Math–Gender Stereotypes in Elementary-School Children

Dario Cvencek and Andrew N. Meltzoff

Institute for Learning & Brain Sciences

University of Washington

Anthony G. Greenwald

Department of Psychology

University of Washington

Abstract

247 children in grades 1–5 (126 girls and 121 boys) completed Implicit Association Tests and parallel self-report measures assessing the association of: (a) *me* with *male* (gender identity), (b) *male* with *math* (math–gender stereotype) and (c) *me* with *math* (math self-concept). Three new findings emerged. First, elementary-school children have absorbed the cultural stereotype that math is for boys. Second, boys' self-concepts differ from girls'—boys identify with math more strongly than do girls. Third, implicit data showed that principles of cognitive balance demonstrated in adults (Heider, 1946) also operate in young children. The findings suggest that the math–gender stereotype develops early and influences math self-concepts prior to ages at which there are actual differences in math achievement.

KEYWORDS: Social Cognition, Stereotypes, Self-Concept, Gender Identity, Mathematics, Sex Differences

Math–Gender Stereotypes in Elementary-School Children

Imagine yourself an elementary-school teacher. One of your female students fails to complete an arithmetic assignment and offers an excuse that "Girls don't do math." What might be a pretext for avoiding homework could also be the outcome of social-cognitive development. Combining prevailing cultural *stereotypes* ("Math is for boys") with the knowledge about one's own *gender identity* ("I am a girl") to influence one's *self-concept* ("Math is not for me") reflects the tendency to achieve what social psychologists (Heider, 1946) call *cognitive balance*.

Social knowledge can be represented as a network of interconnections among concepts (Greenwald et al., 2002). In the foregoing example, three aspects of social cognition are involved. The first is the association between *math* and *boy* or *girl*. If this takes a societally characteristic form (e.g., *math* = *boy*), it can be called a *math–gender stereotype.* The second involves *gender identity*, defined as the association between *me* and either *boy* or *girl*. The third is a *math self-concept*, the association between *self* and *math*.

The interplay among math–gender stereotype, gender identity, and math self-concept has been studied in children chiefly using self-report measures. American elementary-school children often reflect the stereotypic pattern for academic self-concepts: For math, girls rate their own ability lower than boys (Fredericks & Eccles, 2002), but do not do so for reading or spelling (Herbert & Stipek, 2005; Heyman & Legare, 2004). This pattern is evident as early as the first grade (Entwistle, Alexander, Pallas, & Cardigan, 1987), even in the absence of differences in math achievement (Herbert & Stipek, 2005). Girls' weaker identification with math may derive from culturally communicated messages about math being more appropriate for boys than for girls (Dweck, 2007; Eccles, 2007; Guiso, Monte, Sapienza, & Zingales, 2008; NSF, 2006; Steele, 2003). These patterns are important developmentally, because as Eccles and others have

shown, children have reduced interest in future academic courses and occupations that are incompatible with their academic self-concept (Denissen, Zarrett, & Eccles, 2007; Frome, Alfeld, Eccles, & Barber, 2006; Killen, Sinno, & Margie, 2006; Liben, Bigler, & Krogh, 2001; Malcom et. al, 2005; Newcombe, 2007).

Previous investigations of children's math–gender stereotype and math self-concept have focused on self-report measures (for an exception, see Ambady, Shih, Kim, & Pittinsky, 2001). The wording of self-report measures often involves asking children how *good* they think they are at something or how much they *like* it, both of which conflate self-concept (non-evaluative association of self) with self-esteem (evaluative association of self). For example, a girl who reports that she is *good* at math may do so because she thinks that she is good at many things (high self-esteem). Similarly, a boy who reports that he *likes* math may do so because he believes that liking math is a positive quality and he sees himself as having many positive qualities. If the focus is children's math self concepts, it is more informative to assess how strongly a child associates *self* with *math* (i.e. whether the child has a strong math self-concept or not), because measures of association strength often conform to the pattern expected for cognitively balanced structures of social cognition (Greenwald et al., 2002). In order to differentiate the constructs more cleanly, we adapted a test used in social psychology that does not require self-report, the Implicit Association Test (IAT) (Greenwald, McGhee, & Schwartz, 1998). The IAT originated within social psychology, but in recent years has been applied in cognitive psychology (Fazio & Olson, 2003), clinical psychology (Teachman, Gregg, & Woody, 2001), and developmental psychology (Dunham, Baron, & Banaji, 2006; Skowronski & Lawrence, 2001). In adults, IAT measures correlate with actual math performance and real-world choices and actions (Greenwald, Poehlman, Uhlmann, & Banaji, 2009).

The IAT is a computerized categorization task that measures relative strengths of associations among concepts. IAT's format allows the measurement of preference for one concept (e.g., math) relative to the preference for a second concept (e.g., reading). The contrasting category is of practical importance in investigations involving academic subjects, because academic choices rarely occur without alternatives. Reading offers itself readily as a contrasting category for math because: (a) reading and math education are mandated from the first grade on, (b) sex differences in self-concepts have been demonstrated most often for math and reading, and (c) standardized tests across many countries have reading and math portions.

In addition to investigating implicit math–gender stereotype, gender identity, and math self-concept via a child IAT, we also examined explicit or self-reported counterparts in the same children. One motivation comes from research suggesting that stereotypes can be separated into two underlying processes—one automatic, unconscious, and implicit and the other controlled, conscious, and explicit (Devine,1989). In adults, positive but weak correlations are observed between implicit and explicit measures, especially in socially sensitive domains such as stereotypes (Hofmann, Gawronski, Gschwendner, & Schmitt, 2005). One of the explanations for this dissociation involves motivational influences: Implicit measures are assumed to be less susceptible to social desirability artifacts. It has also been suggested that early developmental experiences may shape implicit more than explicit cognition (Liben & Bigler, 2002; Rudman, 2004), again suggesting the value of using both implicit and explicit measures in the same study.

The present research also builds on a social psychological theory—Heider's (1946) balance theory—that has been used to explain how people are motivated to change their attitudes in order to achieve a balance between their perceptions of self, other people, and objects of mutual interest. Although balance theory has been used to show how perceptions of others can lead to

attitude change in adults (Eveland, 2002), it has not been systematically applied in child development research (but see Dunham, Baron, & Banaji, 2007). Within this theoretical framework, associative interconnections among concepts (corresponding to attitudes and stereotypes; see Greenwald et al., 2002) are assumed to self-organize in ways that reflect *cognitive balance*. For example, a boy who associates self with *male*, and also associates math with *male*, should be prone to develop the additional connection of *self* with *math*. Heider's theorizing about cognitive balance focused more on the role of cognitive balance in modifying existing links than on creating new links.

