HHS Public Access
Author manuscript
Q J Exp Psychol (Hove). Author manuscript; available in PMC 2018 April 01.

Published in final edited form as:
Q J Exp Psychol (Hove). 2017 April ; 70(4): 686–702. doi:10.1080/17470218.2016.1151060.

The Relationship Between Non-Symbolic Multiplication and Division in Childhood

Koleen McCrink,
Psychology Department, Barnard College, Columbia University

Patrick Shafto, and
Department of Mathematics and Computer Science, Rutgers - Newark

Hilary Barth
Psychology Department, Wesleyan University

Abstract

Children without formal education in addition and subtraction are able to perform multi-step operations over an approximate number of objects. Further, their performance improves when solving approximate (but not exact) addition and subtraction problems that allow for inversion as a shortcut (e.g., a + b − b = a). The current study examines children's ability to perform multi-step operations, and the potential for an inversion benefit, for the operations of approximate, non-symbolic multiplication and division. Children were trained to compute a multiplication and division scaling factor (*2 or /2, *4 or /4), and then tested on problems that combined two of these factors in a way that either allowed for an inversion shortcut (e.g., 8 * 4 / 4) or did not (e.g., 8 * 4 / 2). Children's performance was significantly better than chance for all scaling factors during training, and they successfully computed the outcomes of the multi-step testing problems. They did not exhibit a performance benefit for problems with the a * b / b structure, suggesting they did not draw upon inversion reasoning as a logical shortcut to help them solve the multi-step test problems.

Keywords

multiplication; division; inversion; cognitive development; memory; quantity; number

Organisms as varied as birds, fish, monkeys, rats, spiders, and humans exhibit an untrained ability to approximately represent the magnitude of a set of objects in their environment (Agrillo, Dadda, Serena, & Bisazza, 2009; Brannon, Wusthoff, Gallistel, & Gibbon, 2001; Cantlon & Brannon, 2007; Cordes, Gelman, & Gallistel, 2002; Meck & Church, 1983; Nelson & Jackson, 2012). Many theorists believe this capacity is supported by a cognitive mechanism that yields an inexact mental representation of the number of items in a set, a magnitude that is independent from both sensory modality (Barth, Kanwisher, & Spelke, 2003; Barth, La Mont, Lipton, & Spelke, 2005; Cantlon, Libertus, Pinel, Dehaene, Brannon,

Correspondence concerning this article should be addressed to Koleen McCrink, Department of Psychology, Barnard College, Columbia University, New York, NY 10027. kmccrink@barnard.edu, Phone: 212 – 854 – 8893, Fax: 212 – 854 – 3601.

& Pelphrey, 2009; Dehaene, 1997) and the spatial extent variables that are often confounded with numerical magnitude (Cordes & Brannon, 2009; Libertus, Starr, & Brannon, 2014; Park, DeWind, Woldorff, & Brannon, 2015). Other theorists argue that non-symbolic number representations (e.g., mental representations of numerical quantity that are distinct from the precise amount conferred by symbols such as Arabic numerals) are actually nonexistent as currently conceptualized, and instead are best conceived of as representations of spatial quantity (e.g., visual measurements of area, perimeter, density, or convex hull of the set of objects: Gebuis & Reynvoet, 2012a, 2012b; Szucs, Devine, Soltesz, Nobes, & Gabriel, 2013; Szucs, Nobes, Devine, Gabriel, & Gebuis, 2013). For any type of non-symbolic quantity, the ability to tell apart two sets is dependent upon the ratio of these sets, not by their absolute difference (e.g., Weber's law or Weber's ratio; Cordes, Gelman, & Gallistel, 2002). This reflects their approximate nature; the psychological distance between 10 and 20 units is equivalent to that between 100 and 200 units.

Approximate quantities are spontaneously used as operands in non-symbolic arithmetic computations, even before the onset of formal schooling (Barth, La Mont, Lipton, Dehaene, Kanwisher, & Spelke, 2006; Barth et al., 2005; Knops, McCrink, & Zitzmann, 2013; McCrink & Wynn, 2004, 2007, 2009; Xenidou-Dervou, van der Schoot, & Lieshout, 2014). For example, infants as young as nine months of age look longer to an incorrect outcome of a large-number addition or subtraction problem, such as when the sum of 5 objects and 5 objects is presented as 5 objects (McCrink & Wynn, 2004, 2009). Preschoolers in Barth et al. (2005, 2006) were able to add two arrays of objects and compare the magnitude of this estimated outcome to another amount; the success of this comparison process was predicated on the ratio of the two outcomes. The children in Barth et al. (2008) could also add an auditory set to a visual one with the same fidelity as adding two visual sets, indicating the presence of an approximate magnitude that was held in memory and updated as the operation progressed.

Children can also perform multi-step arithmetic over mentally represented quantities, in which multiple operations are strung together before a final outcome is computed. The storage and updating required for these complex operations may place a substantial load on working memory, and perhaps as a result, corresponding developmental changes are seen in the ability to perform multi-step operations. Infants - whose working memory is immature (Kibbe & Leslie, 2013; Moher & Feigenson, 2013; Rosenberg & Feigenson, 2013) - can perform one operation over a set but not two (Baillargeon, 1994; Moher, Tuerk, & Feigenson, 2012; Piantadosi & Aslin, under review), and their adding and subtracting is readily interrupted by in-scene variables such as seeing other objects that are tangential to the operation (Cheries, Wynn, & Scholl, 2006; Uller, Carey, Huntley-Fenner, & Klatt, 1999). At the age of five, without formal training in mathematics, children can perform multi-step mental approximate addition and subtraction successfully (Gilmore & Spelke, 2008). A recent study (Xenidou-Dervou, van Lieshout, & van der Schoot, 2014) experimentally manipulated the level and type of working memory load during non-symbolic addition in young children, and found that tasks which loaded upon the central executive had the greatest negative impact on performance, highlighting its key role in maintaining and updating numerical information.

The ability to perform scaling operations such as multiplication and division over non-symbolic representations of large numbers of discrete objects has recently been documented as well (Barth, Baron, Spelke, & Carey, 2009; McCrink & Spelke, 2010, 2015). These studies utilize a serial estimate-then-operate paradigm similar to those in Barth et al. (2005, 2006) and Knops et al. (2013): children are presented with one initial array of objects that is scaled up (multiplication) or down (division) in number of items via a "magical" event. For example, in a non-symbolic halving paradigm, a child would view an initial array of 20 dots transformed to 10, then 32 dots to 16, then 8 dots to 4. To test whether the child had successfully extracted this relationship from the presented examples, a new array would be presented, and the child would have to estimate the number of objects present (though unseen) following the new array's transformation. In some of these paradigms (e.g., McCrink & Spelke, 2010) the magical event is a joining or breaking of the initial array, in which each object breaks into distinct objects (e.g., one object breaks into 4 objects for a quartering operation); in this way they harness a many-to-one correspondence schema found to be helpful in learning symbolic multiplication and division (Nunes, Bryant, Burman, Bell, Evans, & Hallett, 2009; Nunes, Bryant, Evans, Bell, Gardner, Gardner, & Carraher, 2007; Park & Nunes, 2001). Children with no formal schooling in multiplication and division are able to double, halve, quadruple, quarter, and even multiply by an uneven factor of 2.5 (Barth et al., 2009; McCrink & Spelke, 2010, 2015).

A step beyond operating over non-symbolic magnitudes is the ability to appreciate the logical relationship between these operations (see Prather & Alibali, 2009, for a review). Indeed, appreciating mathematical logic is a critical first step of formal mathematical success (Nunes et al., 2007), and several studies have found that an appreciation of arithmetic logic for approximate, non-symbolic amounts emerges before children understand arithmetic logic in a symbolic context (Gilmore & Spelke, 2008; Kibbe & Feigenson, 2015). One such case is the understanding of the inverse relationship between arithmetic operations. Work on the understanding of the inverse relationship between addition and subtraction in an explicit, symbolic context with small numbers finds some evidence for rudimentary procedural understanding as early as the preschool years (Klein & Bisanz, 2000; Sherman & Bisanz, 2007; Vilette, 2002; Rasmussen, Ho, & Bisanz, 2003), with conceptual understanding developing well into middle childhood, likely as a result of formal training in arithmetic (Bisanz & Lefevre, 1990; Bryant, Christie, & Rendu, 1999; Gilmore, 2006; Siegler & Stern, 1998).

Young children are more likely to use the inversion principle to solve problems when they are presented without the use of symbolic numbers. Six-year-olds successfully benefitted from inversion when concrete items such as blocks were added to and subtracted from an array, and when no concrete referents ("invisible men") were used (Bryant et al., 1999). Even very young children are sensitive to the logic of inversion during addition and subtraction when invoking non-symbolic quantity representations (Gilmore & Spelke, 2008). Gilmore and Spelke (2008) presented preschoolers with a variety of complex, multi-step arithmetic problems, some of which lent themselves to being easily "solved" through inversion shortcuts. For example, a sufficiently sophisticated problem-solver can readily compute the answer to a problem such as 8 + 6 − 6 (a + b − b) without actually updating the initial operand. In contrast, the problem 8 + 4 − 6 (a + b − c), requires updating. The authors found that children, operating over approximate quantities, exhibited better performance for the subset of problems in which inversion could be employed as a logical strategy.

Work on inversion reasoning during symbolic multiplication and division suggests a disuse of, and lack of trust in, inversion processes until adulthood (Robinson & Dube, 2009; Robinson, Ninowski, & Gray, 2006). Even in adulthood, inversion logic is still used less frequently for multiplication and division than for addition and subtraction problems (Robinson & Ninowski, 2003; Robinson et al., 2006). In fact, adults and adolescents experience no carry-over benefit from understanding inversion in addition and subtraction as it applies to multiplication and division (Robinson & Ninowski, 2003; Robinson et al., 2006), indicating that for explicit, symbolic operations, the logic of inversion is not providing a broad conceptual foundation upon which to draw. This disconnect is surprising if one assumes that underlying arithmetic logic which transcends operations (such as inversion, or commutativity) should influence performance for all relevant problems.

