

Research Article

Primary School Age Students' Spontaneous Comments about Math Reveal Emerging Dispositions Linked to Later Mathematics Achievement

Michèle M. M. Mazzocco,^{1,2} Laurie B. Hanich,³ and Maia M. Noeder⁴

¹ Institute of Child Development, University of Minnesota, 51 East River Parkway, Minneapolis, MN 55455, USA

² Johns Hopkins University School of Education Baltimore, MD 21218, USA

³ Department of Educational Foundations, Millersville University of Pennsylvania, P.O. Box 1002, Millersville, PA 17551, USA

⁴ Department of Psychology, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106-7123, USA

Correspondence should be addressed to Michèle M. M. Mazzocco, mazzocco@umn.edu

Received 12 June 2012; Revised 10 August 2012; Accepted 10 August 2012

Academic Editor: Ann Dowker

Copyright © 2012 Michèle M. M. Mazzocco et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To longitudinally explore children's developing beliefs towards mathematics, we asked 207 children to define "math" and "reading" at grades 2 and 3 and coded for spontaneous references to likability or difficulty of math (or reading) in their definitions. We found that children attributed more difficulty to math than to reading despite their relatively neutral comments on the likability of either subject. Children described math and reading with comparable degrees of specificity, but girls' definitions were more specific than boys'. Relative to their peers, children with mathematics learning disability (MLD) provided less specific definitions overall, were more likely to describe math as more difficult than reading, and were more likely to show a decrease in likability ratings of math (but not reading) from grades 2 to 3. Grade 2 ratings predicted math ability at grade 3, more so than predictors from grade 3. These findings, although based on informal analyses not intended to substitute for validated assessments of disposition, support the notions that distinct aspects of dispositions towards math emerge in early childhood, are revealed through casual discourse, and are predictive of later math achievement outcomes. This further supports current interests in developing formal measures of academic disposition in early childhood.

1. Introduction

A productive disposition towards mathematics is an essential component of mathematics proficiency [1]. Like many elements of successful mathematics outcomes, the construct of a "productive disposition" is multifaceted. The National Research Council [1] defines it as "the tendency to see sense in mathematics, to perceive math as both useful and worthwhile, and to believe that steady effort in learning mathematics pays off." At a minimum, this description captures features of positive attitudes about mathematics, seeing sense in mathematics, the belief that effort is needed to support math learning, and a perspective that math is useful within and beyond school experiences—regardless of

whether it is liked or disliked or achieved with minimal or great effort. Associations between these and other distinct features of a productive disposition are likely to be dynamic, but the nature and emergence of the elements themselves are not yet fully understood, particularly in early childhood. Here we report on a brief, exploratory study of whether second and third graders' definitions of mathematics shed light on their emerging dispositions and are related to their later mathematics achievement.

The recent interest in measuring beliefs about mathematics in early childhood (e.g., [2, 3]) has yielded evidence of a reciprocal relationship between math interest or beliefs and math ability, even in preschoolers [4]. Accordingly, in this study we ask whether children's spontaneous conversations

about math reveal tendencies in their beliefs. Specifically, we ask if children's definitions of "math" reveal children's beliefs about whether math is easy or difficult, liked or disliked, or useful. We ask whether these potential disposition features apply to math specifically or to school subjects in general by asking our participants to also define "reading." By comparing definitions of math (and reading) among children with a wide range of math achievement levels (including children with mathematics learning disability (MLD)), we test the plausibility that early indicators of children's beliefs about math contribute to predicting their future mathematics outcomes. We base this hypothesis on prior research demonstrating that elementary school students' beliefs about their ability and subjective task values in mathematics predict their current and future activity choices [5, 6].

Until recently, dispositions towards math and related constructs, such as math anxiety, were often not evaluated before middle school [7] or high school [8]. This is troubling in light of evidence of age-related changes in children's motivational patterns. Whereas most students begin schooling enthusiastically [1], a shift from intrinsic to extrinsic motivational orientations emerges at about grade 3 (approximately 8 years of age) [9, 10]. This type of progression has been reported for measures of attitudes and curiosity [11, 12], perceived competence [13–15], and beliefs about intelligence [16], and suggests that early childhood (prior to and up to grade 3) may be an essential period for establishing, maintaining, or at least initiating the foundation of a positive disposition towards mathematics. Moreover, while developing productive dispositions is important for children across a wide range of achievement levels, this may be especially important for students at risk for low mathematics achievement and/or MLD.

To explore how productive learning dispositions are cultivated in early childhood, we utilized an expectancy-value model of achievement motivation [17] which asserts a proportional relationship between task engagement, expectations for success, and subjective task values. Specifically, the amount of effort children invest in an activity (e.g., mathematics engagement) is a product of their expectations for success and the extent to which they value the task. These components are influenced by children's perceptions of competence, task difficulty, and affect, and have been shown to directly affect achievement-related behaviors (as reviewed elsewhere [17]). In the present study, we did not objectively measure students' efforts nor did we rely on self-report measures related to subjective task value. Rather, we used open-ended questioning to specifically test whether indicators of beliefs and values related to mathematics emerge from spontaneous conversations with young children. In particular, we focused on the extent to which children's responses reflect affect (likability), perceived task difficulty, and values reflected in the specificity with which their responses relate to the usefulness of mathematics. The expectancy-value model provided a useful framework for considering the types of beliefs on which to focus in coding children's responses, and for interpreting our findings given the similarities between our variables of interest and those proposed as part of this theoretical model.

What is the relevance of spontaneous conversations of this type in the context of evaluating children's beliefs about math? First, early conversations about math provide a mechanism by which adults can deliberately attend to the potential influences they have on their children's or students' math-oriented beliefs. Second, in addition to using conversations to nurture the development of a healthy math disposition, conversations can be monitored for indicators of a child's emerging beliefs. In other words, early conversations are useful as a platform for adults' messages to children, but also for attending to children's comments—spontaneous or structured.

Why are children's early comments about math potentially important? In view of emerging evidence that early beliefs about mathematics are related to later achievement outcomes, efforts to steer children from paths towards negative outcomes should begin in early childhood, when children's beliefs may be more malleable rather than deeply rooted. Consequences of negative beliefs towards math are frequently described in terms of secondary-school behaviors, such as avoidance of elective mathematics courses after grade 10 (age 15 years) [18]. But the antecedents of such avoidance likely occur much earlier. For example, a negative relationship between math anxiety and math achievement has been observed among elementary school aged children with high working memory capacity [19, 20], and among elementary aged students who believe math-gender stereotypes with female students showing weaker identification with mathematics than their male counterparts despite similar levels of achievement [21, 22]. Gender stereotypes have been shown to influence mathematics achievement gains in first or second grade (ages 6 and 7 years), at least among young girls who conform to the belief that females underperform in math relative to males, and particularly among girls who receive instruction from a teacher with elevated levels of anxiety about mathematics [2]. These early associations warrant attention; information from informal conversation may be among the earliest indicators to direct adults' attention to students' emerging dispositions.