The present research is novel in two ways: (a) it examines the role of cognitive balance in creating new associative links and (b) it does so by examining the creation of new links in children during elementary school, which is an important time for the emergence and sculpting of the concepts under investigation.

One other study has examined cognitive balance in children. Dunham et al. (2007) have examined the development of self-esteem, group identity, and group attitude in Hispanic-American children. Their results are noteworthy in at least two respects: (a) they show support for the predictions of cognitive balance in disadvantaged minorities and (b) the strength of cognitive balance did not appear to vary with age. We return to both of these issues in our General Discussion section.

The three specific aims of our study were to: (a) design new measures of children's math–gender stereotypes and math self-concepts by adapting adult work from social psychology, (b) assess children's math–gender stereotypes and math self-concepts during elementary-school years, and (c) examine whether the cognitive balance described by Heider's (1946) theory

operates in children to link gender identity and math–gender stereotypes to emerging self-concepts.

Method

Participants

247 children (126 girls, 121 boys) from grades 1–5 were tested. All children were recruited through private and public elementary schools from the same school district in an urban area in the Northwest United States. All of the children were typical ages for their grade level (i.e. no child repeated a grade). The typical age ranges for children attending Grade 1 was 6 years of age; the typical age for a child attending Grade 5 was 10 years of age. At each grade, sample sizes ranged between 24–26 for boys and 24–25 for girls. According to parental report the children were 83.3% White, 9.6% Asian, and 7.1% African American. Families received $10 for participation.

Procedure

Children were tested individually while seated at a desk facing a computer (either a 43 or 48 cm screen). Each test session began with a 3–5 minute description of the study, during which children were familiarized with the test apparatus. The children were told that they would be “asked some questions” and then “play a computer game.” They were told that they would see and hear words during the game and would have to press a button to “let the computer know which word it is.” The procedure started with the administration of the self-report measures followed by the administration of the IATs.

Math-gender Stereotype Measures

Self-report. The self-report math–gender stereotype measure was administered as two Likert-scale questions using images from Harter and Pike’s (1984) Pictorial Scale. For each

question, children were shown two pictures of a child and responded by reporting: (a) which character (boy or girl) they believed possessed an attribute (e.g., liking math) to a greater degree, and (b) whether the character possessed the attribute "a little" or "a lot." This was done by their pointing to one of two circles (1.1 cm and 2.3 cm in diameter). One question requested selecting the boy or girl character as "liking to do math more." The other question requested selecting the boy or girl character as "liking to read more." All self-report questions were memorized by the experiment and said aloud to the children. The two scores were subtracted from one another to arrive at the explicit score with lower and upper bounds of –2 and +2; positive values indicated that the child picked the same sex character as liking to do math more. Appendix A provides a full list of all names used in self-report measures.[1]

Child IAT. We adapted the standard adult IAT (Greenwald et al., 1998) for use with children. An IAT score (*D*) (Greenwald, Nosek, & Banaji, 2003) is calculated by comparing the speed with which children categorize exemplars from four categories under two instructional conditions that vary assignments of the four categories to two computer response keys, one operated with the left hand and the other with the right hand. The measure is based on the principle that it is easier to give the same response to items from two categories if the two categories are mentally associated than if they are not. Figure 1 provides a pictorial representation of the child IAT.

During the math–gender stereotype IAT, children first practiced sorting girl and boy names. They responded to *girl* names (Emily, Jessica, Sarah, Ashley) by pressing a response button on the left side of the keyboard (i.e., "D") and to *boy* names (Michael, Andrew, David, Jacob) by pressing a response button on the right side of the keyboard (i.e., "K"). After that,

children practiced sorting *math* words (addition, numbers, graph, math) and *reading* words (read, books, story, letters) using the same two response buttons (Greenwald et al., 1998).

Following these two *single* discrimination tasks, children completed two *combined* discrimination tasks in which all four categories were used. During the combined tasks, two of the four categories were mapped onto the same response key. In one condition, *math* words and *boy* names shared one response key, with *reading* words and *girl* names sharing the other. The second condition switched the key assignments for the *math* and *reading* categories. All single discrimination tasks consisted of 16 trials and all combined tasks consisted of 24 trials. Positive scores indicated stronger association of *math* with *own gender* than with *opposite gender*.[2] Greenwald et al.'s (2003) scoring algorithm constrains the resulting *D* measure to have bounds of –2 and +2.[3]

The keyboard was furnished with two large buttons to replace the computers' "D" and "K" keys. Stickers with left-pointing and right-pointing arrows on those buttons indicated their use for left and right responses. To reduce need for reading, each stimulus word—spoken in a female voice—was synchronized with the onset of the written word on the screen. The intertrial interval was 500 ms. All words used as IAT stimuli were pre-tested with elementary-school children for familiarity and comprehension. To ensure that children understood each IAT task, error responses were followed by a red question mark appearing on the computer screen. After committing an error children could not advance to the next trial until they provided the correct response. As is standard in IAT procedures, trial latency was recorded to the correct response. Appendix A provides the list of all IAT stimuli.

Additional Measures: Gender Identity and Math Self-Concept

Self-report. Two additional self-report measures were administered as 2-item Likert-scales. The measure of gender identity asked children to select a boy or girl character as being more like themselves. Positive values indicated that the child picked the boy character. The math self-concept measure asked the child to choose whether a character of one's own gender solving a math problem was more or less like oneself than a same-gender character who was reading. Positive values indicated that the child picked the same sex character that was doing math. For the self-report measures, the order of the math–gender stereotype, gender identity, and math self-concept measures was counterbalanced across children. The order of characters assigned to left and right sides and the names used for each character were also counterbalanced across children. Order of administering self-report measures did not influence scores (all ps > .52) and was therefore not used as a design factor in analyses to be reported.

Child IAT. Two additional IAT measures were administered. During the gender identity IAT, children classified the words representing *me, not-me, boy,* and *girl*. In one instructional condition, *me* words and *boy* names shared a response key, with *not-me* words and *girl* names sharing the other response key. In the other instructional condition, two of the response assignments were reversed, such that *me* words and *girl* names shared one key while *not-me* words and *boy* names shared the other key. Positive scores indicated stronger association of *me* with *boy* than with *girl*. During the math self-concept IAT children classified the words representing *me, not-me, math,* and *reading*. In one instructional condition, *math* and *me* words shared a response key, as did *reading* and *not-me* words. In the other instructional condition, left versus right assignment of *me/not-me* words was reversed. Positive scores indicated stronger association of *me* with *math* relative to *reading*. For the implicit measures, there were 16

counterbalancing conditions. The gender identity IAT and math self-concept IAT were counterbalanced in the first and third position, with the math–gender stereotype IAT administered in the second position. Within each IAT, order of the two instructional conditions was counterbalanced. The spatial orientation of categories assigned to left and right was counterbalanced across subjects and IATs. Order of administration did not influence scores on any implicit measures (all *p*s > .66) and was therefore not retained as a design factor in analyses to be reported.