The current set of experiments examines the level of competence, and logical reasoning, that children exhibit when performing multi-step non-symbolic scaling over estimated quantities. We systematically manipulate the number of transformations required during the scaling problem and, subsequently, the load on working memory (one step problems vs. two-step problems), as well as the availability of the inversion principle as a means to solve complex two-step problems (non-inversion problem sets vs. inversion-based problem sets). Children were trained separately in multiplication and division (one-step problems) and then required to combine these processes in two-step problems. Half of the children were trained on inverse scaling operations (e.g., they learned about quadrupling, and then quartering), and the other half were not (e.g., they learned about doubling and then quartering). The design includes problems that would be more difficult if children were using approximate representations, instead of calculating exact amounts (e.g., discriminating 8 from 4 vs. 28 from 24, problems of an equal difficulty level for exact calculation but very different difficulty levels for approximate calculation due to the difference in ratio of the correct outcome to a comparison value).

This paradigm allows us to achieve two main goals. First, we can determine whether children perform chained, serial, multi-step computations using the non-symbolic scaling ability previously demonstrated by Barth et al. (2009) and McCrink and Spelke (2010, 2015). We can determine the robustness of this untrained and computationally complex phenomenon in light of this population's still-developing working memory capacity. If children are able to hold these two operations simultaneously in mind, performance on multi-step testing trials should be above chance. If, on the other hand, the updating process of holding in memory and manipulating interim outcomes is too challenging, performance for the testing trials should be significantly worse than single-operation trial performance, and might not even rise above chance levels. A second goal is to address whether the previously documented lack of symbolic inversion reasoning is due to interference from explicit, symbolic, or rote processes rather than a more fundamental lack of access to inversion logic. If children can use inversion reasoning during addition and subtraction problems in the non-symbolic domain, perhaps this useful tool will be available for non-symbolic scaling as well. To do this, we implement here a design which allows us to compare performance for multi-step problems that are conducive to inversion reasoning to problems that are not. If children are able to apply an untrained, spontaneous logic of inversion to multiplicative reasoning, over approximate quantities, the children who view inversion problems (e.g., quadrupling then quartering) should perform better during testing than children who view multi-step problems in which inversion is not logically appropriate (e.g., quadrupling then halving). If, on the other hand, the documented disconnect between addition / subtraction inversion logic and multiplication / division logic is a true conceptual block that transcends numerical format, children's performance should be similar on inversion and non-inversion problems.

Experiment 1

Method

Participants—66 7-and 8-year-old children (33 females, 33 males; average age 7 years 2 months) were recruited via a large mailing database of New York City, Hartford CT, and surrounding suburbs and through partnerships with local New York City schools. The children had low (N = 10), middle (N = 41) and high (N = 15) SES, as assessed by proxy measures of school lunch participation rates and maternal education levels when noted by the parent.[1] Participants were screened for formal education in the concepts of multiplicative/divisive inversion relationships; none had prior formal schooling in inversion. The final sample included 16 females and 16 males in the Inversion conditions (M age 7 years; 3 mos) and 16 females and 16 males in the Non-Inversion conditions (M age 7 years; 2 months). Two additional participants were excluded from the final sample due to lack of concentration (1 subject), and experimenter error (1 subject).

Design and procedure—The experiment consisted of three main sections: multiplication training, division training, and multi-step testing (which combined multiplication and division). Children were randomly assigned to either an Inversion condition or Non-Inversion condition. In the Inversion condition, the child was shown movies depicting multiplication and division factors that were the inverse of each other (Doubling/ Halving; Quadrupling/ Quartering; see Figure 1). In the Non-Inversion condition, the child was trained on non-inverse factors (Doubling/ Quartering; Quadrupling/ Halving). There were 48 trials total, separated into 12 training trials for each operation (multiplication, division), and 24 test trials, all based on the paradigm used in Barth et al. (2009). Each trial consisted of an initial array on the left side of the screen, which was occluded and transformed by a particular factor (training: *2, *4, /2, /4), or factors (testing: *2 then / 2, *4 then /4, *4 then /2, *2 then /4), via a "magic wand", after which a comparison array appeared on the right side of the screen. The child was then asked to indicate which set of objects was more numerous (with the transformed array still occluded and never directly observed).

[1]Children whose school population qualified for free or reduced lunch at 75% or above rates were coded as low SES, 6–74% as middle SES, <5% as high SES (creating 3 larger bins from the 8-category measure often used in education research (e.g., Lubienski & Lubienski, 2006) and gathered on large-scale government questionnaires and used by the National Center for Education Statistics). Children whose maternal education levels consisted of high school graduate or below were coded as low SES, some college as middle SES, and college graduate or beyond as high SES (creating 3 bins from a 4-category measure commonly used by public health and education researchers who access large databases such as the California Maternal and Infant Health Assessment or MIHA (see Heck, Braveman, Cubbin, Chavez, & Kiely, 2006); the lowest two bins were collapsed due to infrequency in our sample).

The comparison arrays differed from the outcome by a distance factor of 2.0 or 1.5, to examine modulation of performance as a function of ratio (larger array: smaller array), a signature of approximate computation. For example, an initial array of 8 might be occluded and doubled via the transforming wand, and then a comparison array of 32 (the outcome of the doubling transformation, with a ratio of 2.0 comparison array : correct array) might appear on the other side of the screen. The correct answer in this case would be the comparison array. A different trial during multiplication training might show an array of 8, doubled, and then a comparison array of 12 (the outcome /1.5, a ratio of 1.5 correct array : comparison array). The correct answer in this case was the transformed outcome array. On 50% of the trials the comparison array was more numerous, and on the other 50% the transformed array was more numerous. 50% of the trials had a ratio of 2.0 (more distant ratios), 50% a ratio of 1.5 (closer ratios) of the transformed, computed outcome: comparison array; because the discriminability of two numbers represented approximately depends on their ratio (e.g. Cordes et al., 2002; Barth et al., 2003), we should observe better performance in larger ratio compared to smaller ratio trials. There were no procedural differences between conditions; the only change was the factor depicted during training for multiplication and division. Multiplication was always introduced first, followed by division.

Multiplicative factor introduction: The child and experimenter were seated together in a quiet testing room at a large table to watch the presentation videos on a Macintosh MacBook Pro laptop computer (17” display screen). The videos were animated using Keynote, with values and shape sizes obtained using the Numerus program (http://cogdevsoc.org/cds-tools/numerus). The child first viewed a video consisting of a single blue rectangle, which grew and shrank multiple times before becoming stationary. The rectangle remained on the screen as an animated wand appeared from off-screen left and waved back and forth over the rectangle while making a “magical” twinkling noise. After several seconds of waving, the wand made a “crack” noise, and the rectangle broke into two (or four, in the *4 condition) pieces of identical size. These newly formed rectangles then shrank and grew as their own individual entities, to instantiate the idea of multiple unique rectangles. The experimenter exclaimed to the child: “Look! The wand made more! It’s our magic multiplying wand. There used to be one rectangle, and now there are two(four). It doesn’t matter if the rectangles are big or small; the wand takes one rectangle and makes it two(four)!” The child then saw a follow-up video, which was identical to the previous video except that the rectangle was occluded by a large rectangle during the waving of the wand. The experimenter paused the movie after the wand waved, and asked the child how many rectangles s/he thought were behind the screen. When the child had correctly answered this question, training trials began. (See Table 1 for the exact values used during the testing and training sessions.)

Multiplication training block: The training block consisted of 12 video presentations in randomized order. In each video, the child saw an array of blue rectangles of varying sizes on the left side of the screen. In the first training video, the experimenter pointed to the blue array and said, “Now there are *this* many rectangles. There are too many to count, so we’re just going to use our imagination and our concentration; it’s not a counting game, it’s an imagination game and we just have to think really hard.” After three seconds, an occluder came up from off-screen and occluded the initial blue array. Three seconds was sufficiently short to stop children from explicitly enumerating the large sets depicted on-screen; any attempts to verbalize exact numbers were disrupted and the child was discouraged from counting. The multiplying wand came out from off-screen left and waved over the occluded array for two seconds, accompanied again by magical "twinkling" noises, and the experimenter said, "Look! Our magic multiplying wand!" After 2 seconds, a comparison array of pink rectangles appeared on the right side of the screen. These arrays contained rectangles of identical size (0.5 cm by 0.5 cm each) and were equated in density (~5 objects per 10 cm^2) throughout all training trials. The comparison array values had a numerical range of 5 – 96 (see Table 1), with corresponding spatial extent value ranges of ~1.25 – ~24 cm^2 area, ~10 – ~192 cm perimeter, 7 – 46 cm convex hull.

The child was then asked to indicate the set with more rectangles, either verbally ("pink", "blue") or by pointing to the right or left side of the screen. To control for experimenter bias, the experimenter phrased this question neutrally ("Where do you think there are more?" "Which side of the screen do you think has more?"). Additionally, the experimenter looked at the child, rather than the screen, until the child provided an answer. The experimenter recorded the child's response, resumed playing of the video to show the occluder drop to reveal the transformed array (e.g., the outcome after scaling up by a factor of two/four, as transformed by the multiplying wand), and provided feedback to the child as to whether his or her response was correct or incorrect. Outcome arrays were similar in area and density to the initial array and were formed by splitting each initially presented rectangle into two/four rectangles, with a 25% variability parameter around this average amount. In this way, the child received visual evidence that a major change between the initial array and revealed outcome array was number, rather than summed area, though the convex hull and perimeter also increased as the number of objects increased. (We aimed to emphasize numerical magnitude, not spatial quantity, as the target of the operation, in order to maximize similarity to previous inversion work that looked at addition and subtraction logic over amounts that were saliently numerical (e.g., Gilmore & Spelke, 2008.)) The outcome arrays ranged from a numerical value of 8–56 objects, 9–13 cm^2 area, 57 – 176 cm perimeter, and 41 – 45 cm convex hull. Overall, 12 training trial movies were shown, after which the division section of the experiment began.

