	18	8
7	19	13
7	20	8
8	21	4
9	22	5

Sensitivity Analyses

Complete-case subsample As a sensitivity analysis, the multilevel main models were re-estimated using a complete-case subsample including only participants with complete data on all variables included in Model 4 (n = 111 children nested within 18 classes; see Table B). The same model specifications as in the main analyses were applied and estimated using restricted maximum likelihood (REML) and bootstrap-based 95% CIs with 50000 replications. Because the class-level variance approached zero in Model 1, the intraclass correlation coefficient (ICC) and conditional R^2 were not defined for Model 1. Variance components, ICCs, and variance reduction indices were calculated relative to the unconditional class model estimated on the same complete-case sample.

Results from the complete-case sensitivity analysis showed a pattern largely consistent with the main analyses. In Model 1, a one standard deviation increase in intellectual helplessness in math was associated with 0.19 *SD* increase in MA, whereas gender and age were not significant predictors. After adding general and test anxiety in Model 2, a one standard deviation increase in test anxiety was associated with 0.25 *SD* increase in MA, while intellectual helplessness in math and general anxiety were not significantly related to MA. In Model 3, test anxiety remained a significant positive predictor, such that a one standard deviation increase in test anxiety corresponded to a 0.26 *SD* increase in MA. Fluid intelligence and verbal working memory were not significantly associated with MA. In Model 4, which additionally included mother's home numeracy activities, mothers' MA, and their education, test anxiety remained the only significant predictor, indicating that a one standard deviation increase in test anxiety corresponded to a 0.24 *SD* increase in MA, whereas the maternal variables were not significantly associated with children's MA.

Across models, systematic differences between schools remained evident, whereas class-level variance was minimal (ICC = .002–.014). The proportion of variance in MA

explained by fixed effects ranged from 57% to 60%, while the total variance explained by both fixed and random effects ranged from 59% to 61%. Overall, the direction and magnitude of effects were highly similar to those obtained in the primary analyses, indicating that the results were robust to the restriction of the sample to complete cases. Relative to the unconditional class model estimated on the complete-case sample (n = 111), predictors reduced between-class variance by approximately 99–100%, while reductions in residual variance were more moderate and approximately 16–22% across models.

Table B. *Multilevel regression models of MA predictors estimated on the complete-case subsample*

DV: mAMAS-E	**Model 1**				**Model 2**				**Model 3**				**Model 4**			
Random effects																
	Variance (*SD*) [95% CI SE]				Variance (*SD*) [95% CI SE]				Variance (*SD*) [95% CI SE]				Variance (*SD*) [95% CI SE]			
Class Intercept	0.000 (0.00) [0, 0.28]				0.001 (0.03) [0, 0.28]				.003 (.05) [0, 0.28]				0.006 (0.07) [0.00, 0.30]			
Residual	0.42 (0.65) [0.55, 0.73]				0.40 (0.63) [0.54, 0.72]				.39 (.62) [0.53, 0.71]				0.40 (0.63) [0.53, 0.72]			
Fixed effects																
	Estimate	SE	[95% CI]	*p*	Estimate	SE	[95% CI]	*p*	Estimate	SE	[95% CI]	*p*	Estimate	SE	[95% CI]	*p*
Intercept	1.11	0.15	[0.83, 1.40]	< .001	1.07	0.15	[0.79, 1.36]	< .001	1.03	0.15	[0.74, 1.32]	< .001	1.02	0.19	[0.65, 1.40]	< .001
School 2	-1.23	0.23	[-1.69, -0.78]	.002	-1.22	0.23	[-1.68, -0.77]	.004	-1.28	0.24	[-1.75, -0.82]	.002	-1.28	0.25	[-1.77, -0.78]	.002
School 3	-1.21	0.22	[-1.63, -0.78]	< .001	-1.20	0.22	[-1.62, -0.77]	< .001	-1.21	0.22	[-1.63, -0.78]	< .001	-1.20	0.23	[-1.65, -0.76]	< .001
School 4	-1.58	0.29	[-2.14, -1.02]	< .001	-1.44	0.28	[-2.01, -0.88]	< .001	-1.35	0.29	[-1.91, -0.78]	< .001	-1.35	0.30	[-1.94, -0.75]	< .001
School 5	-1.07	0.22	[-1.50, -0.64]	< .001	-0.99	0.22	[-1.41, -0.57]	< .001	-0.95	0.22	[-1.38, -0.53]	< .001	-0.95	0.23	[-1.39, -0.50]	< .001
School 6	-2.05	0.23	[-2.51, -1.60]	< .001	-1.87	0.24	[-2.34, -1.40]	< .001	-1.78	0.25	[-2.27, -1.31]	< .001	-1.77	0.26	[-2.27, -1.27]	< .001
School 7	-1.59	0.21	[-1.99, -1.19]	< .001	-1.49	0.21	[-1.90, -1.07]	< .001	-1.42	0.23	[-1.86, -0.97]	< .001	-1.43	0.24	[-1.90, -0.96]	< .001