Internal Consistency

For implicit measures, Cronbach's alpha was calculated from two *D* measures computed for matched 24-trial subsets of each IAT. Cronbach's alpha coefficients for the math–gender stereotype, gender identity, and math self-concept IATs were $\alpha = .74$, $\alpha = .89$, and $\alpha = .78$ respectively. For the self-report measures, Cronbach's alpha coefficients for gender identity and math self-concept were $\alpha = .93$ and $\alpha = .79$ respectively. The two items of the self-reported math–gender stereotype scale measured two distinct constructs (gender stereotype towards math versus gender stereotype towards reading). Thus, the expectation was for low internal consistency of the self-reported math–gender stereotype measure, which was the case, $\alpha = .03$.

Data Reduction

Implicit measures ($N = 247$; See Footnote 1) were analyzed after excluding participants who met any one of three exclusion criteria: (a) 10% or more of their responses faster than 300 ms, (b) error rate of 35% or greater in at least one of the three IATs or (c) average response latency 3 *SD*s above the mean response latency for the whole sample in at least one of the three IATs. These criteria excluded 25 (10.1%) of participants. This was done to reduce noise in the data by excluding participants who would be identified as outliers on the basis of pre-established

criteria, consistent with the usual IAT procedures with adults (Greenwald et al., 2003). Self-report data ($N = 231$; See Footnote 1) were analyzed after excluding data from 11 participants (4.7%) due to excessively slow responding (either 30 seconds or more to respond to three or more self-report items or 90 seconds or more to respond to one or more of them). The analyses following data reduction provided increased power compared to analyses of the full sample, but the pattern of significant results and the conclusions drawn from them remained unchanged.

Results

How Well Did the Child IAT Work?

Previous literature has established gender identity using self-report measures in elementary school and younger (Martin, Ruble, & Szkrybalo, 2002; Slaby, & Frey, 1975). Evidence that our child IAT and self-report measures are effective derives from the fact that we obtained very orderly results that are in line with published findings. Figure 2 displays the results for both the implicit and self-report measures. As expected, boys associated *me* with *boy* more strongly than did girls on both the implicit measure, $t(220) = 15.35$, $p < .0001$, and self-report, $t(218) = 18.81$, $p < .0001$. The IAT and self-report measures of gender identity were strongly correlated, $r = .64$, $p < .0001$.

Math–gender Stereotype

On the implicit measure, boys associated *math* with *own gender* and girls did not, $t(220) = 6.46$, $p < .0001$. Similarly, on self-report measure boys were more likely to pick the same gender character as “liking to do math more” than were girls, $t(218) = 4.75$, $p < .0001$. These results mean that both boys and girls indicated stronger association of math with boys than with girls—evidence for math–gender stereotype. The overall implicit-explicit correlation for the math–gender stereotype measure, was positive but small, $r = .14$, $p < .05$.

Math Self-concept

There was evidence for gender-distinctive math self-concepts. On the implicit measure, boys associated *me* with *math* more than did girls, $t(220) = 2.63$, $p < .01$ and on the self-report measure boys identified more with a picture of a same gender character who was solving a math problem than did girls, $t(218) = 3.31$, $p < .01$. The overall implicit-explicit correlation for math self-concept measure was also positive but small, $r = .28$, $p < .0001$.

Cognitive Balance

Building on Heider's (1946) balance theory, Greenwald and colleagues provided a rigorous statistical analysis to test for *balanced identity* (Greenwald et al., 2002). Greenwald et al.'s balanced identity analysis method was used here to examine whether the interrelations among math–gender stereotype, gender identity, and self-concept reflected the theorized configuration. The theoretical expectation is that, if a balanced configuration exists among three constructs, any one of the three constructs should be predictable from the multiplicative product of measures of the other two constructs. In other words, when two concepts are associated with the same third concept, the association between those two concepts should strengthen (Greenwald et al., 2002). This balanced identity configuration was strongly confirmed in the data for implicit measures but not for self-report measures. The strength of cognitive balance with implicit measures increased as a function of school grade (see below for details).

Statistical tests of the balanced identity involve a two-step hierarchical linear regression. The measure of each of three constructs was predicted solely from the multiplicative product of the other two in Step 1, with the two other variables added individually in Step 2 (Greenwald et al., 2002). Greenwald et al. (2002) described a sequence of 4 tests to evaluate a *pure multiplicative* model (which asserts that the multiplicative product of two variables is the sole

predictor of some effect). This method deliberately bypasses the standard statistical procedure of testing a product term after first entering its component variables as predictors. Justification of this procedure as a method for testing a pure multiplicative theoretical model has been given in detail by Greenwald et al. (2002) and Greenwald, Rudman, Nosek, and Zayas (2006).

With three measures (math–gender stereotype, gender identity, and math self-concept) this analysis can be done with each of the three as criterion in the 2-step hierarchical regression and the other two as multiplicative predictors. Each of these three regressions provides opportunities to pass the 4 tests to confirm the theoretically expected pattern: (a) the Multiple *R* should be statistically significant at Step 1 and the regression coefficient of the product term (b_1) should be numerically positive; (b) the value of b_1 should remain numerically positive at Step 2; (c) the increase in criterion variance explained at Step 2 should not be statistically significant; and (d) neither regression coefficient associated with the individual predictors (b_2 *and* b_3) should differ from zero at Step 2 (see Greenwald et al., 2002 for the logic and rationale for this analytic approach).

For implicit measures, all three analyses satisfied these four criteria, conforming to expectations from the pure multiplicative model. Parallel tests using the self-report measures passed the 4-criterion test only for one of the three measures (gender identity) as criterion. The principles of cognitive balance in children thus appear to be stronger on implicit measures, similar to the general pattern of multiple previous tests using adult data (Greenwald et al., 2002).

To examine developmental changes in balanced identity, we conducted a set of three 3-step regressions, in which: (a) each criterion was first regressed on school grade (1–5) in Step 1, (b) the multiplicative product of other two predictors was added in Step 2, and (c) the product of grade with Step 2's product term was entered as a predictor in Step 3. This regression provides

two relevant statistical tests. Significance of Step 2's product term indicates that the balanced identity pattern is apparent even with the addition of school grade as a preliminary predictor. Significance of Step 3's 3-component product term indicates a linear developmental progression across school grades. For implicit measures, the Step 2 product term was statistically significant in all three tests using each of the three measures as criterion, $t(218) > 2.50$, $p < .01$. The Step 3 product term was positive in all three, but statistically significant in only two of the three tests, $t(218) > 2.03$, $p < .05$.