Division factor introduction: This section of the experiment introduced the division concept to the child. The division videos were identical to the multiplication videos, except that a) the wand was different in color and shape, and was referred to by the experimenter as the "magic dividing wand," and b) the initial training video presented two rectangles growing and shrinking, and finally coalescing into one rectangle following the waving of the dividing wand. Although adults balk at the use of the term "dividing" to describe a joining of objects, the children readily accepted this term, likely because a) it was used minimally, and b) we were careful to highlight the fact that the dividing was happening over *number*, not area, saying the phrase "The wand made LESS... There used to be more, but now there are less.", even as the overall area stayed exactly the same. The procedure here was identical to the multiplicative introduction videos; the child first watched a video in which the entire division process proceeded unobscured, followed by a second video in which the rectangles were occluded and the child was asked how many s/he believed were behind the screen. Once the child correctly predicted the remaining number of rectangles, the experiment moved on to the division training videos.

Division training block: This training block consisted of 12 training videos. These videos were identical to those shown in the multiplication training block, except that the wand being shown was the green division wand, and the original blue arrays were scaled down, not up, in number. These videos were also randomized prior to presentation. Following the appearance of the pink comparison array, the experimenter still asked the child to indicate which side of the screen had more rectangles. The pink comparison arrays consisted of .5 × .5 cm squares with a density of ~5 objects per 10 cm 2. The comparison array values had a numerical range of 3 – 48 (see Table 1), with corresponding spatial extent value ranges of ~.75 – ~12 cm^2 area, ~6 – ~96 cm perimeter, 9 – 37 cm convex hull. The outcome arrays, constructed by taking the initial arrays and joining two (for /2) or four (for /4) ranged from a numerical value of 4–28 objects, 5–17 cm^2 area, 36–130 cm perimeter, and 40 – 46 cm convex hull. Following completion of all 12 training videos, the child and experimenter got up to take a quick break and stretch. The experiment then moved on to the testing block.

Multi-step testing block: The testing block consisted of 24 videos. In these videos, an array of blue rectangles, similar to those seen in the training blocks, were presented and then occluded after 3 seconds. Following occlusion, the child saw the multiplication wand wave over the occluded array and disappear off-screen left (a 2-second event), followed one second later by the division wand which waved over the array and disappeared off-screen left (a 2-second event). Following the appearance of both the multiplication and division wands, an array of pink rectangles appeared 3 seconds later on the right side of the screen. In a change from the training trials, after the child indicated which side of the screen s/he believed to have more shapes, the occluder did not drop to reveal the correct transformed amount beneath the screen. After the child answered, they were not told whether or not they answered correctly; instead, the experimenter provided uniformly positive feedback (“Good job! You’re doing great! Let’s try another”) and moved on to the next video. Furthermore, the comparison arrays did not look like those shown in the training blocks; rather, they were similar in terms of area, contour length, and convex hull, with the goal that the child would be attentive to numerical magnitude over these variables. That is, each set of comparison arrays within each multi-step testing problem (e.g., 8, 14, 18, 22, 28, or 40 *2/2 or *4/4) had similar area, perimeter, and convex hull (e.g., identical with some random variability programmed in), though the measurements across each problem differed. Density is challenging to hold steady across a range of small and large values, given placement constraints on large items, and if one holds convex hull steady across arrays density will vary with the number of objects. Thus, more numerous sets within a range of comparison array values tended to have greater density. If the child were responding on the basis of area, contour length, or convex hull instead of the numerosity of the comparison area during the training blocks (recall that these variables scaled up or down alongside the number of objects in the comparison array for training), they would perform only at chance during test trials, because in these test trial comparison arrays that information does not vary within a given scaling problem. For example, in the test videos of 14 * 2 / 2 = 7 vs. 14, 14 * 2 / 2 = 9 vs. 14, 14 * 2 / 2 = 21 vs. 14, and 14 * 2 / 2 = 28 vs. 14, the values for 7, 9, 21, and 28 objects were all 55–60 cm sq area, 54–56 cm^2 contour length, ~ 44–47 cm convex hull, and the density for the 7 object array was ~ one-quarter that of the 28 object array. Density did vary with number (more numerous arrays were denser), though this would be unlikely to be brought over as a strategy from the training trials, because in those training trials density was equated across all comparison arrays and thus could not be used to help solve the greater/fewer task at hand. Thus, the overall area and contour length of the comparison arrays were all equidistant from the (imagined, never-revealed) outcome arrays on these spatial extent variables. Again, take the example of the test trials 14 * 2 / 2 = 7 vs. 14, 14 * 2 / 2 = 9 vs. 14, 14 * 2 / 2 = 21 vs. 14, and 14 * 2 / 2 = 28 vs. 14. The comparison arrays had an area of 55–60 cm^2, 54–56 cm contour length, ~ 44–47 cm convex hull, and the density for the 7 object array was ~ 1 object per 10 cm^2 and for the 28 object array ~3.5 object per 10 cm^2. The (potentially imagined, based on what they saw at training) outcome array has an area of 9 cm^2, perimeter 82 cm, convex hull 45 cm, density ~2.25 objects per 10 cm^2. Though the children never saw the outcome arrays during the testing blocks, control of this factor discouraged the potential strategy of comparing the area or contour length of the imagined original outcome array to the presented comparison array. In this way, children were encouraged to attend to numerical values, rather than perceptual variables that vary along with number. The initial values used in the testing section were different from those in the training blocks, in order to avoid any performance boosts from rote memorization of the training trials (see Table 1 for the specific values used.)

Results

The average percentage correct for each participant was calculated and binned according to experimental session (Multiplication training, Division training, Multi-step Testing), and ratio of correct outcome array: comparison array (2.0, 1.5). These scores were entered into a repeated-measures ANOVA with within-subject factors of Session (Multiplication, Division, Multi-Step Testing) and ratio of correct array: comparison array (2.0, 1.5), and between-subjects factor of gender (male, female) and Inversion condition (Inversion, Non-Inversion). There was no significant main effect of, or interactions with, gender; this variable is not further included in the analyses. There was a significant main effect of Session ($F(2, 124) = 7.74$, $p = .001$, $\eta_p^2 = .11$); Multiplication (81%, $SEM = 1.8$) was significantly higher than Division (74%, $SEM = 1.7$; $p = .02$) and Multi-Step Testing (73%, $SEM = 1.6$; $p < .01$), which were comparable to each other ($p = 1.0$; see Figure 2). (These and all reported follow-up paired comparisons were Bonferonni-corrected for multiple comparisons.) There was also a significant main effect of Ratio ($F(1,62) = 46.61$, $p < .001$, $\eta_p^2 = .43$), with performance on Ratio 2.0 trials better than 1.5 trials (81% (SEM =1.3) vs. 71%, (SEM = 1.5), $p < .001$, 95% CI for the difference between these means [7, 13]). Using one-sample tests, we compared performance at each of the two ratio categories for each experimental session to chance performance (50%); all were significantly higher than would be expected by chance (Multiplication 2.0 86% ($SEM = 2.2$), Multiplication 1.5 77% ($SEM = 2.4$), Division 2.0 80%, ($SEM = 2.0$), Division 1.5 69% ($SEM = 2.4$), Test 2.0 79% ($SEM = 1.9$), Test 1.5 68% ($SEM = 1.8$, all ps < .001). There was no main effect of or interaction with the Inversion condition; children who received a testing session in which this process was available performed similarly to those who did not have this process available (74% (SEM =

1.7) vs. 73% (*SEM* = 1.7); $F(1,62) = .84$, $p = .36$). To assess the strength of the evidence in favor of the null hypothesis, we also computed the Bayes Factor (scaled JZS Bayes factor; Rouder et al., 2009). This analysis suggested that the evidence were more than three times more likely to have occurred under the null hypothesis ($B_{01} = 3.71$, using the default setting of the expected scale parameter for the effect size, .707, which is consistent with effect sizes observed in previous research, see Gilmore & Spelke, 2008).

A possible alternate strategy to solving problems in this type of paradigm is to observe whether the comparison array is especially high or low, and use this knowledge to decide whether the comparison array or transformed array is likely to be more numerous (e.g. see Barth, Beckmann, & Spelke, 2008; Barth et al., 2009; McCrink & Spelke, 2010). For example, a child who sees a comparison array of 4 objects appear could make an educated guess, based on the previous range of values shown, that 4 objects is a relatively low number and they should choose the transformed array; this strategy involves no actual multiplicative operations. To test for this possibility, we calculated performance for problems whose comparison arrays fell at the extreme end of values (the 3 highest, and lowest, values), and compared performance on these trials against performance on trials in which the comparison arrays were mid-range (the 6 middle values). The mid-range and extreme-range categories each contained equal numbers of more and less difficult ratio trials. For each experimental session, paired t-tests shows that extreme trial performance was better than mid-range trial performance (Multiplication, 88% (*SD* = 17) vs. 74% (*SD* = 20); Division, 83% (*SD* = 16) vs. 66% (*SD* = 20); Multi-Step Testing, 85% (*SD* = 14) vs. 61% (*SD* = 16); all *p*s < .001), indicating that children were likely attentive to the distribution of possible outcomes and using this information to guide their answers. However, this strategy could not explain children's successful performance entirely: for each experimental session, performance for both the mid-range and the extreme-range trials was better than predicted by chance (all *p*s < .001). To address the possibility that it is only on these challenging mid-range trials that children might reach for inversion-based reasoning, we entered the mid-range test trials percentage correct into a one-way ANOVA with Inversion condition as a between-subjects variable. There was no effect of inversion condition ($F(1,62) = .54$, $p = .47$.)

Discussion: Experiment 1

The children in this experiment were able to non-symbolically multiply and divide estimated numbers of objects. They were also able to combine these two different types of scaling to estimate the answer to a multi-step non-symbolic arithmetic problem. These computations are likely supported by a system which represents approximate discrete quantities; the scenarios presented emphasized number as the relevant to-be-scaled variable (with the spatial variables of area, contour length, density, and convex hull equated at critical junctures during numerical comparison), and performance varied as a function of the ratio of the correct outcome to the comparison array. Children's performance when stringing together two scaling operations was comparable to that of the more difficult one-step operation (division), indicating a capacity to serially update interim outcomes without introducing a significant amount of noise into the quantity representation.