School 8	-1.61	0.49	[-2.59, -0.64]	< .001	-1.58	0.49	[-2.54, -0.62]	.002	-1.53	0.49	[-2.50, -0.56]	.003	-1.52	0.51	[-2.52, -0.52]	.004
Gender	-0.08	0.14	[-0.34, 0.19]	.579	-.12	0.13	[-0.38, 0.14]	.361	-.09	0.14	[-0.36, 0.18]	.506	-0.09	0.14	[-0.37, 0.18]	.513
Age	0.09	0.07	[-0.05, 0.22]	.199	.10	0.07	[-0.03, 0.23]	.147	0.11	0.07	[-0.03, 0.24]	.116	0.11	0.07	[-0.03, 0.24]	.131
MATH	**-0.19**	**0.09**	**[-0.36, -0.02]**	**.031**	-.11	0.09	[-0.29, 0.06]	.192	-0.10	0.09	[-0.28, 0.07]	.257	-0.10	0.09	[-0.29, 0.08]	.266
IHMS	**0.19**	**0.08**	**[0.04, 0.34]**	**< .01**	.07	0.09	[-0.10, 0.25]	.408	0.08	0.09	[-0.09, 0.25]	.377	0.09	0.09	[-0.09, 0.27]	.334
STAI-6					-.001	0.07	[-0.14, 0.14]	.982	0.01	0.07	[-0.13, 0.16]	.838	0.01	0.07	[-0.13, 0.16]	.839
CTAS-7					**0.25**	**0.09**	**[0.07, 0.43]**	**.008**	**0.26**	**0.09**	**[0.08, 0.44]**	**.006**	**0.24**	**0.10**	**[0.06, 0.43]**	**.012**
RAVEN									-0.12	0.07	[-0.25, 0.02]	.103	-0.11	0.07	[-0.25, 0.03]	.121
WISC-R									0.11	0.07	[-0.03, 0.25]	.143	0.11	0.07	[-0.03, 0.25]	.139
MAQA-M													0.01	0.07	[-0.13, 0.16]	.835
HNAS-M													0.04	0.07	[-0.09, 0.16]	.590
EDU-M													0.01	0.16	[-0.30, 0.32]	.968
n	111 children, 18 classes				111 children, 18 classes				111 children, 18 classes				111 children, 18 classes			
ICC	NA				.002				.007				.014			
CVR (%)	100				99.78				99.38				98.75			

RVR (%)	15.85	20.32	21.98	20.03
Marginal *R*2	.568	.593	.602	.595
Conditional *R*2	NA	.592	.605	.601
REML	230.7	229.9	233.1	241.8

*$p < .05$, **$p < .01$, ***$p < .001$, listwise deletion within each model. DV – dependent variable, Age – the mean age at first measurement in months, mAMAS-E – MA, STAI-6 – general anxiety, CTAS-7 – test anxiety, IHMS – intellectual helplessness in mathematics, RAVEN – fluid intelligence, WISC-R – verbal working memory, MATH – math achievement, MAQA-M – mothers' MA, HNAS-M – mothers' home numeracy activities, EDU-M – mothers' education. Gender was coded 0 – boys, 1 – girls; EDU-M was coded 0 – lower and 1 – higher education. CVR – class variance reduction (%) and RVR – residual variance reduction (%) were calculated relative to the unconditional class model as the proportional decrease in variance components after adding predictors. They were calculated relative to the unconditional class model estimated on the sample $n = 111$.