Developmental Order of Emergence

To obtain adequate statistical power for examination of developmental change, pairs of adjacent grades were combined. Figure 3 shows that gender identity was robustly evident at the earliest grades on both implicit and self-report measures. Implicit and self-reported math–gender stereotypes were also evident throughout the developmental sequence. The developmental changes in math self-concepts are evident in Figure 3. As shown, there are different developmental patterns for implicit versus self-report measures for math self-concepts. Implicit math self-concept showed decreasing sex difference with increasing grade; self-reported math self-concept showed increasing sex difference with increasing grade.

Analyses testing for the presence of sex differences at each grade level showed that, for implicit gender identity, boys associated *me* with *boy* more strongly than girls did in Grades 1–2, $t(83) = 10.32$, $p < .0001$, and similar t tests assessing boys versus girls remained highly significant (see Figure 2) through increasing grades (all ps < .0001). Self-report gender identity showed the same pattern: Boys were more likely to pick a boy character as being more like themselves than were girls in Grades 1–2, $t(87) = 12.42$, $p < .0001$, and similar t tests remained highly significant at later grade levels (all ps < .0001). The finding of clear evidence for gender

identity on both implicit and self-report measures is useful in showing that, even at the earliest grades examined in this study, children could follow directions for both of these measures and were responding in an orderly fashion.

Sex differences indicating the presence of math–gender stereotype were also apparent for Grades 1–2, although weaker than gender identity. On the implicit measure, boys associated *math* with *own gender* more strongly than girls did in Grades 1–2, $t(83) = 3.91$, $p < .001$, and similar t tests assessing boys versus girls remained significant at each adjacent 2-grade level thereafter (all ps < .001). For self-report measures, boys were more likely than girls to pick the same sex character as liking to do math in Grades 1–2, $t(87) = 2.66$, $p < .01$, and this remained stable for subsequent grade levels (all ps < .05).

The data for math self-concepts suggests that it may emerge later than the other two constructs. On the implicit measure of math self-concepts, boys associated *me* with *math* more strongly than did girls in Grades 1–2, $t(83) = 2.30$, $p < .05$. This sex difference remained significant in Grades 2–3, $t(87) = 2.85$, $p < .01$, but was not significant in later grades (ps >. 23). For the self-report measure of math self-concepts, girls were more likely to pick a character who was reading as being more like themselves in Grades 1–2, but this was not statistically significant ($p > .14$). Evidence for sex differences on the self-report measure first became statistically significant in Grades 4–5, $t(87) = 2.52$, $p < .05$.

Discussion

In this study of elementary-school children, we distinguished between math–gender stereotypes and math self-concepts using both implicit and explicit measures in the same children. The findings confirm that our child IAT (and self-report) procedures are effective inasmuch as they provide the expected evidence of gender identity. The methods allow us to

uncover three new findings. First, the math–gender stereotype previously found to be pervasive in North American samples was found in elementary-school children on both implicit and self-report measures. Second, elementary-school girls showed a weaker identification with math than boys on both implicit and self-report measures (math self-concept). Third, the implicit measures suggested that math–gender stereotypes combine with gender identity to shape children's strength of identification with math and that the expected cognitive balance increased in strength from first to fifth grade. These findings suggest that the math–gender stereotype develops early and differentially influences boys' versus girls' self-identification with math *prior* to ages at which differences in math achievement emerge.

Math–gender Stereotypes

The current demonstration of math–gender stereotypes during elementary school years extends previous work on this topic (Aronson & Good, 2003; Muzzatti & Agnoli, 2007). For example, in one study of stereotypes in elementary grades, Ambady et al., (2001) found that the activation of female identity (e.g. coloring a picture of a girl holding a doll) significantly impeded girls' performance on a subsequent math test. The children in that study did not explicitly report awareness of the cultural stereotype. We provide a potentially more direct way of measuring whether children have assimilated the cultural stereotype: We tested their explicit awareness of the stereotype that "math is for boys," and the results showed that both boys and girls explicitly subscribe to this view. Like Ambady et al., we found evidence that such stereotypes operate at an implicit level as well (but we did so with a different implicit measure—the IAT).

Conceptualizing the cultural stereotype as an association between *math* and *boy* addresses a central issue raised in the child stereotype literature (Signorella, Bigler, & Liben, 1993). It has

been proposed that self-reported gender stereotypes in children may indicate the mere *awareness* of stereotypes as opposed to personal *endorsement* of those stereotypes. We found that the implicit stereotype about math (i.e., *boys = math*) was only weakly correlated with the self-reported stereotype (i.e., "boys like to do math more"). This dissociation is most appropriately interpreted as dissociation between explicit and implicit stereotypes about math ability. However, a demonstration of such dissociation in children may inform the distinction between awareness and endorsement: Implicit measures may be tapping children's cultural *awareness* of stereotypes, whereas open-ended self-report measures may be more indicative of children's personal *endorsement* of those stereotypes. Further research is warranted exploring the development and interrelation between implicit and explicit knowledge of stereotypes in young children (Greenwald, & Nosek, 2008; Liben & Bigler, 2002).

Math Self-concept

The definition of math self-concept used in the current study differentiates children's identification with math from more global beliefs about themselves, such as self-esteem (Marsh, Craven, & Debus, 1991; Wigfield, Battle, Keller, & Eccles, 2002). This distinction is useful because domain-specific self-concepts (e.g., math self-concepts) may have different motivational and behavioral consequences than global beliefs about the self (Heyman & Dweck, 1998). For example, older children are more prone to make specific and stable attributions about their academic competence than younger children (Rholes, Newman, & Ruble, 1990; Ruble, & Dweck, 1995). Math self-concepts may therefore operate differently in older children than in younger children. Our methods may be useful for uncovering conditions under which children of different ages make dispositional attributions about themselves (e.g. about math competence) based on self-concepts, and how such attributions motivate performance (Blackwell,

Trzesniewski, & Dweck, 2007; Dweck, 1999; Heyman, 2008; Ruble & Dweck, 1995) and aspirations about the self in the future.

Developmental Changes

As expected, there was robust evidence for the presence of gender identity across the entire Grade 1–5 spectrum and significant evidence even as early as Grades 1–2 on both the implicit and explicit measures. These findings for gender identity are consistent with previous research (see Ruble & Martin, 1998 for a review). Moreover, these findings are useful because they establish that, even at the youngest grades we tested, the children could understand directions for implicit and self-report measures and provided interpretable data for both.