In the multiplicative reasoning paradigm of this study, children did not show evidence of inversion-based reasoning such as that seen in non-symbolic addition and subtraction problems (e.g., Gilmore & Spelke, 2008). Children who sequentially witnessed a doubling and then a halving operation performed like children who saw a doubling and a quartering operation: they did not exhibit a "bonus" for invertible operations. One potential - albeit unlikely - reason for this pattern could be that children were not challenged by these problems; they were above-chance consistently even on the hardest ratios, and even when alternate strategies were not available. Perhaps children reach for strategies such as inversion when they feel challenged and feel that their usual computations may fail (Siegler & Schrager, 1987; Siegler, 1987). To test this hypothesis, we lowered the ratio values from Experiment 1 from 2.0 and 1.5 to 1.5 and 1.25, in order to reduce the discriminability of the transformed array and the comparison array. If problem difficulty prompts children to make a logical strategy choice, and inversion reasoning for scaling is in their mental toolbox, then this increase in task difficulty should make inversion reasoning more likely.

Experiment 2

Method

Participants—65 7-and 8-year-old children (37 females, 28 males; average age 7 years 6 months) were recruited via a large mailing database of New York City, Hartford CT, and surrounding suburbs and through partnerships with local New York City schools. The children had low (N = 39) or middle (N = 30) SES, as assessed by school lunch participation rates and maternal education levels when noted by the parent.[2] Participants were screened for formal education in the concepts of multiplicative/divisive inversion relationships; none were excluded due to prior formal schooling in inversion. One additional participant was excluded from the final sample due to lack of concentration and inability to finish the task. The final sample included 18 females and 14 males in the Inversion conditions (M age 7 years; 9 mos) and 19 females and 13 males in the Non-Inversion conditions (M age 7 years; 6 mos).

Design and procedure—All materials and design details were identical to Experiment 1, with the sole difference of the values chosen as the comparison arrays (see Table 1). The comparison arrays were now structured to have ratios of 1.5 and 1.25 (see Table 1 for the range of values used), the latter discrimination ratio causing difficulty even for some adults (Halberda, Ly, Wilmer, Naiman, Germine, 2012). The construction of the stimuli followed the same constraints as that of Experiment 1, with comparison arrays for the training sessions possessing equal density and item size (but convex hull, area, and contour length that varied alongside number of objects), and comparison arrays at test possessing equal area, contour length, and convex hull (but density and item size that varied with number of objects). In this way, if children had been attentive to those factors at one phase of the

[2]Because the children in Experiment 2 were drawn from a slightly less diverse population, with a lower mean SES, we ran preliminary analyses to determine whether there was an overall impact or significant interactions with SES as a dependent variable in the model, but there was not; children in low- (66%), middle- (68%) and high-SES (78%) groupings performed comparably ($F(2,108) = 2.01$, $p = .14$) with no interaction with SES and experiment with the critical variable of inversion condition ($F(1,108) = 2.6$, $p = .11$).

experiment they would be a neutral, uninformative, source of scaling information during the subsequent testing phase.

Results

As before, subjects' performance was calculated and binned according to experimental session and ratio of outcome : comparison array. These scores were entered into a repeated-measures ANOVA with Experimental Session (Multiplication, Division, Multi-Step Testing) and Ratio (1.5, 1.25) as within-subject factors, and gender (male, female) and Inversion condition as between-subject factors. There was a significant main effect of Session (F(2, 120) = 3.18, p = .045, η_p^2= .05); multiplication performance (73%, SEM = 3.2) was marginally higher than division performance (61%, SEM = 3.6,, p = .05, 95% CI of the difference between the means [−.6, 24]), and multi-step test performance was intermediate (68%, SEM = 3.1, both ps = .61; see Figure 2). There was a significant main effect of Ratio of correct outcome : comparison array (F(1, 60) = 8.96, p = .004, η_p^2 = .13); the children exhibited higher performance in ratio 1.5 (73%, SEM = 2.5) relative to ratio 1.25 (62%, SEM 2.8, p < .01, 95% CI [4, 19].) There was a significant main effect of gender (F(1,60) = 7.51, p = .01, η_p^2 = .11); males (72%, SEM = 2.9) performed better than females (62%, SEM = 2.5, p < .01, 95% CI of the difference between the means[3, 18].) Using one-sample tests, we compared performance at each ratio for each experimental session to chance performance (50%); all sessions, with the exception of the Division 1.25 (51%;, SEM = 4.1, p = .88), were significantly higher than would be expected by chance (Multiplication 1.5 77% (SEM = 3.8), Multiplication 1.25 66% (SEM = 4.8), Division 1.5 71% (SEM = 4.6), Division 1.25 51% (SEM = 5.2), Test 1.5 68% (SEM = 4.6), Test 1.25 65% (SEM = 4.5); all ps < .01). There was no significant main effect of Inversion condition (F(1,60) = .36, p = . 55); children in Inversion conditions had a score of 71% correct (SEM = 2.7), non-Inversion conditions 65% correct (SEM = 2.7). Analysis using Bayes Factors suggest that the observed data were more than three times more likely to have occurred under a model assuming the null hypothesis is correct (B_{01} = 3.43, assuming a scale parameter of .707). There was a three-way interaction between Inversion condition, Session, and gender (F(2,120) = 3.28, p = .04, η_p^2= .05). However, this is simply a reflection of baseline between-groups variation in the division training session for girls in the Inversion condition relative to the Non-Inversion condition. Girls who were subsequently placed into the Inversion condition (49%, SEM = 6.8) performed more poorly during division training than girls who were placed into Non-Inversion conditions (70%, SEM = 6.8; p = .04).

To test for the alternate strategy of using the extreme ends of the range of comparison array values, we calculated performance for problems whose comparison arrays fell at the extreme end of values (the 3 highest, and lowest, values). We then compared this performance via paired t-tests to performance on trials in which the comparison arrays were mid-range, and therefore not subject to this alternate strategy (the 6 middle values). The mid-range and extreme-range categories each contained equal numbers of more and less difficult ratio trials. For the multiplication and testing experimental sessions the end-range performance was higher than mid-range performance (Multiplication, 79% (SD = 31) vs. 64% (SD = 39); p = .02; Division, 65% (SD = 39) vs. 59% (SD = 40); p = .39; Multi-Step Testing, 80% (SD = 27) vs. 59% (SD = 38); p < .01), indicating that children were likely attentive to the distribution of possible outcomes and uses this information to guide their answers. To test whether this strategy could explain performance entirely, we used one-sample tests to compare both the mid- and end-range problem performance for each experimental session to that of chance; middle and end-range trials for multiplication (64% and 79%), and end-range trials for division (65%) and multi-step testing (80%), were significantly higher than 50% (all *p*s < .01). The mid-range trials for division (59%) and multi-step testing (59%) were marginally higher than chance (*ps* = .09, .06 respectively). To address the possibility that it is only on these challenging mid-range trials that children might reach for inversion-based reasoning, we entered the mid-range test trials percentage correct into a one-way ANOVA with Inversion condition as a between-subjects variable. There was no effect of inversion condition ($F(1,62) = 2.56$, $p = .12$.)

Discussion: Experiment 2

The children in this experiment were clearly impacted by the increased difficulty of our ratio manipulation. When the ratio of the computed outcome and comparison array was too low, they had a difficult time distinguishing the transformed array from the comparison array, and performed at chance levels during the most-difficult division trials. This performance benefit for multiplication over division was also found in Experiment 1. However, one should use caution before concluding that scaling down the number of objects in a set is harder than scaling up. First, the division training was always given after the multiplication training in this design, which could readily lead to interference between the two processes. Second, a similar scaling paradigm used by McCrink & Spelke (2010, 2015) found no differences in overall performance as a function of operation type (multiplication / division) when the two processes were presented in isolation across children. The children in the current experiment also exhibited signs of using an extraneous, non-arithmetic range-based strategy when faced with these challenging division and multi-operation testing trials. This supports the idea that children may reach for alternate strategies to support their arithmetic reasoning in challenging situations (Siegler & Schrager, 1987; Siegler, 1987), even during non-symbolic math tasks. Despite this increased difficulty, they still exhibited no sign of utilizing a logical understanding of the inverse nature of the problem. Children who viewed an array that was quadrupled in number, and then immediately halved, were just as good at computing the outcome as those who saw a sequential doubling and then halving series.

General Discussion

The experiments reported here illustrate that children in the early years of formal schooling are strikingly good at performing multiple serial non-symbolic multiplication and division problems. They can readily quadruple, quarter, halve, and double the number of objects in a set, and they do so in a variety of scaling combinations. Their performance varied with the ratio of the computed outcome and a comparison array, indicating that this scaling is done in an approximate manner. In both experiments, multiplication transformations were more readily computed than division transformations, and multi-step problems were performed with precision similar to that of single-step multiplication problems. Despite being able to flexibly scale non-symbolic quantities by several multiplicative and divisive factors, children apparently failed to appreciate the conceptual relationship between these two operations, and were unable to use the logic of inversion to improve their performance.

The finding that performance on two-step problems - in which both multiplication and division transformations were presented - was similar to performance on single-operation problems was unexpected. Multi-step problems require two updates to the quantity representation while it is being transformed, increasing the number of arrays to be represented during problem solving. To our knowledge, there is only a small body of work on the development of combinatorial operations. The majority of these studies are with infants (Baillargeon et al., 1994; Moher, Teurk, & Feigenson, 2012; Piantadosi, Palmeri, & Aslin, under review) and find a lack of prowess in making more than one transformation to a represented object or set of objects. By early childhood this ability appears to be in place: Piantadosi and Aslin (under review) have found that by 4 years of age, children can combine two transformations over an object, and Gilmore and Spelke (2008) established that 5-year-olds were above chance at computing the outcomes to blended addition and subtraction problems. However, operational combinations may come at a cost to children; comparing the datasets of Gilmore and Spelke (2008) to those of Barth et al. (2006), obtained with similarly aged children and using similarly difficult ratios of outcomes to comparison arrays, suggests a decline in performance from single operations (69%) to multi-step operations (59%).