Children's math ability self-perceptions become well established in elementary school (e.g., [6]), but parent expectancies and attributions may influence children's math beliefs [23, 24] before and after the onset of formal schooling. Parents who believe that math and science are male-oriented domains tend to overestimate their sons' math and science performance and underestimate their daughters' performance in math and science at least in mid- to late-elementary school [25]. In middle school students, this underestimation may impact self-concepts and self-underestimation of math ability [26, 27]. These findings suggest that teacher and parent behaviors *can* and *do* play a major role in student learning, and that efforts to counter such effects on negative beliefs about math and achievement-related perceptions should begin in early childhood [23]. Listening to children's dialogues about math may contribute to these efforts and provide opportunity for intercession.

There are a number of ways to evaluate math beliefs. Ratings of children's math self-perception and their enjoyment of mathematics are objectively measurable in elementary

school aged children [13–15, 28] and earlier [4]. Here, we propose that discourse may serve not only as a means for parents and teachers of young children to help shape a positive disposition towards math through adult-initiated “math talk,” but also as a means for adults to glean information about a child’s emerging disposition, despite the fact that dialogues are no substitute for formal screening and diagnostic assessments of risk for poor math outcomes.

What aspects of children’s early dispositions towards mathematics are revealed during open-ended discussions, and are these beliefs related to formal math outcomes? To address these questions, in the present study we used a simple and straightforward approach to elicit children’s spontaneous comments about math. Importantly, we avoided explicit prompting for information about beliefs about math in order to focus on children’s spontaneous comments. Although findings to emerge from this approach do not reveal causal pathways to successful mathematics, evidence of the mere presence of relevant information from spontaneous speech has implications for the role of early conversations about mathematics among children and their parents, teachers, and care providers in shaping or supporting a child’s disposition towards mathematics.

2. Method

2.1. Participants. Participants were drawn from a larger longitudinal study of mathematics achievement and mathematics learning disabilities during the primary school aged years, described elsewhere in more detail [29]. Briefly, the participants from the larger study were recruited from 23 kindergarten classrooms across seven schools in one large metropolitan public school district. The participating schools were selected based on their relatively low rates of mobility and free or reduced lunch participation, to diminish attrition and decrease the likelihood that poor math performance observed among participants was linked to low socioeconomic status. Of the 445 kindergartners from these classrooms invited to participate, 249 enrolled (129 girls). Participation in the present study was limited to the 207 children (107 girls) who completed annual assessments during both grades 2 and 3. Of the 207 included, most (86.5%) participants were white, 8% were black, 3% were Asian, and the remaining 2.5% represented other or mixed ethnicities.

2.2. Materials and Procedures. During each year of the study, children were tested individually by one of five female examiners. Testing occurred during two or three sessions up to 45 minutes each, in a room separate from the classroom or other distracting activities. The testing battery included a range of standardized and experimental assessments. The measures relevant to the present report are two measures of mathematical ability, and a vocabulary probe administered to glean information about each participant’s beliefs about mathematics.

2.2.1. Test of Early Math Ability-Second Edition (TEMA-2, [30]). The TEMA-2 is a standardized measure of formal and

informal mathematical ability normed for use with children ages 2 to 8 years, 11 months. Items from the TEMA-2 include basic knowledge items such as counting principles, calculation and fact retrieval items, and items testing place value concepts or word problem solving. Total correct scores are converted to age-referenced standard scores for which test-retest reliability is 0.94 [30].

We administered the TEMA-2 to participants in the longitudinal study, from kindergarten through grade 3. In the present study, we used the standard scores from all four grades to assign children to one of three groups: children with mathematical learning disability (MLD), children with low mathematics achievement (LA), and children with age appropriate math achievement (typically achieving or TA). For the participants who exceeded the upper age level for the TEMA-2 ceiling at grade 3, we calculated prorated standard scores using regression-models to predict age 9 years and age 9.5 years outcomes, based on data from over 200 children who completed the TEMA-2 during all four years of the study. The criteria for these participant groups, described elsewhere in detail [31], reflect the growing consensus that MLD classification is strengthened by considering scores over time, rather than one score at a single assessment. Children with TEMA-2 scores below the 10th percentile of the study sample were classified as having MLD ($n = 18$) and children with scores between the 11th to 25th percentile were classified as having low achievement in mathematics (LA; $n = 26$). These criteria were met during at least two of the four grades tested, provided that scores during any remaining years fell within the 95th percentile confidence interval of the 10th or 25th percentile cut off, respectively. Children who did not meet these criteria and who scored above the 25th percentile (or within the 95th percentile confidence interval for this range) were classified as typically achieving (TA; $n = 163$).

2.2.2. Woodcock Johnson—Revised Calculations (WJ-R Calc). The Woodcock Johnson Psycho-Educational Battery-Revised [32] is a widely used standardized test of formal academic skills and is normed for use with children ages 2 to 21 years. The WJ-R Calculation subtest, normed for ages 6 and older, is an untimed paper-and-pencil test that includes a range of arithmetic problems presented in order of increasing difficulty. We administered this subtest at grades 1 and 3. Due to floor effects, we did not administer the test at kindergarten; due to time constraints, it was not administered at grade 2. Internal consistency reliability for the WJ-R Calc subtest ranges from $r = .89$ to $r = .93$ across ages 6–9 years [33]. In the present study, we used the standard scores from the WJ-R Calc subtest at grade 3 as the dependent variable in select analyses, controlling for earlier math performance (WJ-R Calc performance at grade 1).

2.2.3. Defining “Mathematics.” During select years of the longitudinal study, the test battery included a standardized expressive vocabulary test from the Stanford Binet Fourth Edition [34] (at grades K to 2) or the Wechsler Abbreviated Scale of Intelligence [35] (at grade 3). These subtests involved

asking children to define a word presented orally by the examiner, beginning and ending at standardized basals and ceilings based on the examinee's age. In accordance with these standardized protocols, we asked each participant, "what is math?" followed by, "what is reading?" without any prompting for any specific or additional information. Only if a child responded with the equivalent of, "I do not know," we added the following prompt: "Just tell me what you think (math/reading) is." These definitions were elicited after a child reached a performance ceiling on the standardized vocabulary subtest, so that the point at which a definition was requested occurred at comparable levels of performance fatigue for all participants. This procedure also avoided unintentional alteration of the standardized vocabulary subtest. Also, by requesting a definition of math at the end of the expressive vocabulary test, we implicitly established expectations for the amount of information expected in response, and could indirectly convey comparable expectations about length of definitions for both "math" and "reading." Definitions of math and reading were obtained during grades 2 and 3.

Children's responses were recorded verbatim and transcribed into a central database with child identifiers removed. This ensured that coders were blind to the children's identity, gender, or math achievement status. Each response was coded independently by two trained coders. Coding responses were compared, and discrepancies were resolved in a coding meeting with the study PI and both coders present. Therefore, all definitions were double coded and, if there was disagreement, by three coders. Coding reliability is reported subsequently.