Multiple imputed data As a second sensitivity analysis, multiple imputation was conducted using the *mice* package in R, generating 20 imputed datasets via chained equations, receiving $n = 166$. For each imputed dataset, the same multilevel models were estimated using REML. Parameter estimates were then pooled across the 20 imputations using Rubin's rules, and standard errors and 95% CIs were derived from the pooled estimates as implemented in the *mice* package. The number of imputations used exceeds the minimum recommended to obtain stable estimates and adequately reflect sampling variability due to missing data (Rubin, 1987; White et al., 2011; van Buuren, 2018). Because Rubin's rules combine within- and between-imputation variability, additional bootstrap procedures were not applied. Variance components (class- and residual-level variance), ICCs, and variance reduction indices were calculated separately for each imputed dataset and then averaged across imputations to obtain pooled model-level statistics.

Results based on the multiply imputed datasets were highly consistent with the primary analyses. In Model 1, a one standard deviation increase in intellectual helplessness was associated with 0.29 *SD* increase in MA. Gender, age, and math achievement were not significant predictors. After adding general and test anxiety in Model 2, test anxiety emerged as a significant positive predictor, such that a one standard deviation increase corresponded to a 0.25 *SD* increase in MA, while general anxiety was not significantly related to MA. Intellectual helplessness remained positively associated with MA, although the effect size was reduced (0.15 *SD* per one *SD* increase). In Model 3, which additionally included fluid intelligence and verbal working memory, test anxiety (0.22 *SD* increase in MA per one *SD* increase) and intellectual helplessness (0.17 *SD* increase) remained significant predictors. Fluid intelligence showed a negative association with MA, such that a one standard deviation increase corresponded to 0.12 *SD* decrease in MA, whereas verbal working memory was not significantly associated with MA. In Model 4, which additionally included mothers' home

numeracy activities, mothers' MA, and their education, intellectual helplessness in math (0.15 *SD* increase in MA per one *SD* increase) and test anxiety (0.25 *SD* increase) remained significant predictors, while the mothers' variables were not significantly related to children's MA.

Across models, systematic differences between schools remained evident, whereas class-level variance was small (ICC = .019–.030). The proportion of variance in MA explained by fixed effects ranged from 58% to 62%, while the total variance explained by both fixed and random effects ranged from 59% to 63%. Relative to the unconditional class model estimated on the imputed sample, the inclusion of predictors reduced between-class variance by approximately 97–98%, while reductions in residual variance were moderate, approximately 24–31%.

Table C. *Multilevel regression models of MA predictors estimated on multiply imputed data*

DV: mAMAS-E	**Model 1**				**Model 2**				**Model 3**				**Model 4**			
Random effects																
	Variance (*SD*)				Variance (*SD*)				Variance (*SD*)				Variance (*SD*)			
Class Intercept	0.012 (0.11)				0.011 (0.10)				0.012 (0.11)				0.007 (0.09)			
Residual	0.42 (0.65)				0.39 (0.62)				0.38 (0.62)				0.39 (0.62)			
Fixed effects																
	Estimate	SE	[95% CI]	*p*	Estimate	SE	[95% CI]	*p*	Estimate	SE	[95% CI]	*p*	Estimate	SE	[95% CI]	*p*
Intercept	1.03	0.13	[0.77, 1.28]	< .001	1.01	0.12	[0.76, 1.25]	< .001	0.99	0.12	[0.75, 1.25]	< .001	0.93	0.16	[0.61, 1.24]	< .001
School 2	-1.43	0.23	[-1.88, -0.98]	< .001	-1.44	0.22	[-1.88, -1.01]	< .001	-1.49	0.22	[-1.93, -1.05]	< .001	-1.47	0.22	[-1.91, -1.04]	< .001
School 3	-1.14	0.22	[-1.57, -0.72]	< .001	-1.15	0.21	[-1.56, -0.75]	< .001	-1.20	0.21	[-1.61, -0.79]	< .001	-1.21	0.21	[-1.62, -0.80]	< .001
School 4	-1.45	0.28	[-2.01, -0.90]	< .001	-1.38	0.27	[-1.91, -0.84]	< .001	-1.34	0.27	[-1.88, -0.81]	< .001	-1.32	0.28	[-1.87, -0.76]	< .001
School 5	-1.22	0.21	[-1.64, -0.80]	< .001	-1.13	0.20	[-1.54, -0.73]	< .001	-1.12	0.21	[-1.53, -0.72]	< .001	-1.06	0.21	[-1.47, -0.64]	< .001
School 6	-1.80	0.20	[-2.19, -1.40]	< .001	-1.65	0.20	[-2.04, -1.26]	< .001	-1.63	0.20	[-2.03, -1.24]	< .001	-1.56	0.21	[-1.98, -1.13]	< .001
School 7	-1.52	0.18	[-1.87, -1.17]	< .001	-1.40	0.18	[-1.74, -1.05]	< .001	-1.36	0.18	[-1.72, -1.01]	< .001	-1.30	0.20	[-1.68, -0.92]	< .001