Sex differences indicating the presence of math–gender stereotype also emerged during Grades 1–2 (see Fig. 3), but were not as strong as observed for gender identity. The exact age of emergence of a math–gender stereotype is difficult to pinpoint in our data, but given the stronger magnitude of sex differences in Grades 1–2 for gender identity than for math–gender stereotype (see also Fig. 3), it appears that gender identity develops *before* Grades 1–2 and that math–gender stereotype emerges *after* gender identity. This time frame would be consistent with previous research on children's: (a) susceptibility to gender stereotypes about math in the first two elementary grades (Ambady et al. 2001) and (b) familiarity with the stereotypes associated with social identities in the early elementary school years (Bigler, Jones, & Lobliner, 1997; Signorella, Bigler, & Liben, 1993).

Sex differences indicating the presence of math self-concepts were not clear-cut and varied according to which measure, implicit or self-report, was used. The overall pattern of data indicates a later emergence of math self-concepts than gender identity and math–gender stereotypes (see also Fig. 3). The findings with the self-report measure are consistent with

previous research showing that sex differences in math self-concepts first emerge in middle to late elementary school (Herbert & Stipek, 2005; Muzzatti & Agnoli, 2007). It remains interesting that this was not apparent on the implicit measure. One possibility is that implicit measures may not always reflect the same underlying processes across development. More specifically, work in aging adults has suggested that older adults' slower speed of responding on cognitive tasks may inflate implicit effects on an Age Identity IAT (but not on an Age Attitude or Self-esteem IATs) (Hummert et al, 2002). Similar mechanisms could affect children's responses on an IAT. It is therefore possible that the IAT effects were overestimated for the implicit math-self concept measure in the younger children.[4]

Cognitive Balance

Based on the current results and previous research (Ruble, Martin, & Beerenbaum, 2006), we assume that gender identity emerges before the first grade. The interesting developmental question is how children's gender identity interacts with the prevailing cultural stereotypes. Two alternatives can be offered based on the current data: (a) Stereotypes may be acquired first and influence self-concepts or (b) early self-concepts may facilitate internalization of stereotypes. The first holds that children who strongly identify with their gender (strong gender identity) are more likely to internalize cultural stereotypes about their gender (math–gender stereotypes), which in turn influences their math self-concepts. Considered from the perspective of girls, this developmental sequence can be expressed as: Me = Girl; Girls≠ Math; *therefore* Me ≠ Math. The second alternative proposes that children with a strong gender identity and a given level of self-identification with math (math self-concept) are more likely to generalize/project their own math identification to others of their own gender (math–gender stereotype). This developmental sequence can be expressed as: Me = Girl; Me≠ Math; *therefore* Girls ≠ Math. We favor the first

alternative, because it is relatively implausible that the weaker effect produces the stronger one (see Figure 2). The current correlational data do not allow us to identify the actual causal mechanisms, and further research is needed on this point.

The current study replicated previous findings of cognitive balance in young children (Dunham et al., 2007). However, the two studies differ in the developmental patterns evident in children's cognitive balance on implicit measures: Our results indicate greater cognitive balance in older children, whereas the results of the Dunham et al. (2007) indicated no changes in the strength of cognitive balance across the ages tested (i.e., 5- to 12-year-olds). The different developmental patterns might result from the different topics under investigation (gender stereotypes versus racial stereotypes). Previous research has shown that children's negative racial attitudes towards an outgroup may decline with age (Dunham, Baron, & Banaji, 2006; Killen, Sinno, & Margie, 2006), whereas the cross-sectional work on gender stereotypes in children has shown gender stereotyping to increase with age (Killen, Sinno, & Margie, 2006; Martin, Wood and Little, 1990). Further research aimed is needed to explore developmental changes in cognitive balance among gender versus racial associations.

The results have implications for educational theory and practice. If principles of cognitive balance operate in children similarly to the way they do in adults, there are implications for children's academic development. In female college students, a balanced configuration of math–gender stereotypes, gender identity, and math self-concepts is associated with their negative attitudes towards mathematics and lower performance on the mathematical portion of the Scholastic Aptitude Test (Nosek, Banaji, & Greenwald, 2002). Societal stereotypes about math could begin to influence children's academic interests and performance well before college (Barron, 2004; Good, Aronson, & Inzlicht, 2003). In elementary school, boys and girls score

equally well on math-achievement tests (Hyde et al., 2008) and girls receive higher math grades (Kimball, 1989). Thus, the sex differences in math self-concepts detected by our tests *precede* rather than follow actual differences in math achievement, and may exert a developmental influence on children's interests, efforts, and aspirations, which could in turn affect achievement. Recent research in adults shows that perceiving an academic field to be at odds with one's identity leads to a sense of not 'belonging' and can deter people from pursuing that field (Cheryan, Plaut, Davies, & Steele, in press). Early differences in identification with math demonstrated here might contribute to how children think of themselves in the future and influence their thinking about who they could be and what they can do.

Where do children's stereotypes about academic subjects come from? Do children learn the stereotypes from their parents, school environments, the media, gaming, and/or primarily from interactions with peers? Such questions are becoming a focus in education and the learning sciences (Lee, 2008; Meltzoff, Kuhl, Movellan, & Sejnowski, in press; Nasir, 2002). Studies designed to address this issue will use ethnographic techniques following individual children in their everyday lives (Bell, Lewenstein, Shouse, & Feder, 2009). At a more global level, societies themselves provide a "natural experiment." Research over the past two decades has shown the pervasiveness of gender stereotypes about math in the U.S. (Killen, Sinno, & Margie, 2007; Liben & Bigler, 2002; Ruble et al., 2006; Wigfield, et al., 2006). A recent study revealed that national estimates of implicit gender–science stereotyping correlate with national sex differences in eight graders' performance on international science assessments, even after accounting for general indicators of societal gender equality (Nosek et al., 2009). We are planning a cross-cultural study using the experimental methods reported here. In Singapore girls score higher than boys on standardized math assessments and both sexes score higher than age/grade-matched

American children (Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald, 2008). In such a society, there may be absence or reversal of sex differences in children's math self-concepts on our tests.

Conclusion

In the present research, young girls showed a weaker identification with math than did age-matched male peers. Such gender imbalances in children's math-self concepts may arise from the early combination of societal influences (cultural stereotypes about gender roles) and intrapersonal cognitive factors (balanced cognitive organization). Future studies will profit from combining the tools and concepts from both social psychology and developmental science (Dunham & Olson, 2008; Olson & Dweck, 2008) to explore the development of academic identity and how it contributes to children's educational success.

APPENDIX A

Words for Self-report Measures

Boy: Michael, Jacob, Joshua, David, Andrew, Robert, Ryan, William.

Girl: Emily, Sarah, Jessica, Ashley, Lauren, Hannah, Rachel, Jennifer.

Words for Implicit Association Tests

Me: my, mine, I, myself.

Not-me: they, them, theirs, other.