2.3. Coding Variables and Guidelines. Each definition was coded separately for three attributes: the extent to which the child's definition spontaneously reflected a like or dislike for math or reading (*likability* code), whether the child's definition made reference to ease or difficulty associated with math or reading (*difficulty* code), and how specifically math or reading concepts, procedures, or uses were described (*specificity* code). One set of scoring criteria was developed for each of the likability and difficulty coding protocols (Tables 1 and 2, resp.), each of which was applied to definitions of math and reading. Two parallel sets of criteria were established for coding the specificity of math (Table 3) and reading (Table 4) definitions.

The full range of possible *likability* codes was -2 (extremely dislikes math or reading) to 2 (extremely likes math or reading), with a code of 0 reflecting a neutral response (Table 1). Children's responses were assigned codes based on whether they were moderate (1) to strong or exaggerated (2) statements; these were further coded as positive (1 or 2) or negative (-1 or -2). Neutral responses included either no negative or positive comments, or included both (e.g., "it is kind of fun, but sometimes it is boring"). Of the four sets of 207 valence codes assigned, there were 2 disagreements among the 2 coders (99.9% agreement), which were resolved with a third coder/arbitrator. No child received a likability code of "2" for math at either grade,

although at both grades, codes of "2" were assigned for reading.

The full range of possible *difficulty* codes was -2 (very difficult) to 2 (very easy), with a code of 0 reflecting no reference to either the difficulty or ease associated with mathematics or reading (Table 2). Of the four sets of 207 difficulty codes, there were 19 disagreements among the 2 coders (99.97% agreement), which were all resolved with a third coder/arbitrator. No child received a code of "2" (very easy) for either math or reading, at either grade.

Specificity codes ranged from 1 to 5 , reflecting irrelevant or otherwise uninformative responses (1) to responses reflecting mathematics (or reading) as a useful tool (5). A summary of these coding criteria and sample responses appears in Tables 3 and 4. Of the four sets of 207 specificity codes, there were 58 disagreements among the 2 coders (93.0% agreement), which were resolved with a third coder/arbitrator.

All data were double entered into two excel spreadsheets, subtracted, and reviewed for discrepancies that were then corrected until the subtraction comparisons yielded no errors.

3. Results

3.1. Defining Math and Reading. First we asked whether qualitative features of children's definitions of math and reading differ as a function of gender or MLD status. We carried out three analyses of variance (ANOVAs), each based on a 2 (subject area: math versus reading) $\times 2$ (grades: 2 and 3) $\times 2$ (gender) $\times 3$ (MLD status: TA, LA, and MLD) design, with repeated measures of the first two factors. The outcome variables were likability codes, difficulty codes, or specificity codes from children's spontaneous definitions of math and reading. Note that the means reported in the text are marginal estimated means whereas means reported in the figures are observed means.

3.1.1. Children's Likability for Math and Reading Definitions. We found no main effects of grade, MLD status, or gender on likability codes, $P > .437$. Although mean codes for reading were slightly more positive than those for math, the effect of subject area was not statistically significant, $F(1, 201) = 3.86$, $P = .051$, partial $\eta^2 = 0.019$. Group means for likability codes were essentially neutral for math (-0.011) and reading (0.048) when collapsed across grade and MLD status.

Two-way interactions emerged for grade \times MLD status, $F(2, 201) = 3.86$, $P < .03$, partial $\eta^2 = 0.037$; and grade \times subject area, $F(1, 201) = 6.67$, $P < 0.02$, partial $\eta^2 = 0.032$, but these effects were also very small. From grade 2 to 3, valence codes did not change among children from the TA group (estimated marginal means = 0.032 and 0.023 , resp.), became slightly more positive among children in the LA group (means = -0.017 and 0.095 , resp.), and slightly more negative among children with MLD (means = 0.013 and -0.036 , resp.). From grade 2 to 3, valence codes did not change as much for math (means = 0.013 and -0.035 , resp.) as they did for reading (means = 0.006 and 0.090 ,

TABLE 1: Likability scores: coding criteria and real examples of math and reading definitions.

Score	Definition	Math examples	Reading examples
–2	Extremely disliked, hated, or dreaded. Includes reference to words such as hate, worst, awful, and so forth.	n/a ^a	“When you sit and read a book bored out of your mind.”
–1	Disliked or avoided. Includes reference to words such as do not like, not fun, boring, bad, and so on.	“Math is something I do not like.”	“Something you do only when you need to do it, you have to read to figure out information.”
0	Neutral feeling or tolerated. Includes no reference to an emotion.	“When you learn math problems.”	“Like when you read a book in school or at home.”
1	Liked or enjoyed. Includes reference to words such as play, fun, like, good, and so forth.	“Math is fun, you could do math at school—you could do it anywhere!”	“You can read for information or just for the fun of it.”
2	Extremely liked, loved, or favored. Includes reference to words such as love, favorite, best, and so forth.	n/a ^a	“Oh I love to read.”

Note: ^aNo responses were coded at this value.

TABLE 2: Difficulty scores: coding criteria and real examples of math and reading definitions.

Score	Definition	Math examples	Reading examples
–2	Really hard. Includes reference to words such as <i>difficult</i> , <i>hard</i> , or <i>complex</i> .	“Something you do in school. Something that’s very, very hard.”	“Like you are reading and you are bad at reading and you read something like hard words.”
–1	Kind of hard. Includes reference to words such as the noun form of <i>work</i> , words like <i>learn</i> , phrases that imply exertion zero, or implicit or explicit reference to cognitive processes.	“It is like problems that you have to solve.”	“It is when there is like lots of sentences and you read it to get information.”
0	Neutral difficulty. Includes no reference to difficulty, refers to verbs such as work, do, use, teach, or “work with” without implying exertion.	“Like math projects. Doing math homework.”	“You have a book and you read the words.”
1	Kind of easy. Includes reference to words such as “not hard” or words implying ease.	“It is easy and you have to do it in your homework.”	“To just lay back and read a book.”
2	Really easy. Includes reference to words such as simple, effortless, or speed.	n/a ^a	n/a ^a

Note: ^aNo responses were coded at this value.

resp.), but the direction of these changes differed with math scores becoming slightly more negative, and reading scores becoming slightly more positive (Figure 1).

Even the strongest interaction to emerge from this analysis was associated with a small effect size, for a three-way interaction between subject area, grade, and MLD status, $F(2,201) = 5.53$, $P < .01$, partial $\eta^2 = 0.052$. Pairwise comparisons indicated that children in the TA group made relatively neutral references to both math and reading at

both grades (means = 0.012 for math at grades 2 and 3; and means = .051 and .034 for reading at grades 2 and 3, resp.), with little change from grades 2 to 3. Among children in the LA group, valence codes for definitions of math became slightly more negative over time (means = .000 to –0.045 at grades 2 to 3, resp.), and slightly more positive for reading during the same time period (means = –0.033 to 0.236). Children with MLD made neutral references when defining reading at both grades (both means = 0.000), whereas the

TABLE 3: Specificity scores: coding criteria and real examples of math definitions.