School 8	-1.60	0.37	[-2.33, -0.88]	< .001	-1.56	0.35	[-2.26, -0.86]	< .001	-1.56	0.35	[-2.26, -0.86]	< .001	-1.79	0.41	[-2.60, -0.99]	< .001
Gender	0.03	0.11	[-0.19, 0.25]	.800	-.04	0.11	[-.25, 0.17]	.722	-.02	0.11	[-0.23, 0.19]	.851	-0.01	0.12	[-0.23, 0.22]	.955
Age	0.08	0.06	[-0.04, 0.20]	.170	0.09	0.06	[-0.03, 0.20]	.127	0.10	0.06	[-0.02, 0.21]	.115	0.10	0.06	[-0.02, 0.22]	.109
MATH	-0.11	0.07	[-0.24, 0.02]	.090	-0.07	0.06	[-0.20, 0.06]	.273	-0.06	0.07	[-0.19, 0.07]	.353	-0.09	0.07	[-0.23, 0.06]	.221
IHMS	**0.29**	**0.06**	**[0.18, 0.41]**	**< .001**	**0.15**	**0.07**	**[0.02, 0.28]**	**.028**	**0.17**	**0.07**	**[0.03, 0.30]**	**.015**	**0.15**	**0.07**	**[0.01, 0.30]**	**.046**
STAI-6					0.05	0.06	[-0.07, 0.16]	.407	0.06	0.06	[-0.05, 0.18]	.265	0.05	0.06	[-0.07, 0.18]	.417
CTAS-7					**0.25**	**0.07**	**[0.11, 0.38]**	**< .001**	**0.22**	**0.07**	**[0.09, 0.36]**	**.001**	**0.25**	**0.08**	**[0.10, 0.40]**	**.002**
RAVEN									-0.12	0.06	[-0.23, -0.01]	.040	-0.12	0.06	[-0.24, 0.01]	.057
WISC-R									0.06	0.06	[-0.06, 0.17]	.309	0.09	0.06	[-0.03, 0.21]	.133
MAQA-M													0.02	0.07	[-0.12, 0.17]	.739
HNAS-M													0.01	0.06	[-0.12, 0.13]	.912
EDU-M													0.08	0.14	[-0.19, 0.36]	.541
n	166 children, 19 classes				166 children, 19 classes				166 children, 19 classes				166 children, 19 classes			
ICC	.0266				.0264				.030				.019			

CVR (%)	97.42	97.66	97.35	98.36
RVR (%)	23.78	30.08	31.43	29.70
Marginal R^2	.579	.615	.623	.618
Conditional R^2	.590	.625	.634	.625
REML	337.22	329.76	332.98	314.37

*$p < .05$, **$p < .01$, ***$p < .001$, listwise deletion within each model. DV – dependent variable, Age – the mean age at first measurement in months, mAMAS-E – MA, STAI-6 – general anxiety, CTAS-7 – test anxiety, IHMS – intellectual helplessness in mathematics, RAVEN – fluid intelligence, WISC-R – verbal working memory, MATH – math achievement, MAQA-M – mothers' MA, HNAS-M – mothers' home numeracy activities, EDU-M – mothers' education. Gender was coded 0 – boys, 1 – girls; EDU-M was coded 0 – lower and 1 – higher education. CVR – class variance reduction (%) and RVR – residual variance reduction (%) were calculated relative to the unconditional class model as the proportional decrease in variance components after adding predictors. They were calculated relative to the unconditional class model estimated on the sample $n = 166$.