In contrast, children in the current task performed similarly on multi-transformation trials, compared to division trials in which only one transformation was demanded. It is possible that the serial nature of the paradigm used in these experiments assisted in lightening the memory load for the children. The children only ever needed to hold one single amount in mind at any one time, up until the comparison array appeared. If we had required them to hold each of these amounts in memory simultaneously, then the difference between the conditions would likely be greater (Feigenson & Yamaguchi, 2009; Moher & Feigenson, 2013). A challenge to this explanation is that the previous work that suggests costly updating during arithmetic (Barth et al., 2006; Gilmore & Spelke, 2008) was also presented serially in a paradigm similar to the current experiments. Perhaps the ability to carry out multiple updates here is a result of our slightly older participant pool (7 years of age, compared to 5). Children in the present study also had more practice with single transformations during training than was available in previous work; this may have emphasized that numerical information about the initial array could be discarded in favor of the transformed amount. Future work is needed to determine whether, and how, multiple scaling transformations impacts the performance of younger children with less mature updating capacities. For example, a valuable follow-up would be to administer a separate working memory task, such as the Automated Working Memory Assessment (or AWMA; Alloway (2007). In this way one could examine how the performance difference between single and multiple transformations varies with WM capacity, and use an individual differences approach to inform the way we teach alternate or supplemental mathematical strategies to children.

The children in this sample did not appear to use inversion reasoning to guide their computations. They experienced no benefit for problems in which they could have used a conceptual understanding of inversion to avoid carrying out mental computations, and instead simply to hold the initial array in mind and compare it to another at test. Bayes factor analyses conducted in each experiment showed evidence for this null effect. Pooling across all 128 children in Experiments 1 and 2, yields a Bayes Factor suggesting that our data are more than four times more likely under the null hypothesis ($B_{01} = 4.38$). Thus it appears from this study that there is a conceptual block between addition/subtraction reasoning and multiplication/division reasoning that transcends the format of the numerical representation (symbolic, or non-symbolic.) It is possible that children explicitly appreciated the logic of inversion, but still chose to use interim calculations, believing them to be more "trustworthy" (as with the adolescent population in Robinson & Dube, 2009). However, we believe this to be improbable. First, children are notoriously poor at providing explicit answers to problems they can intuitively solve even in infancy (such as where and when an object will move according to gravity; Keen, 2003). Second, both Kibbe and Feigenson (2015) and Gilmore and Spelke (2008) found that arithmetic logic in the non-symbolic, untrained domain precedes the same logic in the symbolic, formally trained domain, making the reverse pattern unlikely here.

Another possible explanation for the lack of inversion benefit is the extensive training that children received on each initial operation. To appreciate inversion in this paradigm, the children needed to essentially ignore the transformations during multi-step testing. During the several dozen training trials on the single-step operations (a necessary aspect of our design, to instantiate the transformation factor of the wands), they may have grown accustomed to discarding information about the initial array, making them less apt to access this information during testing trials. However, one could imagine the reverse would also be true; seeing dozens of trials of transformations that are exactly the physical (breaking, joining) and mathematical (*2, /2) inverse of each other should substantially increase their willingness to conceive of the operations as inverse to each other.

Thus, the findings reported here echo what is consistently found in the literature on symbolic inversion: a distinct lack of transfer between addition and subtraction inversion logic and multiplication and division logic (Robinson & Ninowski, 2003; Robinson et al., 2006; Robinson & Dube, 2009). These children were older than the children in Gilmore and Spelke (2008), who exhibited inversion reasoning in the realm of approximate addition and subtraction, yet they failed to exhibit it here in a logically identical paradigm. The design here was between-subjects, and therefore the children who were placed into the inversion conditions could know immediately from the training sessions that the scaling factors for multiplication and division were the inverse of each other. With this knowledge in hand, there is logically no need to perform any scaling process or updating of the initial array. Further, the children studied in these experiments were no strangers to the use of shortcuts to solve problems, especially challenging ones; their usage of non-scaling, range-based strategies increased when the required discrimination required were more difficult.

An unresolved question, then, is *why* children cannot recognize and/or employ inversion logic for non-symbolic scaling. A potential explanation is that the conceptual logic is there, as it is with adding and subtracting – but it is masked by a lack of confidence in the domain of scaling. Children are, generally, presented with the viewpoint that multiplication and division are more challenging than addition and subtraction. For example, multiplication facts and division facts are more likely to be taught in a memorized, rote way than addition and subtraction facts (e.g., the prevalence of "times tables"), and addition / subtraction concepts are introduced several years before multiplication / division concepts in the course of early schooling (kindergarten vs. third grade). We believe this may have the unintended effect of leading children and even adults (Robinson & Ninowski, 2003) to be less likely to reach for arithmetic intuition when multiplying or dividing than when adding and subtracting, even when it is applicable. They would rather be "safe than sorry." This explanation leads to the counterintuitive hypothesis that preschool children, or any other unschooled population, might be *more* likely than older children to reach for inversion-based strategies when multiplying and dividing. After all, they have not been exposed to these contrasting ideas of arithmetic difficulty and formal solving strategies for scaling and adding/subtracting, and do appear to appreciate the principles of inversion for other types of non-symbolic arithmetic..

In sum, these results converge with those of Barth et al. (2009) and McCrink & Spelke (2010), who found that young children were able to non-symbolically multiply and divide and to do so with a variety of scaling factors. They support the robustness of this phenomenon: children with little to no formal experience with multiplication and division were able to serially compute multiple scaling transformations. These results also suggest that the divide between different types of inversion understanding found in the symbolic literature is a deep one, which extends to non-symbolic reasoning as well. This information is essential knowledge, given the contradictory findings that non-symbolic mathematical skills are sometimes linked to formal mathematical skills (Gilmore, Attridge, De Smedt, & Inglis, 2014; Halberda, Mazzocco, & Feigenson, 2008; Mazzocco, Feigenson, & Halberda, 2011; Park & Brannon, 2013; Piazza, Facoetti, Trussardi, Berteletti, Conte, Lucangeli, Dehaene, & Zorzi, 2010) and sometimes not (Bonny & Lourenco, 2013; De Smedt & Gilmore, 2011; Holloway & Ansari, 2009; Libertus, Feigenson, & Halberda, 2013; Sasanguie, De Smedt, Defever, & Reynvoet, 2012; Szucs, Devine, Soltesz, Nobes, & Gabriel, 2014) as well as the proposition that these skills form a foundation for some early symbolic mathematical skills (as in Gilmore, McCarthy, & Spelke, 2007). By delineating the boundaries of children's arithmetic intuitions, and taking into account their strengths and weaknesses, we may be able to design effective ways to improve their conceptual understanding at critical points in development.

Acknowledgments

We are grateful to the parents and children who participated in the study, as well as PS 204 for their support of our research. We thank Michal Boyars, Wesley Birdsall, Jenn Garcia, and Jessica Taggart for their assistance in collecting data.

This work was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development under grant R15 HD065629-01 to the first author and by NSF DRL-0950252 to the second author.

References

Agrillo C, Dadda M, Serena G, Bisazza A. Use of number by fish. PloS one. 2009; 4:e4786. [PubMed: 19274079]

Alloway, T. Automated Working Memory Assessment. London: Pearson Assessment; 2007.

Baillargeon R. Physical reasoning in young infants: Seeking explanations for impossible events. British Journal of Developmental Psychology. 1994; 12:9–33.

Barth H, Baron A, Spelke E, Carey S. Children's multiplicative transformations of discrete and continuous quantities. Journal of Experimental Child Psychology. 2009; 103:441–454. [PubMed: 19289237]

Barth H, Beckmann L, Spelke ES. Nonsymbolic, approximate arithmetic in children: Abstract addition prior to instruction. Developmental Psychology. 2008; 44(1466)

Barth H, Kanwisher N, Spelke E. The construction of large number representations in adults. Cognition. 2003; 86:201–221. [PubMed: 12485738]

Barth H, La Mont K, Lipton J, Spelke ES. Abstract number and arithmetic in preschool children. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102:14116–14121. [PubMed: 16172388]

Barth H, La Mont K, Lipton J, Dehaene S, Kanwisher N, Spelke E. Non-symbolic arithmetic in adults and young children. Cognition. 2006; 98:199–222. [PubMed: 15876429]

Bisanz, J.; LeFevre, JA. Strategic and nonstrategic processing in the development of mathematical cognition. In: Bjorklund, DF., editor. Children's strategies: Contemporary views of cognitive development. Hillsdale, NJ: Erlbaum; 1990. p. 213-244.

Bonny JW, Lourenco SF. The approximate number system and its relation to early math achievement: Evidence from the preschool years. Journal of Experimental Child Psychology. 2013; 114:375–388. [PubMed: 23201156]

Brannon EM, Wusthoff CJ, Gallistel CR, Gibbon J. Numerical subtraction in the pigeon: Evidence for a linear subjective number scale. Psychological Science. 2001; 12:238–243. [PubMed: 11437307]

Bryant P, Christie C, Rendu A. Children's understanding of the relation between addition and subtraction: Inversion, identity, and decomposition. Journal of Experimental Child Psychology. 1999; 74:194–212. [PubMed: 10527554]

Cantlon JF, Brannon EM. Basic math in monkeys and college students. PLoS biology. 2007; 5:e328. [PubMed: 18092890]

Cantlon JF, Libertus ME, Pinel P, Dehaene S, Brannon EM, Pelphrey KA. The neural development of an abstract concept of number. Journal of Cognitive Neuroscience. 2009; 21:2217–2229. [PubMed: 19016605]

Cheries EW, Wynn K, Scholl BJ. Interrupting infants' persisting object representations: an object-based limit? Developmental Science. 2006; 9:F50–F58. [PubMed: 16911437]

Cordes S, Brannon EM. Crossing the divide: Infants discriminate small from large numerosities. Developmental Psychology. 2009; 45:1583. [PubMed: 19899916]

Cordes S, Gelman R, Gallistel CR. Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. Psychological Bulletin and Review. 2002; 8:698–707.

De Smedt B, Gilmore C. Defective number module or impaired access? Numerical magnitude processing in first graders with mathematical difficulties. Journal of Experimental Child Psychology. 2011; 108:278–292. [PubMed: 20974477]

Dehaene, S. The number sense. New York, NY: Oxford University Press, Inc.; 1997.