Score	Definition	Examples
1	No response, a circular response, a response unrelated to math, or another uninformative response.	“You do your math.” “I cannot explain it.” “I do not know.”
2	Response is unspecific or only indirectly related to math as a primary school subject. The response may include references to activities performed in relation to math or in a math class but with no discernible reference to math concepts or procedures.	“You play games and stuff.” “Science.” “Like you do something. . .and you have pennies and money.”
3	Unelaborated basic concepts or mechanics of math. Includes reference to real numbers, operations, math problems, or learning math.	“It has to do with numbers and sizes and fractions.” “It means taking away, subtracting, and multiplication.” “When you do all sorts of problems like divide, multiply, fractions, decimals.” “In my class we do like a drill so it like refreshes our memory from the other day.”
4	Elaborated concepts of math.	“Like if you have a word problem, like Jim has 18 apples and eats 3, you use math to solve it.”
5	Concept of math as a useful tool.	“Math is something that people do and they have to know math to be able to get a job and do other stuff.”

largest difference over time was the shift towards more negative ratings of math among the MLD group (means = 0.026 to -0.071 from grades 2 to 3, a difference of -0.097).

In summary, at grades 2 and 3, boys and girls do not spontaneously make reference to extremely positive or negative sentiments about math or reading when defining either term, but there is a slight tendency for more negative comments about math (versus reading) among children with MLD, and, to a lesser degree, among children with low achievement in math.

3.1.2. Children's Reference to the Difficulty of Math and Reading. We also evaluated spontaneous comments about the difficulty versus easiness of math or reading in our second set of repeated measures ANOVAs. Here, the main effect of subject area was significant, $F(1,201) = 16.02$, $P < .0001$, partial $\eta^2 = 0.074$. Children's definitions of math included more references to difficulty (mean code = -0.347) than did their definitions of reading (mean code = -0.171). The main effect of MLD status reflected significantly more negative references (i.e., more ratings reflecting difficulty) to math and reading in the MLD group (mean rating across subject areas = -0.378), relative to the LA and TA groups (whose means ratings were -0.174 and -0.225 , resp.), $F(2,201) = 3.40$, $P < .05$. Still, the effect was small, partial $\eta^2 = 0.033$. Pairwise comparisons showed that math and reading ratings

combined were significantly more negative among the MLD versus LA or TA groups, $P < .03$, but did not differ between the TA and LA groups, $P = .361$.

There was no main effect of grade or gender on difficulty ratings, $P > .87$. No interactions were significant, with the exception of a weak three-way interaction between subject area \times grade \times gender, $F(1,201) = 4.26$, $P < .05$, partial $\eta^2 = .021$. Here, both boys and girls reported slightly greater difficulty for math than reading at both grades. This difference was more pronounced for girls than for boys at grade 2 but not at grade 3. Over time, boys showed a slight shift towards reporting greater difficulty from grades 2 to 3, for both math and reading; among girls, there was a slight shift towards reporting less difficulty for math (and no change in reports of difficulty for reading) from grades 2 to 3 (Figure 2).

In sum, at both grades 2 and 3, both boys and girls make references to math or reading as difficult (or as requiring “work” or exertion), but there is a tendency for math to be described as slightly more difficult than reading, and a tendency for children with MLD to rate both math and reading as more difficult compared to ratings assigned by their non-MLD peers.

3.1.3. Likability and Difficulty as Distinct Constructs. It is possible that likability and difficulty ratings simply represent

TABLE 4: Specificity scores: coding criteria and real examples of reading definitions.

Score	Definitions	Examples
1	No response, a circular response, or a response unrelated to reading or otherwise uninformative.	"Reading means you read a lot." "I do not know that one."
2	Response related to unspecific reading activities, or indirectly related to reading activities. Includes reference to activities that may be performed in relation to reading or for a reading class but without a clear reference to the concept of reading.	"It means that you do reading projects." "It is if you have a book report you can report."
3	Unelaborated basic principles or mechanics of reading. Includes reference to sounding out and/or to reading materials or to learning to read.	"To like sound out something." "You read a book, you read a piece of paper." "When you like read words and books and sound out letters." "Something you do to get smarter at knowing words and learning what the words mean and how to pronounce the words and how to say them. That is why you have teachers to help you pronounce, read, and spell the words."
4	Elaborated concepts of reading. Includes reference to the extraction of meaning from written material through the act of reading.	"Like if you have a book, it has words, and you read the words not to just look and say, but to know what the words say and what the story is about."
5	Concept of reading as a useful tool. Must demonstrate usage of reading or usage of meaning derived from reading.	"To like look at books and some books help you make things and repair your house and some people just read for fun."

TABLE 5: Observed means or definition codes among total study sample ($n = 207$).

	Math		Reading	
	Mean (SD)	Range	Mean (SD)	Range
Grade 2				
Likability	0.01 (0.14)	-1 to 1	0.03 (0.30)	-2 to 2
Difficulty	-0.29 (0.52)	-2 to 0	-0.11 (0.33)	-2 to 0
Specificity	2.85 (0.473)	1 to 4	2.96 (0.63)	1 to 5
Grade 3				
Likability	0.00 (0.14)	-1 to 1	0.06 (0.39)	-2 to 2
Difficulty	-0.32 (0.50)	-2 to 0	-0.21 (0.43)	-2 to 1
Specificity	3.03 (0.66)	1 to 5	2.97 (0.61)	1 to 5

Note: Standard deviations shown in parentheses.

a general reporting tendency, rather than specific constructs. For example, each measure may reflect a positive or negative disposition towards academic subjects in general, in which case likability codes for math and reading should be positively correlated, and difficulty codes for math and reading should also be correlated. Alternatively, the measures may reflect an even broader tendency for positive or negative reporting in general, in which case likability and difficulty codes should be correlated with each other. Finally, if the codes represent stable, subject-domain sentiments, then

math (or reading) likability scores should be correlated across grades.