Table D

Comparison of predictor significance across three analytic approaches

Predictor	**Main analysis**				**Complete-case**				**Multiple imputation**			
	M1	**M2**	**M3**	**M4**	**M1**	**M2**	**M3**	**M4**	**M1**	**M2**	**M3**	**M4**
Gender	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*
Age	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*
MATH	*ns*	*ns*	*ns*	*ns*	* (−)	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*
IHMS	* (+)	*ns*	* (+)	*ns*	* (+)	*ns*	*ns*	*ns*	* (+)	* (+)	* (+)	* (+)
STAI-6	—	*ns*	*ns*	*ns*	—	*ns*	*ns*	*ns*	—	*ns*	*ns*	*ns*
CTAS-7	—	* (+)	* (+)	* (+)	—	* (+)	* (+)	* (+)	—	* (+)	* (+)	* (+)
RAVEN	—	—	* (−)	*ns*	—	—	*ns*	*ns*	—	—	* (−)	*ns*
WISC-R	—	—	*ns*	*ns*	—	—	*ns*	*ns*	—	—	*ns*	*ns*
MAQA-M	—	—	—	*ns*	—	—	—	*ns*	—	—	—	*ns*
HNAS-M	—	—	—	*ns*	—	—	—	*ns*	—	—	—	*ns*
EDU-M	—	—	—	*ns*	—	—	—	*ns*	—	—	—	*ns*

*$p < .05$; *ns* – non-significant predictor, M1– M4 – Models 1–4, — predictor not included in model, (+) positive direction; (−) negative direction. Main analysis: $n = 111–155$; Complete-case: $n = 111$; Multiple imputation: $n = 166$. Age – the mean age at first measurement in months, mAMAS-E – MA, STAI-6 – general anxiety, CTAS-7 – test anxiety, IHMS – intellectual helplessness in mathematics, RAVEN – fluid intelligence, WISC-R – verbal working memory, MATH – math achievement, MAQA-M – mothers' MA, HNAS-M – mothers' home numeracy activities, EDU-M – mothers' education. Gender was coded 0 – boys, 1 – girls; EDU-M was coded 0 – lower and 1 – higher education.

Model Comparison Using Maximum Likelihood

The four models reported in the main analyses were estimated using restricted maximum likelihood (REML). Because REML log-likelihoods are not comparable across models that differ in their fixed-effect structure, we re-estimated all models using full maximum likelihood (ML; REML = FALSE in lmer), which yields directly comparable log-likelihoods and thus permits formal model comparisons. Likelihood ratio tests (LRT) and information criteria (AIC and BIC) are valid only when all compared models are fitted to the same observations. Because of item-level missingness, the four main models in the main analyses were fitted to different samples (Model 1, $n = 155$; Model 2, $n = 155$; Model 3, $n = 154$; Model 4, $n = 111$). Therefore, here we conducted the model comparisons on the common complete-case sample ($n = 111$), that is, the same subsample used for the complete-case sensitivity analysis (Table B).

The unconditional class model (M0) and Models 1 to 4 were re-estimated with ML on this sample and compared with sequential likelihood-ratio tests and with AIC and BIC; Akaike and Bayesian model weights were computed across Models 1 to 4. For reference, marginal R^2 is also reported. Because the class-level variance was estimated at zero in every conditional model, the random-effect variance could not be partitioned and conditional R^2 was undefined; the reported R^2 therefore reflects the fixed effects alone.

The models comparisons are presented in Table E. Adding the sociodemographic (gender, age) and math domain-specific (math achievement, intellectual helplessness in math) variables, together with school, to the unconditional model substantially improved model fit, $\chi^2(11) = 62.46$, $p < .001$. Adding domain-general affective variables (general and test anxiety) further improved fit, $\chi^2(2) = 8.15$, $p = .017$. In contrast, adding domain-general cognitive variables (fluid intelligence and verbal working memory) did not improve fit, $\chi^2(2) = 4.35$, $p = .114$. Finally, adding family-related variables (mothers' home numeracy activities, mothers'

MA, and mothers' education) did not improve fit either, $\chi^2(3) = 0.35$, $p = .949$. The information criteria converged with this pattern. AIC was lowest for Model 3 and, very similarily, Model 2 (229.75 and 230.10) and highest for Model 4 (235.40). The more conservative BIC was lowest for the parsimonious Models 1 and 2 (272.18 and 273.45) and penalized the more complex Models 3 and 4 (278.52 and 292.30). Expressed as model weights, Models 2 and 3 together carried 92% of the Akaike weight (.42 and .50), whereas the Bayesian weights concentrated on Models 1 and 2 (.64 and .34). Marginal R^2 rose from .596 (Model 1) to .625 (Model 2) and then changed only little (.639 in Model 3; .640 in Model 4); on the common sample the index is therefore essentially flat from the affective block onward, and the small apparent decline at Model 4 seen in Table B (and in Table 3 of the main text) reflects the different per-model samples used there rather than a genuine loss of fit.