Feigenson L, Yamaguchi M. Limits on infants' ability to dynamically update object representations. Infancy. 2009; 14:244–262.

Gebuis T, Reynvoet B. The role of visual information in numerosity estimation. PLoS ONE. 2012; 7(5)

Gebuis T, Reynvoet B. Continuous visual properties explain neural responses to non-symbolic number. Psychophysiology. 2012; 4(11):1481–1491.

Gilmore CK. Investigating children's understanding of inversion using the missing number paradigm. Cognitive Development. 2006; 21:301–316.

Gilmore C, Attridge N, De Smedt B, Inglis M. Measuring the Approximate Number System in children: Exploring the relationships among different tasks. Learning and Individual Differences. 2014; 29:50–58.

Gilmore CK, Spelke ES. Children's understanding of the relationship between addition and subtraction. Cognition. 2008; 107:932–945. [PubMed: 18281029]

Gilmore CK, McCarthy SE, Spelke ES. Symbolic arithmetic knowledge without instruction. Nature. 2007; 447:589–591. [PubMed: 17538620]

Halberda J, Ly R, Wilmer JB, Naiman DQ, Germine L. Number sense across the lifespan as revealed by a massive Internet-based sample. Proceedings of the National Academy of Sciences. 2012; 109:11116–11120.

Halberda J, Mazzocco MM, Feigenson L. Individual differences in non-verbal number acuity correlate with math achievement. Nature. 2008; 455:665–668. [PubMed: 18776888]

Heck K, Braveman P, Cubbin C, Chavez G, Kiely J. Socioeconomic status and breastfeeding initiation among California mothers. Public Health Reports. 2006; 121(1):51–59.

Holloway ID, Ansari D. Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's math achievement. Journal of Experimental Child Psychology. 2009; 103:17–29. [PubMed: 18513738]

Keen R. Representation of Objects and Events Why Do Infants Look So Smart and Toddlers Look So Dumb? Current Directions in Psychological Science. 2003; 12:79–83.

Kibbe MM, Feigenson L. Young children 'solve for x'using the Approximate Number System. Developmental Science. 2015; 18:38–49. [PubMed: 24589420]

Kibbe MM, Leslie AM. What's the object of object working memory in infancy? Unraveling 'what' and 'how many'. Cognitive Psychology. 2013; 66:380–404. [PubMed: 23770623]

Klein JS, Bisanz J. Preschoolers doing arithmetic: The concepts are willing but the working memory is weak. Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale. 2000; 54:105–116. [PubMed: 10881394]

Knops A, Zitzmann S, McCrink K. Examining the presence and determinants of operational momentum in childhood. Frontiers in Psychology. 2013; 4:325. [PubMed: 23772216]

Libertus ME, Feigenson L, Halberda J. Numerical approximation abilities correlate with and predict informal but not formal mathematics abilities. Journal of Experimental Child Psychology. 2013; 116:829–838. [PubMed: 24076381]

Libertus ME, Starr A, Brannon EM. Number trumps area for 7-month-old infants. Developmental Psychology. 2014; 50:108–112. [PubMed: 23647413]

Lubienski C, Lubienski S. Charter, Private, Public Schools and Academic Achievement: New Evidence From NAEP Mathematics Data. National Center for Education Statistics, Institute of Education Sciences. 2006

Mazzocco, Feigenson, & Halberda. Impaired acuity of the approximate number system underlies mathematical learning disability (Dyscalculia). Child Development. 2011; 82(4):1224–1237. [PubMed: 21679173]

McCrink K, Spelke ES. Core multiplication in childhood. Cognition. 2010; 116:204–216. [PubMed: 20537618]

McCrink K, Spelke ES. Non-symbolic division in childhood. Journal of Experimental Child Psychology. 2015; 142:66–82. [PubMed: 26513326]

McCrink K, Wynn K. Large-number addition and subtraction by 9-month-old infants. Psychological Science. 2004; 15:776–781. [PubMed: 15482450]

McCrink K, Wynn K. Ratio abstraction by 6-month-old infants. Psychological Science. 2007; 18:740–745. [PubMed: 17680947]

McCrink K, Wynn K. Operational momentum in large-number addition and subtraction by 9-month-olds. Journal of Experimental Child Psychology. 2009; 103:400–408. [PubMed: 19285683]

Meck WH, Church RM. A mode control model of counting and timing processes. Journal of Experimental Psychology: Animal Behavior Processes. 1983; 9:320–334. [PubMed: 6886634]

Moher M, Feigenson L. Factors influencing infants' ability to update object representations in memory. Cognitive Development. 2013; 28:272–289. [PubMed: 24049245]

Moher M, Tuerk AS, Feigenson L. Seven-month-old infants chunk items in memory. Journal of Experimental Child Psychology. 2012; 112:361–377. [PubMed: 22575845]

Nelson XJ, Jackson RR. The role of numerical competence in a specialized predatory strategy of an araneophagic spider. Animal Cognition. 2012; 15:699–710. [PubMed: 22526693]

Nunes T, Bryant P, Burman D, Bell D, Evans D, Hallett D. Deaf children's informal knowledge of multiplicative reasoning. Journal of Deaf Studies and Deaf Education. 2009; 14:260–277. [PubMed: 18948568]

Nunes T, Bryant P, Evans D, Bell D, Gardner S, Gardner A, Carraher J. The contribution of logical reasoning to the learning of mathematics in primary school. British Journal of Developmental Psychology. 2007; 25:147–166.

Park JH, Nunes T. The development of the concept of multiplication. Cognitive Development. 2001; 16:763–773.

Park J, Brannon EM. Training the approximate number system improves math proficiency. Psychological Science. 2013 0956797613482944.

Park J, DeWind NK, Woldorff MG, Brannon EM. Rapid and Direct Encoding of Numerosity in the Visual Stream. Cerebral Cortex. 2015:bhv017.

Piazza M, Facoetti A, Trussardi A, Berteletti I, Conte S, Lucangeli D, Dehaene S, Zorzi M. Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. Cognition. 2010; 116(1):33–41. [PubMed: 20381023]

Prather RW, Alibali MW. The development of arithmetic principle knowledge: How do we know what learners know? Developmental Review. 2009; 29:221–248.

Rasmussen C, Ho E, Bisanz J. Use of the mathematical principle of inversion in young children. Journal of Experimental Child Psychology. 2003; 85:89–102. [PubMed: 12799163]

Robinson KM, Dubé AK. Children's understanding of addition and subtraction concepts. Journal of Experimental Child Psychology. 2009; 103:532–545. [PubMed: 19167015]

Robinson KM, Ninowski JE. Adults' understanding of inversion concepts: How does performance on addition and subtraction inversion problems compare to performance on multiplication and division inversion problems? Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale. 2003; 57:321. [PubMed: 14710869]

Robinson KM, Ninowski JE, Gray ML. Children's understanding of the arithmetic concepts of inversion and associativity. Journal of Experimental Child Psychology. 2006; 94:349–362. [PubMed: 16674969]

Rosenberg RD, Feigenson L. Infants hierarchically organize memory representations. Developmental Science. 2013; 16:610–621. [PubMed: 23786478]

Rouder JN, Speckman PL, Sun D, Morey RD, Iverson G. Bayesian t tests for accepting and rejecting the null hypothesis. Psychonomic bulletin & review. 2009; 16(2):225–237. [PubMed: 19293088]

Sasanguie D, De Smedt B, Defever E, Reynvoet B. Association between basic numerical abilities and mathematics achievement. British Journal of Developmental Psychology. 2012; 30:344–357. [PubMed: 22550952]

Piantadosi ST, Palmeri H, Aslin R. Limits on composition of conceptual operations in 9-month-olds. 2015 Manuscript under review.

Piantadosi ST, Aslin R. Compositional reasoning in early childhood. 2015 Manuscript under review.

Sherman J, Bisanz J. Evidence for use of mathematical inversion by three-year-old children. Journal of Cognition and Development. 2007; 8:333–344.

Shrager J, Siegler RS. SCADS: A model of children's strategy choices and strategy discoveries. Psychological Science. 1998; 9:405–410.

Siegler RS. The perils of averaging data over strategies: An example from children's addition. Journal of Experimental Psychology: General. 1987; 116:250.

Siegler RS, Stern E. Conscious and unconscious strategy discoveries: A microgenetic analysis. Journal of Experimental Psychology: General. 1998; 127:377. [PubMed: 9857493]

Szucs D, Devine A, Soltesz F, Nobes A, Gabriel F. Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. Cortex. 2013; 49(10):2674–2688. [PubMed: 23890692]

Szucs D, Devine A, Soltesz F, Nobes A, Gabriel F. Cognitive components of a mathematical processing network in 9-year-old children. Developmental Science. 2014; 17:506–524. [PubMed: 25089322]

Szucs D, Nobes A, Devine A, Gabriel FC, Gebuis T. Visual stimulus parameters seriously compromise the measurement of approximate number system acuity and comparative effects between adults and children. Frontiers in Developmental Psychology. 2013

Uller C, Carey S, Huntley-Fenner G, Klatt L. What representations might underlie infant numerical knowledge? Cognitive Development. 1999; 14:1–36.

Vilette B. Do young children grasp the inverse relationship between addition and subtraction?: Evidence against early arithmetic. Cognitive Development. 2002; 17:1365–1383.