To explore which of these alternatives is supported, we ran three sets of four correlations, using 12 Spearman rank tests, with alpha adjusted to 0.004 based on multiple correlations (.05/12). With respect to the three alternatives posed above, we found only weak, partial support for the notion of a general academic valence bias, based on the weak positive correlation between likability codes for math and reading observed at grade 2 (Spearman $Rho = .150$,

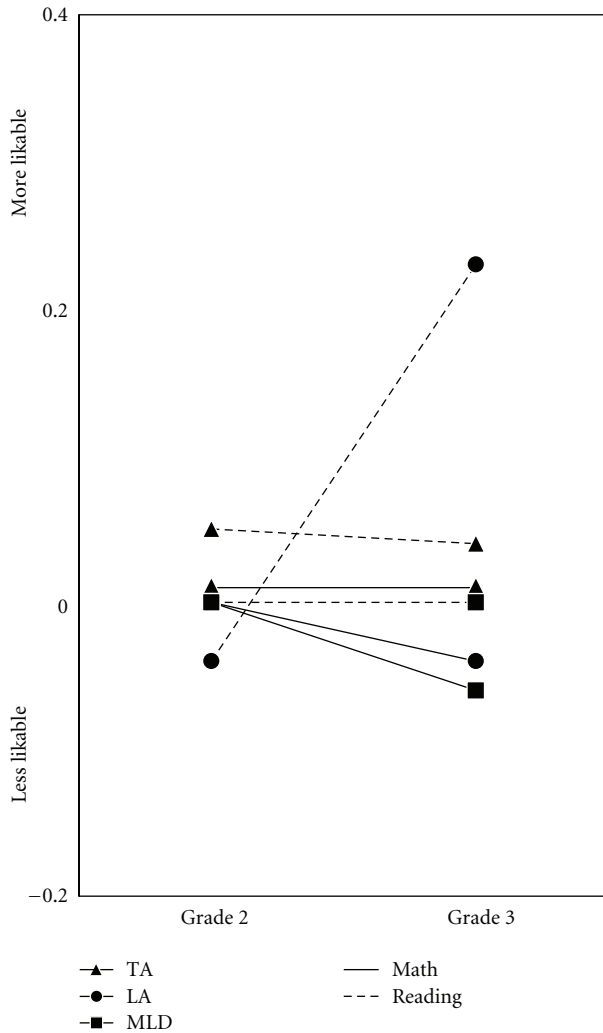


FIGURE 1: Observed means for likability scores, from grades 2 and 3, depict the three following interactions: grade \times MLD status, grade \times subject area, and grade \times MLD status \times subject area.

$P = .031$) but not at grade 3 (Spearman $Rho = .0$, $P = 1.0$); and positive correlations between difficulty codes for math and reading at Grades 2 (Spearman $Rho = .160$, $P = .022$) and 3 (Spearman $Rho = .311$, $P = .0001$). Only the latter correlation at grade 3 met our adjusted significance criteria for multiple correlations.

There was less support for a broad reporting disposition bias, because likability and difficulty codes were not correlated with each other, P ranging from .096 to .594, with the exception of the association between math likability and difficulty at grade 3 only, $-.148$, $P = .033$, which did not meet criteria adjusted for multiple correlations.

There was far more support for subject-specific likability (or difficulty) over time, suggesting that our codes are indicative of stable, and subject-domain beliefs, at least over the short term (from grade 2 to 3): reading likability at grade 2 was correlated with reading likability at grade 3 (Spearman $Rho = .239$, $P = .001$), and reading difficulty at grades 2 and 3 were also correlated with each other

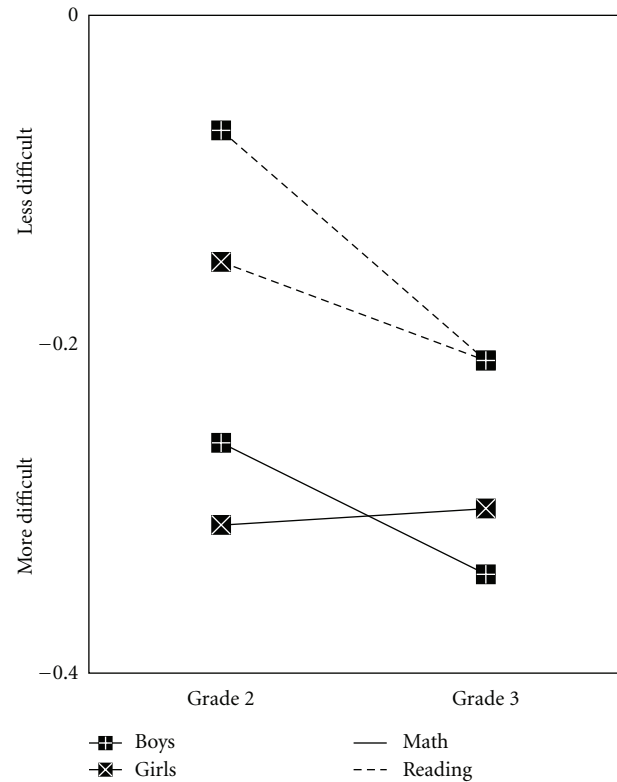


FIGURE 2: Observed means for difficulty scores in grades 2 and 3 depict the significant main effects of subject area and MLD status, and the significant three way interaction between grade, gender, and subject area.

(Spearman $Rho = .210$, $P = .002$). Likewise, math likability at grades 2 and 3 were correlated with each other, (Spearman $Rho = .251$, $P < .0003$), although math difficulty scores were not correlated across grades ($P = .816$). All three of the significant results met our criteria adjusted for multiple correlations. Grade level means are reported in Table 5.

3.1.4. Subject-Area Specificity of Children's Definitions of Math and Reading. The third set of repeated measures ANOVAs concerned the specificity of children's descriptions of math or reading, with codes reflecting noninformative to elaborate descriptions (exemplified in Tables 3 and 4). There was no main effect of subject area on this aspect of children's definitions, $P = .49$. Responses ranged from 1 to 5 for both math and reading, with estimated marginal mean ratings of 2.86 for math and 2.90 for reading, collapsed across all other variables. There was no effect of grade on the specificity of definitions (means = 2.85 to 2.91, at grades 2 and 3, resp.), $P = .29$.

Gender and MLD status each contributed significantly to variance in the specificity of definitions. Girls gave slightly more specific definitions of mathematics or reading (mean = 3.00) than did boys (mean = 2.75), $F(1, 201) = 10.53$, $P = .001$, partial $\eta^2 = 0.050$. The main effect of MLD status, $F(2, 201) = 4.30$, $P < .02$, partial $\eta^2 = 0.041$, reflected more specific definitions of math or reading by children in the TA

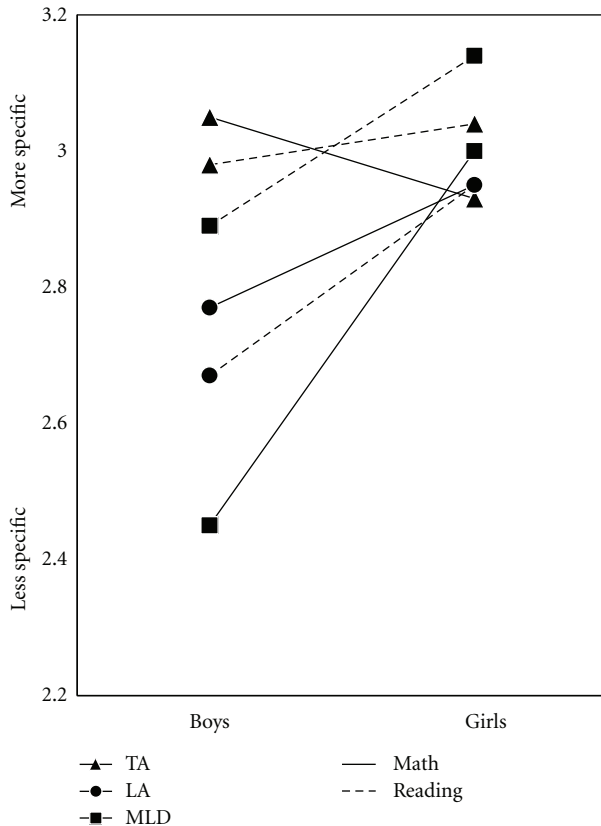


FIGURE 3: Observed means for specificity scores depict significant main effects of MLD status and gender, and the significant MLD status \times gender interaction. Subject area is not significant.

group (mean = 3.00), relative to the MLD or LA groups ($P < .02$), which did not differ from each other, $P = .68$ (means = 2.84 and 2.80, resp.).