Because Model 3 (with added domain-general cognitive block) does not involve the family variables, it could additionally be tested on the larger child-complete sample ($n = 154$), in which statistical power is greater. The domain-general cognitive block remained non-significant there, $\chi^2(2) = 4.96$, $p = .084$ (on that sample the preceding steps were $\chi^2(11) = 80.68$, $p < .001$ for the sociodemographic and math domain-specific block and $\chi^2(2) = 17.19$, $p < .001$ for the affective domain-general block). The non-significance of the domain-general cognitive block on the common sample is therefore not attributable solely to the reduced size of the $n = 111$ subsample.

Note that all of the conditional models (Models 1 to 4) were singular: the class-level variance was estimated at the boundary (approximately zero) once school was included as a fixed effect, consistent with the small ICCs reported in the main paper and in Table B. This does not affect the validity of the likelihood-ratio tests, because every compared model has the identical class random-effect structure, so the tests concern only the added fixed effects. Note also that, the comparison was conducted on the common complete-case sample so that all

models are fitted to the same observations; the conclusions therefore apply to that subsample and are further supported by the larger-sample robustness check reported above.

Overall, the model comparison indicates that, beyond the domain-general affective block, the sequentially added domain-general cognitive and family blocks did not provide a statistically meaningful improvement in terms of model fit. This is consistent with the analytic aim of the present study, namely to assess the stability of the affective predictors under successive theory-driven controls rather than to maximize the variance explained in children's MA. This also mirrors the non-significant coefficients of the domain-general cognitive and family predictors reported in Table 3. The comparison presented here thus reinforces the conclusion that test anxiety is the central predictor of children's MA.

Table E. *Model comparison using full maximum likelihood (common complete-case sample, n = 111)*

Model	Variables added	k	−2LL	AIC	BIC	AIC wt	BIC wt	$\Delta\chi^2$ (df)	*p*	Marg. R^2
M0	Unconditional model (intercept + class)	3	268.71	274.71	282.84	—	—	—	—	—
M1	M0 + Gender, age, IHMS, MATH (+ school)	14	206.25	234.25	272.18	.05	.64	62.46 (11)	< .001	.596
M2	M1 + STAI-6, CTAS-7 (+ school)	16	198.10	230.10	273.45	.42	.34	8.15 (2)	.017	.625
M3	M2 + RAVEN, WISC-R (+ school)	18	193.75	229.75	278.52	.50	.03	4.35 (2)	.114	.639
M4	M3 + MAQA-M, HNAS-M, EDU-M (+ school)	21	193.40	235.40	292.30	.03	< .001	0.35 (3)	.949	.640

M0 – unconditional model; M1–M4 – Models 1–4. All models were re-estimated with full maximum likelihood (ML) on the common complete-case sample (*n* = 111 children. Variables added: M1 adds gender, age, IHMS — intellectual helplessness in mathematics, MATH – math achievement, and school; M2 adds STAI-6 – general anxiety and CTAS-7 – test anxiety; M3 adds RAVEN – fluid intelligence and WISC-R – verbal working memory; M4 adds MAQA-M – mothers' MA, HNAS-M – mothers' home numeracy activities, and EDU-M – mothers' education. *k* – number of estimated parameters; −2LL – minus twice the log-likelihood; AIC – Akaike information criterion; BIC – Bayesian information criterion; AIC wt and BIC wt – Akaike and Bayesian model weights, computed across Models 1–4 (the unconditional model is excluded); $\Delta\chi^2$ (df) – sequential likelihood-ratio test of each model against the preceding (simpler) model, with its degrees of freedom; *p* significance; Marg. R^2 – marginal R^2 (variance explained by the fixed effects).

References

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys.* John Wiley & Sons Inc. http://dx.doi.org/10.1002/9780470316696

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, *30*(4), 377–399. https://doi.org/10.1002/sim.4067

van Buuren, S. (2018). *Flexible imputation of missing data* (2nd ed.). Chapman and Holl/CRC Press.