Xenidou-Dervou I, Lieshout EC, Schoot M. Working Memory in Nonsymbolic Approximate Arithmetic Processing: A Dual-Task Study With Preschoolers. Cognitive Science. 2014; 38:101–127. [PubMed: 23855416]

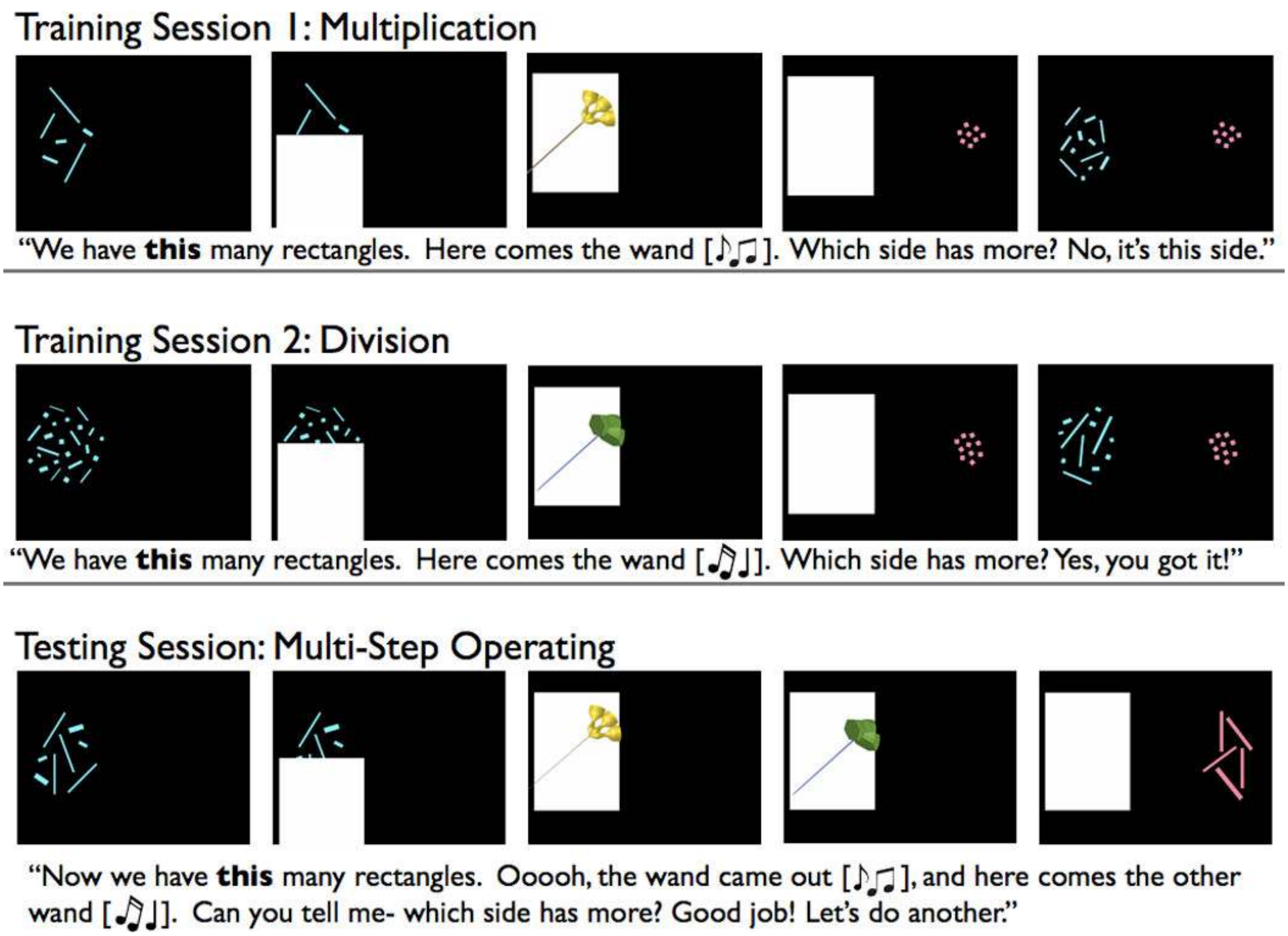

Figure 1.

Schematic of training and testing procedure for the Inversion condition Times 2, Divided by 2. The children were given a series of training trials for both multiplication and division problems, in which the answer was revealed after they judged whether the transformed (blue) or comparison (pink) array was more numerous. They then viewed a multi-step problem in which both operations were performed over the occluded initial array.

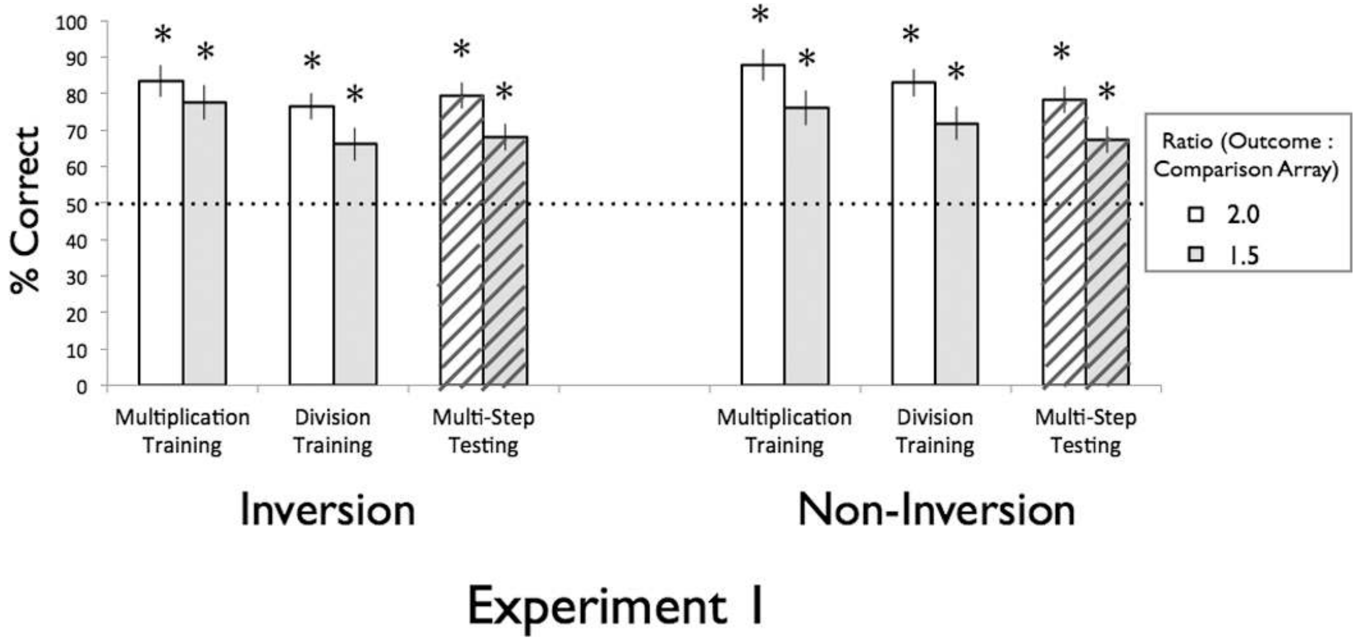

Figure 2.

Children's performance at training (multiplication and division) and testing (both operations) for the Inversion and Non-Inversion conditions in Experiment 1. The overall percentage correct is plotted for each condition (Inversion, Non-Inversion), session (multiplication training, division training, multi-step testing), and ratio of larger array to smaller array at comparison (Ratio: 2.0, 1.5). Error bars indicate +/− 1 SEM, and asterisks indicate $p < .05$ when comparing performance against chance level (the dotted line).

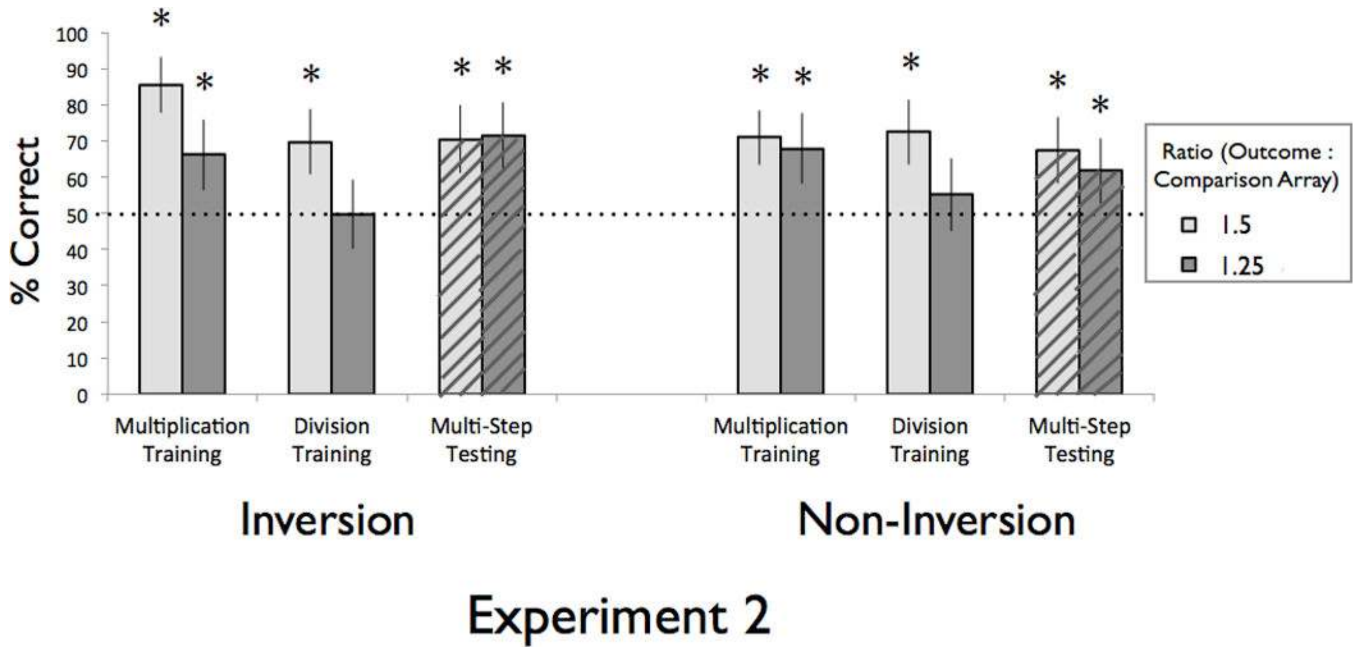

Figure 3.

Children's performance at training (multiplication and division) and testing (both operations) for the Inversion and Non-Inversion conditions in Experiment 2. The overall percentage correct is plotted for each condition (Inversion, Non-Inversion), session (multiplication training, division training, multi-step testing), and ratio of larger array to smaller array at comparison (Ratio: 1.5, 1.25). Error bars indicate +/− 1 SEM, and asterisks indicate $p < .05$ when comparing performance against chance level (the dotted line).