The only interaction to reach statistical significance was a two way interaction between MLD status and gender, $F(2, 201) = 5.96$, $P < .004$, partial $\eta^2 = 0.056$. This interaction reflects that girls' ratings did not differ across MLD groups (with estimated marginal mean ratings of 2.99, 2.96, and 3.07 across the TA, LA, and MLD groups, resp.), whereas boys' ratings did (means = 3.02, 2.72, and 2.52, resp.; Figure 3).

3.2. Predicting Mathematics Achievement. Do these exploratory measures predict future or concurrent math performance? We carried out two regression models, each comprised of three predictors of grade 3 WJ-R math Calculation score as the outcome variable of interest. These models were based on predictors obtained at grade 2 or grade 3 (grade level predictors examined separately). When combined, grade 2 likability, difficulty, and specificity codes accounted for approximately 5% of the variance in WJ-R Calculation scores obtained at grade 3, $F(3, 206) = 5.03$, $P < .003$, adjusted $r^2 = .055$. However, likability and difficulty codes did not contribute significantly to the model, $P > .12$,

whereas specificity codes did, $t(206) = 3.11$, $\beta = .212$, $P < .01$.

Model strength improved significantly when grade WJ-R Calculation score was included as a predictor, $F(4, 206) = 32.47$, $P < .0001$, adjusted $r^2 = .379$. Still, the small contribution made by the specificity score remained significant, $t(206) = 2.40$, $\beta = .134$, $P < .02$.

To examine whether *concurrent* predictors would be stronger than predictors obtained one year prior, the analysis was repeated using grade 3 likability, difficulty, and specificity codes to predict grade 3 WJ-R Calculation scores. The model strength for the grade 3 predictor model was weaker than the aforementioned grade 2 predictor model, $F(3, 206) = 3.00$, $P < .04$, adjusted $r^2 = .028$. Here, both the specificity and likability codes predicted grade 3 WJ-R Calculation scores, but effects were small ($t(206)s$ (β) = 2.12 (.147), and 1.99 (.138), resp., $P < .05$), and these effects disappeared when the grade 2 WJ-R Calculation scores were added to the model, $P > .31$.

To evaluate whether the effect to emerge from grade 2 predictors was specific to math disposition, we carried out parallel regression models using grade 2 or grade 3 likability, difficulty, and specificity codes from *reading* definitions, to predict WJ-R Calculation scores. Neither model accounted for significant variation in grade 3 WJ-R Calculation scores, adjusted $r^2 < .02$, $P > .10$.

4. Discussion

Our research questions examined primary school aged students' spontaneous comments about mathematics and specifically tested whether these comments reveal emerging dispositions linked to students' later mathematics achievement. Additionally, we tested the notion that beliefs about math may differ among children with versus without MLD. For comparative purposes, children's spontaneous comments about reading were also obtained. Our approach involved a straightforward design to elicit spontaneous comments of the kind that may emerge during casual verbal dialogue between young children and their parents or teachers. Finally, by repeating this procedure during two consecutive grades, we were able to test short term stability or reliability of the codes we collected.

Despite the subtlety of our approach, our findings indicate that children's spontaneous comments about math are informative even if they are no substitute for a structured and validated assessment of disposition. Specifically, we found that subject-specific comments about likability (or difficulty) appear to be stable beliefs, at least from grade 2 to 3. Lower likability ratings (i.e., more negative comments) were evident in definitions of math (versus reading) among children with MLD, and, to a lesser degree, among children with LA in math; but overall, reference to liking or disliking mathematics was absent from most of the participants' spontaneous comments of math and variability was greater among definitions of reading.

With regard to *difficulty*, children made more references to difficulty in definitions of math versus reading, collapsed

across grades, groups, and gender. Children rarely, if ever, spontaneously remarked that math was easy but this was also true of reading. Children with MLD were more likely to make references to difficulty of math than their peers, but they also made more references to reading as difficult. This finding is consistent with reports of high co-occurrence of math and reading difficulties [36] and with reports that children with MLD recognize their academic weakness in mathematics, as determined by ratings on self-perception measures [37]. However, Hanich and Jordan [37] found that, on measures of reading, children with math difficulties and normal reading achievement do not rate their competence lower than children without math difficulties, although children with co-morbid math and reading difficulty do. They hypothesized that children in the math and reading difficulty group in their study were more accurate at evaluating their reading performance than math performance given that many of the children in this group were receiving special education services for reading related problems, which may have contributed to the formation of children's achievement-related beliefs.

Analyses related to *specificity* of children's definitions of math and reading showed that, in general, children report unelaborated conceptions regarding the usefulness or importance of both math and reading as a tool. The only contributions of MLD status to the *specificity* rating from children's definitions of math concerned slightly less specific definitions by the MLD and LA groups, relative to the TA group, more so for boys versus girls; but this finding also emerged from definitions of reading. Greater specificity in definitions for math and reading were also observed among girls versus boys, which may simply reflect girls' superior expressive vocabulary during the school age years.

Of the four main predictor variables examined across these analyses—MLD status, subject area, gender, and grade—the only variable to account for variability in all three outcome variables was MLD status. Main effects or interactions involving MLD status emerged for likability, difficulty, and specificity. Although it is erroneous to infer that brief, open-ended questions like the ones used in our study are appropriate for diagnosis of MLD, future and ongoing work is needed to address the dynamic role(s) of a productive disposition and long term math achievement outcomes.

There were no gender differences in likability ratings for math or reading, but boys and girls reported slightly greater difficulty for math than reading at both grades. From grades 2 to 3, boys shifted towards reporting slightly greater difficulty for both math and reading over time, whereas girls shifted towards reporting slightly less difficulty for math over time. This may reflect differences in shift towards behavioral compliance rather than a specific shift in beliefs about mathematics, but neither explanation is supported by our data. The effect was small and warrants more in depth evaluation, especially in light of research that has found primary school aged boys identify more strongly with mathematics and have higher self-concepts than girls despite similar levels of mathematics achievement [22]. Other research conducted with eighth graders found an absence of gender differences

in attitudes towards mathematics and achievement, although boys were more likely to attribute their math achievement to their intellectual abilities than girls [38].