Boy: Michael, Andrew, David, Jacob.

Girl: Emily, Jessica, Sarah, Ashley.

Math: addition, numbers, graph, math.

Reading: read, books, story, letters.

Footnotes

[1] These self-report measures were not administered to 16 of the 247 subjects, because they have not been developed at the same time as implicit measures.

[2] For some analyses, positive scores on the math–gender stereotype measures indicated stronger association of *math* with *male*.

[3] The implicit data were analyzed separately using two alternative approaches for computing the *D* measure by adding penalties to error trials (Greenwald et al., 2003): *D–600 ms* penalty as well as the *D–*2*SD* penalty measures. For all three IATs, the *D–600* and *D–2SD* were not statistically significant from the *D–as is* measure (all *p*s > .26). The *D–as is* measure is therefore used throughout the text.

[4] We are thankful to an anonymous reviewer for suggesting this alternative interpretation.

References

Ambady, N., Shih, M., Kim, A., & Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science*, *12*, 385−390.

Aronson, J., & Good, C. (2003). The development and consequences of stereotype vulnerability in adolescents. In F. Pajares, & T. Urdan (Eds.), *Adolescence and education: Vol. 2. Academic motivation of adolescents* (pp. 299–330). Greenwich, CT: Information Age Publishing.

Barron, B. (2004). Learning ecologies for technological fluency: Gender and experience differences. *Journal of Educational Computing Research, 31*, 1-36

Bell, P., Lewenstein, B., Shouse, A. W., & Feder, M. A. (Eds.). (2009). *Learning science in informal environments*. Washington. DC: National Academy Press.

Bigler, R. S., Jones, L. C., & Lobliner, D. B. (1997). Social categorization and the formation of intergroup attitudes in children. *Child Development, 68*, 530–543.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, 78, 246–263.

Cheryan, S., Plaut, V.C., Davies, P., & Steele, C.M. (in press). Ambient belonging: How stereotypical environments impact gender participation in computer science. *Journal of Personality and Social Psychology.*

Denissen, J. J. A., Zarrett, N. R., & Eccles, J. S. (2007). I like to do it, I'm able, and I know I am: Longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child Development*, *78*, 430–447.

Devine, P.G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psycholog*y, *56*, 5–18.

Dunham, Y., Baron, A.S., & Banaji, M.R. (2006). From American city to Japanese village: A cross-cultural investigation of implicit race attitudes. *Child Development*, *77*, 1268–1281.

Dunham, Y., Baron, A.S., & Banaji, M.R. (2007). Children and social groups: A developmental analysis of implicit consistency in Hispanic Americans. *Self and Identity*, *6*, 238–255.

Dunham, Y. & Olson, K. (2008). The importance of origins: Why cognitive development is central to a mature understanding of social psychology. *The Open Psychology Journal*, *1*, 59–65.

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia: Psychology Press.

Dweck, C. S. (2007). Is math a gift? Beliefs that put females at risk. In S. J. Ceci, & W. M. Williams (Eds.), *Why aren't more women in science? Top researchers debate the evidence* (pp. 47−55). Washington, DC: American Psychological Association.

Eccles, J. S. (2007). Where are all the women? Gender differences in participation in physical science and engineering. In S. J. Ceci, & W. M. Williams (Eds.) *Why aren't more women in science? Top researchers debate the evidence* (pp. 199–210). Washington DC: American Psychological Association.

Entwistle, D. R., Alexander, K. L., Pallas, A. M., & Cardigan, D. (1987). The emergent academic self-image of first graders: Its response to social structure. *Child Development*, *58*, 1190–1206.

Eveland, W. P., Jr. (2002). The impact of news and entertainment media on perceptions of social reality. In J. P. Dillard & M. Pfau (Eds.), *The persuasion handbook: Developments in theory and practice* (pp. 691–727). Thousand Oaks, CA: Sage.

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, *54*, 297–327.

Fredericks, J. A., & Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence: Growth trajectories in two male-sex-typed domains. *Developmental Psychology*, *38*, 519–533.

Frome, P. M., Alfeld, C. J., Eccles, J. S., & Barber, B. L. (2006). Why don't they want a male-dominated job? An investigation of young women who changed their occupational aspirations. *Educational Research and Evaluation*, *12*, 359–372.

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., & Brenwald, S. (2008). *Highlights from TIMSS 2007: Mathematics and science achievement of U.S. fourth- and eighth-grade students in an international context.* Washington, DC: U.S. Department of Education.

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Applied Developmental Psychology*, *24*, 645−662.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, *109*, 3–25.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.

Greenwald, A. G., & Nosek, B.A. (2008). Attitudinal dissociation: What does it mean? In R. E. Petty, R. H Fazio, & P. Briñol, (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 65–82). Hillsdale, NJ: Erlbaum.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197–216.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *97*, 17–41.

Greenwald, A. G., Rudman, L. A., Nosek, B. A., & Zayas, V. (2006). Why so little faith? A reply to Blanton and Jaccard’s (2006) skeptical view of testing pure multiplicative theories. *Psychological Review*, *113*, 170–180.

Guiso, L., Monte, F., Sapienza, P., & Zingales, L. (2008). Culture, gender, and math. *Science*, *320*, 1164–1165.

Harter, S., & Pike, R. (1984). The pictorial scale of perceived competence and social acceptance for young children. *Child Development*, *55*, 1969–1982.

Heider, F. (1946). Attitudes and cognitive organization. *The Journal of Psychology*, *21*, 107–112.

Herbert, J., & Stipek, D. T. (2005). The emergence of gender differences in children's perceptions of their academic competence. *Applied Developmental Psychology*, *26*, 276–295.

Heyman, G. D. (2008). Talking about success: Implications for achievement motivation. *Journal of Applied Developmental Psychology*, *29*, 361–370.

Heyman, G. D., & Dweck, C. S. (1998). Children's thinking about traits: Implications for judgments of the self and others. *Child Development*, *69*, 391–403.

Heyman, G. D., & Legare, C. H. (2004). Children's beliefs about gender differences in the academic and social domains. *Sex Roles*, *50*, 227–239.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin*, *31*, 1369–1385.

Hummert, M. L., Garstka, T. A., O'Brien, L. T., Greenwald, A. G., & Mellott, D. S. (2002). Using the Implicit Association Test to measure age differences in implicit social cognitions. *Psychology and Aging*, *17*, 482–495.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, *321*, 494–495.

Killen, M., Margie, N. G., & Sinno, S. (2006). Morality in the context of intergroup relationships. In M. Killen & J. Smetana (Eds.), *Handbook of moral development* (pp. 155-183). Mahwah, NJ: Erlbaum.