Table 1

Exact values used for Experiments 1 and 2

Experiment 1			
Inversion Conditions:		**Non-Inversion Conditions:**	
Times 2, Divided by 2	**Times 4, Divided by 4**	**Times 2, Divided by 4**	**Times 4, Divided by 2**
Multiplication Training	**Multiplication Training**	**Multiplication Training**	**Multiplication Training**
4 * 2 vs. 5, 12	4 * 4 vs. 8, 32	4 * 2 vs. 5, 12	4 * 4 vs. 8, 32
6 * 2 vs. 6, 24	6 * 4 vs. 16, 36	6 * 2 vs. 6, 24	6 * 4 vs. 16, 36
8 * 2 vs. 8, 32	8 * 4 vs. 21, 48	8 * 2 vs. 8, 32	8 * 4 vs. 21, 48
10 * 2 vs. 13, 30	10 * 4 vs. 20, 80	10 * 2 vs. 13, 30	10 * 4 vs. 20, 80
12 * 2 vs. 16, 36	12 * 4 vs. 24, 96	12 * 2 vs. 16, 36	12 * 4 vs. 24, 96
14 * 2 vs. 14, 56	14 * 4 vs. 37, 84	14 * 2 vs. 14, 56	14 * 4 vs. 37, 84
Div. Training	**Div. Training**	**Div. Training**	**Div. Training**
16 / 2 vs. 4, 16	16 / 4 vs. 3, 6	16 / 4 vs. 3, 6	16 / 2 vs. 4, 16
24 / 2 vs. 8, 18	24 / 4 vs. 3, 12	24 / 4 vs. 3, 12	24 / 2 vs. 8, 18
32 / 2 vs. 11, 24	32 / 4 vs. 4, 16	32 / 4 vs. 4, 16	32 / 2 vs. 11, 24
44 / 2 vs. 11, 44	44 / 4 vs. 7, 17	44 / 4 vs. 7, 17	44 / 2 vs. 11, 44
48 / 2 vs. 12, 48	48 / 4 vs. 8, 18	48 / 4 vs. 8, 18	48 / 2 vs. 12, 48
56 / 2 vs. 19, 42	56 / 4 vs. 7, 28	56 / 4 vs. 7, 28	56 / 2 vs. 19, 42
*** 2 / 2 Testing**	*** 4 / 4 Testing**	*** 2 / 4 Testing**	*** 4 / 2 Testing**
8 * 2 / 2 vs. 4, 5, 12, 16	8 * 4 / 4 vs. 4, 5, 12, 16	8 * 2 / 4 vs. 2, 3, 6, 8	8 * 4 / 2 vs. 8, 11, 24, 32
14 * 2 / 2 vs. 7, 9, 21, 28	14 * 4 / 4 vs. 7, 9, 21, 28	14 * 2 / 4 vs. 3, 5, 11, 14	14 * 4 / 2 vs. 14, 19, 42, 56
18 * 2 / 2 vs. 9, 12, 27, 36	18 * 4 / 4 vs. 9, 12, 27, 36	18 * 2 / 4 vs. 4, 6, 14, 18	18 * 4 / 2 vs. 18, 24, 54, 72
22 * 2 / 2 vs. 11, 15, 33, 44	22 * 4 / 4 vs. 11, 15, 33, 44	22 * 2 / 4 vs. 5, 7, 17, 22	22 * 4 / 2 vs. 22, 29, 66, 88
28 * 2 / 2 vs. 14, 19, 42, 56	28 * 4 / 4 vs. 14, 19, 42, 56	28 * 2 / 4 vs. 7, 9, 21, 28	28 * 4 / 2 vs. 28, 37, 84, 112
40 * 2 / 2 vs. 20, 27, 60, 80	40 * 4 / 4 vs. 20, 27, 60, 80	40 * 2 / 4 vs. 10, 13, 30, 40	40 * 4 / 2 vs. 40, 53, 120, 160
Experiment 2			
Inversion Conditions:		**Non-Inversion Conditions:**	
Times 2, Divided by 2	**Times 4, Divided by 4**	**Times 2, Divided by 4**	**Times 4, Divided by 2**
Multiplication Training	**Multiplication Training**	**Multiplication Training**	**Multiplication Training**
4 * 2 vs. 6, 10	4 * 4 vs. 11, 24	4 * 2 vs. 6, 10	4 * 4 vs. 11, 24
6 * 2 vs. 8, 18	6 * 4 vs. 19, 30	6 * 2 vs. 8, 18	6 * 4 vs. 19, 30
8 * 2 vs. 11, 24	8 * 4 vs. 26, 40	8 * 2 vs. 11, 24	8 * 4 vs. 26, 40
10 * 2 vs. 16, 25	10 * 4 vs. 27, 60	10 * 2 vs. 16, 25	10 * 4 vs. 27, 60
12 * 2 vs. 19, 30	12 * 4 vs. 32, 72	12 * 2 vs. 19, 30	12 * 4 vs. 32, 72

14 * 2 vs. 19, 42	14 * 4 vs. 45, 70	14 * 2 vs. 19, 42	14 * 4 vs. 45, 70
Div. Training	**Div. Training**	**Div. Training**	**Div. Training**
16 / 2 vs. 5, 12	16 / 4 vs. 3, 5	16 / 4 vs. 3, 5	16 / 2 vs. 5, 12
24 / 2 vs. 10, 15	24 / 4 vs. 4, 9	24 / 4 vs. 4, 9	24 / 2 vs. 10, 15
32 / 2 vs. 13, 20	32 / 4 vs. 5, 12	32 / 4 vs. 5, 12	32 / 2 vs. 13, 20
44 / 2 vs. 15, 33	44 / 4 vs. 9, 14	44 / 4 vs. 9, 14	44 / 2 vs. 15, 33
48 / 2 vs. 16, 36	48 / 4 vs. 10, 15	48 / 4 vs. 10, 15	48 / 2 vs. 16, 36
56 / 2 vs. 22, 35	56 / 4 vs. 9, 21	56 / 4 vs. 9, 21	56 / 2 vs. 22, 35
*** 2 / 2 Testing**	*** 4 / 4 Testing**	*** 2 / 4 Testing**	*** 4 / 2 Testing**
8 * 2 / 2 vs. 5, 6, 10, 12	8 * 2 / 2 vs. 5, 6, 10, 12	8 * 2 / 4 vs. 2, 3, 5, 6	8 * 4 / 2 vs. 11, 13, 20, 24
14 * 2 / 2 vs. 9, 11, 18, 21	14 * 4 / 4 vs. 9, 11, 18, 21	14 * 2 / 4 vs. 5, 6, 9, 11	14 * 4 / 2 vs. 19, 22, 35, 42
18 * 2 / 2 vs. 12, 14, 23, 27	18 * 4 / 4 vs. 12, 14, 23, 27	18 * 2 / 4 vs. 6, 7, 11, 14	18 * 4 / 2 vs. 24, 29, 45, 54
22 * 2 / 2 vs. 15, 18, 28, 33	22 * 4 / 4 vs. 15, 18, 28, 33	22 * 2 / 4 vs. 7, 9, 14, 17	22 * 4 / 2 vs. 29, 35, 55, 66
28 * 2 / 2 vs. 19, 22, 35, 42	28 * 4 / 4 vs. 19, 22, 35, 42	28 * 2 / 4 vs. 9, 1,, 18, 21	28 * 4 / 2 vs. 37, 45, 70, 84
40 * 2 / 2 vs. 27, 32, 50, 60	40 * 2 / 2 vs. 27, 32, 50, 60	40 * 2 / 4 vs. 13, 16, 25, 30	40 * 4 / 2 vs. 53, 64, 100, 120

Note. Detailed here are the values used for the multiplication and division training trials, as well as the values for problems used during the multi-step testing movies, separated by experiment (Experiment 1: Ratios of 2.0 and 1.5; Experiment 2: Ratios of 1.5 and 1.25) and condition (Inversion, Non-Inversion).

Table 2

Children's performance in Experiments 1 and 2

Scaling Factors	Training/Testing Session:	Experiment 1			Experiment 2		
		Weber Ratio	% Correct	Std. Error	Weber Ratio	% Correct	Std. Error
Multiplied by 2, Divided by 2 (Inversion Condition)	Multiplication Training:	2.0 1.5	81.2 76	6.3 7.6	1.5 1.25	92.7 66.7	6.3 7.6
	Division Training:	2.0 1.5	83.3 75	6.9 8	1.5 1.25	82.3 41.7	6.9 8
	Multi-Step Testing:	2.0 1.5	79.2 70.3	7 6.9	1.5 1.25	61.5 75.5	7 6.9
Multiplied by 4, Divided by 4 (Inversion Condition)	Multiplication Training:	2.0 1.5	85.4 79.2	6.3 7.6	1.5 1.25	78.1 65.6	6.3 7.6
	Division Training:	2.0 1.5	69.8 57.3	6.9 8	1.5 1.25	55.2 49	6.9 8
	Multi-Step Testing:	2.0 1.5	79.7 65.6	7 6.9	1.5 1.25	77.6 64.6	7 6.9
Multiplied by 2, Divided by 4 (Non-Inversion Condition)	Multiplication Training:	2.0 1.5	88 72.9	6.3 7.6	1.5 1.25	75 63.5	6.3 7.6
	Division Training:	2.0 1.5	84.4 70.8	6.9 8	1.5 1.25	71.9 62.5	6.9 8
	Multi-Step Testing:	2.0 1.5	81.8 68.2	7 6.9	1.5 1.25	69.8 65.6	7 6.9
Multiplied by 4, Divided by 2 (Non-Inversion Condition)	Multiplication Training:	2.0 1.5	87.5 79.2	6.3 7.6	1.5 1.25	63.5 70.8	6.3 7.6
	Division Training:	2.0 1.5	81.3 72.9	6.9 8	1.5 1.25	74 50	6.9 8
	Multi-Step Testing:	2.0 1.5	74.5 66.1	7 6.9	1.5 1.25	63 54.9	7 6.9

Note. Presented here are the percentages correct exhibited by the children for the multiplication / division training trials and multi-step testing movies, as a function of Experiment, Ratio at comparison, session, and scaling factor.