It is quite clear that the effects that emerged in this exploratory study are small. Yet the fact that *any* such findings would emerge given the simplistic nature of our questioning is somewhat intriguing. Our ratings were based on children's unstructured responses to brief, open-ended questions, "what is math?" and "what is reading?" and their definitions hint at potentially distinct elements of a productive disposition towards math, evident at second and third grades. This finding is consistent with factors to emerge from formal assessments based research. For instance, Adelson and McCoach [28, 39] found evidence of two factors from their math attitudes survey given to upper elementary students (grades 3 to 5), reflecting an "enjoyment of math" construct (similar to our "likability" codes) and a "math self-perceptions" construct (similar to both our "difficulty" and "specificity" codes). Their development of a structured survey supports the notion that individual differences in dispositions towards math exist and can be measured at grade 3 and above; we add that this work can likely be extended downward to grade 2 and below. As our work and that of others indicates, some dispositional factors may be math-specific ("math is difficult") whereas others may generalize to academic subjects.

4.1. Implications for Further Studies. Although our findings do not indicate causal pathways, they are consistent with evidence that primary school aged students' interests or beliefs about mathematics affect their achievement level or reflect risk status for future math outcome [2, 4]. The predictive contribution of grade 2 specificity ratings to math scores at grade 3 suggests associations emerging in early childhood. One possibility is that *opinions* about the importance or challenges of math precede and direct learning or performance success, as has been implicated in work with adults [40]; alternatively, earlier ability levels may influence emerging opinions, such that these features operate dynamically throughout schooling. There is some evidence that students' perceived competence ratings and ratings of enjoyment of mathematics are independent of achievement in the early years, but are related to math performance at the end of elementary school [41]. Our findings run counter to this conclusion, and thus support the contention that research on dispositional contributions to mathematics achievement must include studies of early childhood.

These findings in early childhood have implications for teachers, parents, and care providers of young children. In general, children in our study expressed relative flat affect with regard to their enjoyment of mathematics, did not hold very elaborate conceptions regarding the usefulness of mathematics, and attributed more difficulty to mathematics than reading. Despite the small effect size that emerged from these findings, these suggest potential negative outcomes when considered in the context of expectancy-value theory. According to the model, these dispositions run counter to facilitating engagement in mathematics activities given

the proportional relationship among components of the model (i.e., engagement = success expectancies \times task value). However, the “neutral” codes for likability were based on spontaneous comments, and children may have responded very differently to prompts eliciting remarks about liking or disliking mathematics. To foster adaptive motivational behaviors, parents and teachers should provide feedback to young children about the task values associated with mathematics, help students develop positive but accurate perceptions of their abilities, and teach children to effectively appraise learning to modify their achievement behaviors.

4.2. Limitations and Conclusions. There are several limitations to our exploratory study. Although our sample size was large and data collection over time was longitudinal, the time period over which we examined definitions of math and reading was limited to two consecutive years, and the number of students with MLD and LA was limited by the defining criteria of these constructs (e.g., MLD occurring in only ~6–10% of the population, as it was in our study). The deliberately simplistic nature of our data collection was consistent with our goal to evaluate more naturalistic and spontaneous versus prompted comments about mathematics likability, difficulty, and specificity, but children’s comments could be elaborated upon through structured probes and conversations. For instance, most children’s definitions received “neutral” codes for likability, but this does not mean that children did not have beliefs about liking or disliking mathematics.

Our research question did not concern the strength of children’s beliefs about math so much as the likelihood of their emergence during conversation. In addition to expressing (or not expressing) beliefs about math and reading, the children in this study remind us of the importance of listening to what they say, even during casual discourse. That is, adult-child discourse provides a means for parents and teachers to both nurture a child’s positive disposition towards math, and monitor the child’s emerging disposition.

Acknowledgments

This work was supported by a Grant from the Spencer Foundation awarded to M. M. M. Mazzocco and L. B. Hanich, based on data collected from a study supported by NIH Grant HD R01 34061 awarded to M. M. M. Mazzocco. The views expressed are solely those of the authors. They would like to thank the children who participated in the study, their parents and teachers, the staff at participating Baltimore County Public School elementary schools; and an anonymous reviewer of an earlier version of this paper. They also acknowledge the outstanding contributions of Gwen F. Myers, Project Manager for the longitudinal study, to this work and to the overall longitudinal research program.

References

- [1] National Research Council, *Adding It Up: Helping Children Learn Mathematics*, Mathematics Learning Study Committee,

- Center for Education, Division of Behavioral and Social Sciences and Education, National Academy Press, Washington, DC, USA, 2001.
- [2] S. L. Beilock, E. A. Gunderson, G. Ramirez, and S. C. Levine, “Female teachers’ math anxiety affects girls’ math achievement,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 5, pp. 1860–1863, 2010.
- [3] T. Kutaka, C. P. Edwards, and W. M. Smith, “The relationship between math competence beliefs and math ability scores in the early primary years,” in *Proceedings of the Annual Meeting of the American Education Research Association*, Vancouver, Canada, April 2012.
- [4] P. H. Fisher, J. Dobbs-Oates, G. L. Doctoroff, and D. H. Arnold, “Early math interest and the development of math skills,” *Journal of Educational Psychology*, vol. 104, no. 3, pp. 673–681, 2012.
- [5] J. S. Eccles and A. Wigfield, “In the mind of the actor: the structure of adolescents’ achievement task values and expectancy-related beliefs,” *Personality and Social Psychology Bulletin*, vol. 21, pp. 215–225, 1995.
- [6] A. Wigfield, “Expectancy-value theory of achievement motivation: a developmental perspective,” *Educational Psychology Review*, vol. 6, no. 1, pp. 49–78, 1994.
- [7] J. Garofalo, “Beliefs and their influence on mathematical performance,” *Mathematics Teacher*, vol. 82, pp. 502–505, 1989.
- [8] A. H. Schoenfeld, “Explorations of students’ mathematical beliefs and behavior,” *Journal for Research in Mathematics Education*, vol. 20, pp. 238–355, 1989.
- [9] J. S. Eccles and C. Midgley, “Changes in academic motivation and self-perceptions during early adolescence,” in *Advances in Adolescent Development: From Childhood to Adolescence*, R. Montemayor, G. R. Adams, and T. P. Gullotta, Eds., pp. 134–155, Sage, Thousand Oaks, Calif, USA, 1990.
- [10] S. Harter, “A new self-report scale of intrinsic versus extrinsic orientation in the classroom: motivational and informational components,” *Developmental Psychology*, vol. 17, no. 3, pp. 300–312, 1981.
- [11] J. H. Corpus, M. S. McClintic-Gilbert, and A. O. Hayenga, “Within-year changes in children’s intrinsic and extrinsic motivational orientations: contextual predictors and academic outcomes,” *Contemporary Educational Psychology*, vol. 34, no. 2, pp. 154–166, 2009.
- [12] A. E. Gottfried, J. S. Fleming, and A. W. Gottfried, “Continuity of academic intrinsic motivation from childhood through late adolescence: a longitudinal study,” *Journal of Educational Psychology*, vol. 93, no. 1, pp. 3–13, 2001.
- [13] J. Eccles, A. Wigfield, R. D. Harold, and P. Blumenfeld, “Age and gender differences in children’s self- and task perceptions during elementary school,” *Child Development*, vol. 64, no. 3, pp. 830–847, 1993.
- [14] J. E. Jacobs, S. Lanza, D. W. Osgood, J. S. Eccles, and A. Wigfield, “Changes in children’s self-competence and values: gender and domain differences across grades one through twelve,” *Child Development*, vol. 73, no. 2, pp. 509–527, 2002.
- [15] A. Wigfield, R. D. Harold, C. Freedman-Doan et al., “Changes in children’s competence beliefs and subjective task values across the elementary school years: a three year study,” *Journal of Educational Psychology*, vol. 89, no. 3, pp. 451–469, 1997.
- [16] D. J. Stipek and D. MacIver, “Developmental changes in children’s assessment of intellectual competence,” *Child Development*, vol. 60, pp. 521–538, 1989.