Killen, M., Sinno, S., & Margie, N. G. (2007). Children's experiences and judgments about group exclusion and inclusion. In R. Kail (Ed.), *Advances in child development and behavior* (Vol. 35, pp. 173–218). New York: Elsevier.

Kimball, M. M. (1989). A new perspective on women's math achievement. *Psychological Bulletin*, *105*, 198–214.

Lee, C. D. (2008). The centrality of culture to the scientific study of learning and development: How an ecological framework in education research facilitates civic responsibility. *Educational Researcher*, *37*, 267–279.

Liben, L. S., & Bigler, R. S. (2002). The developmental course of gender differentiation: Conceptualizing, measuring, and evaluating constructs and pathways. *Monographs of the Society for Research in Child Development*, *67*(2, Serial No. 269*).*

Liben, L. S., Bigler, R. S., & Krogh, H. R. (2001). Pink and blue collar jobs: Children's judgments of job status and job aspirations in relation to sex of worker. *Journal of Experimental Child Psychology*, *79*, 346–363.

Malcom S., Teich, A. H., Jesse, J. K, Campbell, L. A., Babco, E. L., & Bell, N. E. (2005). *Preparing women and minorities for the IT workforce: The role of nontraditional educational pathways*. Washington, DC: American Association for the Advancement of Science.

Marsh, H. W., Craven, R. G., & Debus, R. L. (1991). Self-concepts of young children 5 to 8 years of age: Measurement and multidimensional structure. *Journal of Educational Psychology*, *83*, 377–392.

Martin, C. L., Ruble, D. N., & Szkrybalo, J. (2002). Cognitive theories of early gender development. *Psychological Bulletin*, *128*, 903–933.

Martin, C. L., Wood, C. H., & Little, J. K. (1990). The development of gender stereotype components. *Child Development*, *61*, 1891–1904.

Meltzoff, A. N. (2007). 'Like me': A foundation for social cognition. *Developmental Science*, *10*, 126–134.

Meltzoff, A. N., Kuhl, P. K., Movellan, J. R., & Sejnowski, T. J. (in press). Foundations for a new science of learning. *Science*.

Muzzatti, B., & Agnoli, F. (2007). Gender and mathematics: Attitudes and stereotype threat susceptibility in Italian children. *Developmental Psychology*, *43*, 747–759.

Nasir, N. S. (2002). Identity, goals, and learning: Mathematics in cultural practice. *Mathematical Thinking and Learning*, *4*, 213–247.

National Science Foundation. (2003). *New formulas for America's workforce: Girls in science and engineering* (National Science Foundation Publication No. 03–207). Arlington, VA.

Newcombe, N.S. (2007). Taking science seriously: Straight thinking about sex differences. In S. J. Ceci & W. Williams (Eds.), *Why aren't more women in science? Top researchers debate the evidence* (pp. 69−77). Washington, DC: American Psychological Association.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology*, *83*, 44–59.

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., et al. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences*, *106*, 10593–10597.

Olson, K. R., & Dweck, C. S. (2008). A blueprint for social cognitive development. *Perspectives on Psychological Science*, *3*, 193–202.

Rholes, W. S., Newman, L. S., & Ruble, D. N. (1990). Understanding self and other: Developmental and motivational aspects of perceiving persons in terms of invariant

dispositions. In E. T. Higgins, & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Vol. II. Foundations of social behavior* (pp. 369–407). New York: Guilford.

Ruble, D. N., & Dweck, C. (1995). Self conceptions, person conceptions, and their development. In N. Eisenberg (Ed.), *Review of personality and social psychology: Social Development*. (Vol. 15, pp. 109–139). Thousand Oaks, CA: Sage.

Ruble, D. N., & Martin, C. L. (1998). Gender development. In N. Eisenberg (Ed.), *Handbook of child psychology. Vol. 3: Social, emotional, and personality development* (pp. 933-1016). New York: Wiley.

Ruble, D. N., Martin, C., & Berenbaum, S. A. (2006). Gender development. In N. Eisenberg (Ed.), *Handbook of child psychology: Vol. 3, Social, emotional, and personality development* (6th edition, pp. 858–932). New York: Wiley.

Rudman, L. A. (2004). Sources of implicit attitudes. *Current Directions in Psychological Science*, *13*, 79–82.

Signorella, M. L., Bigler, R. S., & Liben, L. S. (1993). Developmental differences in children's gender schemata about others: A meta-analytic review. *Developmental Review*, *13*, 147–183.

Skowronski, J. J., & Lawrence, M. A. (2001). A comparative study of the implicit and explicit gender attitudes of children and college students. *Psychology of Women Quarterly*, *25*, 155–165.

Slaby, R. G., & Frey, K. S. (1975). Development of gender constancy and selective attention to same-sex models. *Child Development*, *46*, 849–856.

Steele, J. (2003). Children's gender stereotypes about math: The role of stereotype stratification. *Journal of Applied Social Psychology*, *33*, 2587–2606.

Teachman, B. A., Gregg, A. P., & Woody, S. R. (2001). Implicit associations for fear-relevant stimuli among individuals with snake and spider fears. *Journal of Abnormal Psychology*, *110*, 236–235.

Wigfield, A., Battle, A., Keller, L. B., & Eccles, J. S. (2002). Sex differences in motivation, self-concept, career aspiration and career choice: Implications for cognitive development. In A. V. McGillicuddy-De Lisi, & R. De Lisi (Eds.), *Biology, society, and behavior: The development of sex differences in cognition* (pp. 93–124). Greenwich, CT: Ablex.

Wigfield, A., Eccles, J., Schiefele, U., Roeser, R., & Davis-Kean, P. (2006). *Development of achievement motivation.* In Eisenberg, N. (Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (6th ed., pp. 933–1002). New York: Wiley.

Figure Captions

Figure 1. For the IAT, items from four categories appear one at a time on a computer, and children respond by pressing a response button. In one task **(A)**, *math* words and *boy* names share a response key, as do *reading* words and *girl* names (*Stereotype Congruent*). In the other task **(B)**, these assignments are reversed—*math* is paired with *girl* (*Stereotype Incongruent*). Children with the math–gender stereotype (i.e. *boy* = *math*) should respond faster to the task **(A)** than **(B)**.

Figure 2. Sex differences for implicit **(A)** and self-report **(B)** measures in 1st–5th grade children. * = significant sex differences. Error bars = *SE*.

Figure 3. Developmental effects for implicit **(A)** and self-report **(B)** measures for gender identity, math–gender stereotypes, and math self-concepts of 1st–5th grade children. * = statistically significant sex differences. *N* = 222 for the implicit and *N* = 220 for the self-report measures. Error bars = standard errors.

Figure 1.

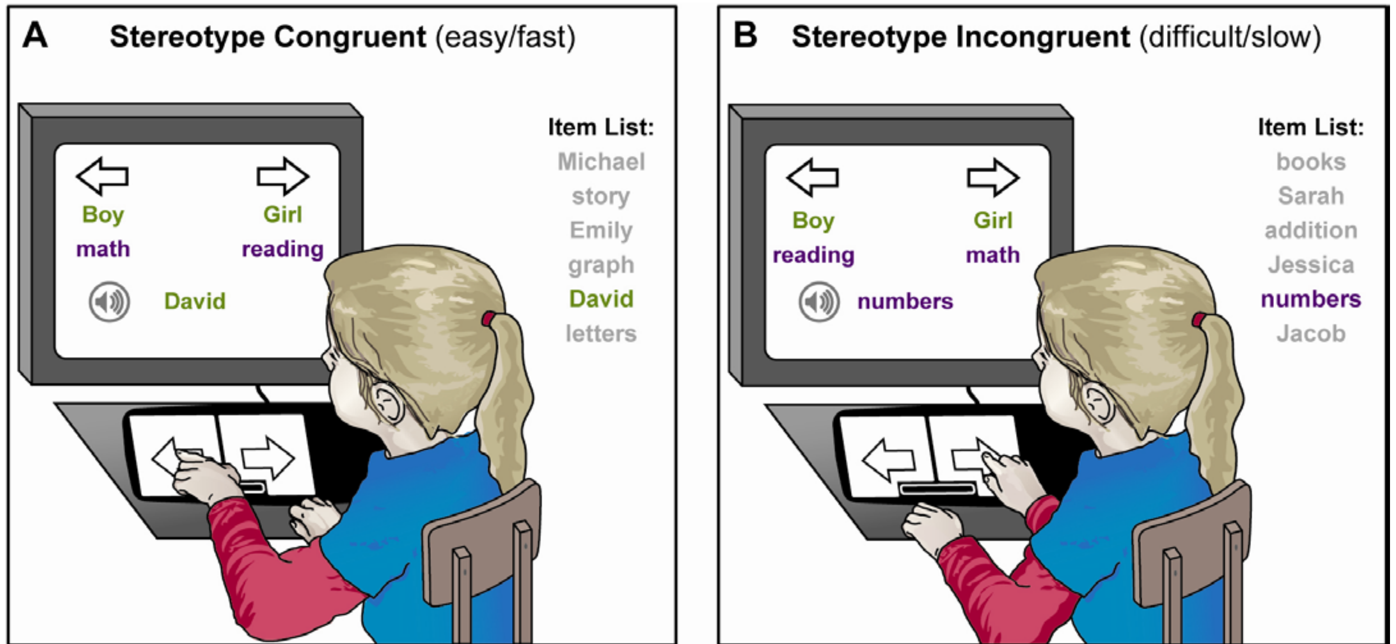

Figure 2.

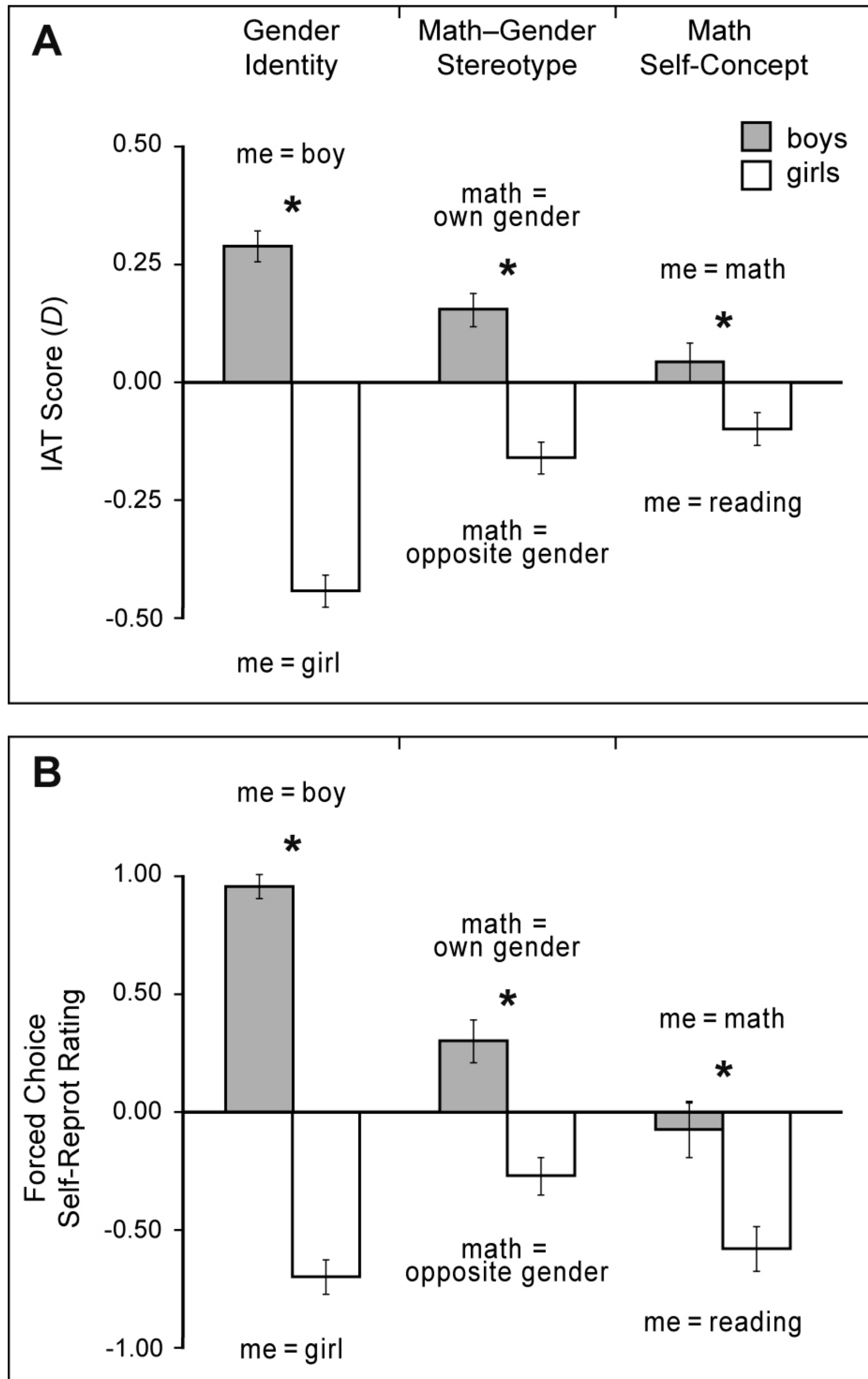

Figure 3.