- [17] A. Wigfield and J. S. Eccles, "Expectancy-value theory of achievement motivation," *Contemporary Educational Psychology*, vol. 25, no. 1, pp. 68–81, 2000.
- [18] M. C. Steffens, P. Jelenec, and P. Noack, "On the leaky math pipeline: comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents," *Journal of Educational Psychology*, vol. 102, no. 4, pp. 947–963, 2010.
- [19] S. Lau, A. D. Liem, and Y. Nie, "Task- and self-related pathways to deep learning: the mediating role of achievement goals, classroom attentiveness, and group participation," *British Journal of Educational Psychology*, vol. 78, no. 4, pp. 639–662, 2008.
- [20] G. Ramirez, E. A. Gunderson, S. C. Levine, and S. L. Beilock, "Math anxiety, working memory and math achievement in early elementary school," *Journal of Cognition and Development*. In press.
- [21] A. G. Greenwald, D. E. McGhee, and J. L. K. Schwartz, "Measuring individual differences in implicit cognition: the implicit association test," *Journal of Personality and Social Psychology*, vol. 74, no. 6, pp. 1464–1480, 1998.
- [22] D. Cvencek, A. N. Meltzoff, and A. G. Greenwald, "Math-gender stereotypes in elementary school children," *Child Development*, vol. 82, no. 3, pp. 766–779, 2011.
- [23] E. A. Gunderson, G. Ramirez, S. C. Levine, and S. L. Beilock, "The role of parents and teachers in the development of gender-related math attitudes," *Sex Roles*, vol. 66, pp. 153–166, 2012.
- [24] M. P. Levpušček and M. Zupančič, "Math achievement in early adolescence: the role of parental involvement, teachers' behavior, and students' motivational beliefs about math," *Journal of Early Adolescence*, vol. 29, no. 4, pp. 541–570, 2009.
- [25] J. Tiedemann, "Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school," *Journal of Educational Psychology*, vol. 92, no. 1, pp. 144–151, 2000.
- [26] R. Bhanot and J. Jovanovic, "Do parents' academic gender stereotypes influence whether they intrude on their children's homework?" *Sex Roles*, vol. 52, no. 9-10, pp. 597–607, 2005.
- [27] P. M. Frome and J. S. Eccles, "Parents' influence on children's achievement-related perceptions," *Journal of Personality and Social Psychology*, vol. 74, no. 2, pp. 435–452, 1998.
- [28] J. L. Adelson and D. B. McCoach, "Development and psychometric properties of the Math and Me survey: measuring third through sixth graders' attitudes towards mathematics," *Measurement and Evaluation in Counseling and Development*, vol. 44, pp. 225–247, 2011.
- [29] M. M. M. Mazzocco and G. F. Myers, "Complexities in identifying and defining mathematics learning disability in the primary school-age years," *Annals of Dyslexia*, vol. 53, pp. 218–253, 2003.
- [30] H. Ginsburg and A. Baroody, *Test of Early Mathematics Ability, PRO-ED*, Austin, Tex, USA, 2nd edition, 1990.
- [31] M. M. Murphy, M. M. M. Mazzocco, L. B. Hanich, and M. C. Early, "Cognitive characteristics of children with Mathematics Learning Disability (MLD) vary as a function of the cutoff criterion used to define MLD," *Journal of Learning Disabilities*, vol. 40, pp. 467–487, 2007.
- [32] R. W. Woodcock and M. B. Johnson, *Woodcock-Johnson PsychoEducational Battery-Revised. Tests of Achievement*, Riverside, Ill, USA, 1989.
- [33] R. W. Woodcock and N. Mather, *Woodcock Johnson Psycho-Educational Battery Revised, Tests of Achievement, Examiner's Manual*, Riverside, Chicago, Ill, USA, 1990.
- [34] R. L. Thorndike, E. P. Hagen, and J. M. Sattler, *Guide for Administering and Scoring the Fourth Edition Stanford Binet Intelligence Scale*, Riverside, Chicago, Ill, USA, 1986.
- [35] D. Wechsler, *Wechsler Abbreviated Scale of Intelligence (WASI)*, The Psychological, San Antonio, Tex, USA, 1999.
- [36] L. B. Hanich, N. C. Jordan, D. Kaplan, and J. Dick, "Performance across different areas of mathematical cognition in children with learning difficulties," *Journal of Educational Psychology*, vol. 93, no. 3, pp. 615–626, 2001.
- [37] L. B. Hanich and N. C. Jordan, "Achievement-related beliefs of third-grade children with mathematics and reading difficulties," *Journal of Educational Research*, vol. 97, no. 5, pp. 227–233, 2004.
- [38] S. N. Geogiou, P. Stavrinides, and T. Kalavana, "Is Victor better than Victoria at maths?" *Educational Psychology Review in Practice*, vol. 23, pp. 329–342, 2007.
- [39] J. L. Adelson and D. B. McCoach, "Measuring the mathematical attitudes of elementary students: the effects of a 4-point or 5-point likert-type scale," *Educational and Psychological Measurement*, vol. 70, no. 5, pp. 796–807, 2010.
- [40] M. Cadinu, A. Maass, A. Rosabianca, and J. Kiesner, "Why do women underperform under stereotype threat? Evidence for the role of negative thinking," *Psychological Science*, vol. 16, no. 7, pp. 572–578, 2005.
- [41] P. Kloosterman and M. C. Cougan, "Students' beliefs about learning mathematics," *The Elementary School Journal*, vol. 94, pp. 375–388, 1994.

