Distribution Agreement

Signature:

______________________________ _______________

Sonia Grayce Rao April 14, 2010

Magnitude and Mathematics: Number, Space, and Mathematical Achievement

by

Sonia G. Rao

Adviser

Dr. Stella Lourenco

Department of
Psychology

Dr. Stella Lourenco
Adviser

Dr. Lawrence Barsalou
Committee Member

Dr. Elena Glazov-Corrigan
Committee Member

April 14, 2010

Magnitude and Mathematics: Number, Space, and Mathematical Achievement

By

Sonia G. Rao

Adviser

Dr. Stella Lourenco

An abstract of
A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Department of
Psychology

2010

Abstract

Magnitude and Mathematics: Number, Space, and Mathematical Achievement

By Sonia G. Rao

There has been substantial research on nonverbal number estimation in adults, including the possibility that individual differences in numerical estimation abilities may correlate with mathematical performance. Additionally, there is some evidence that spatial and numerical processing systems overlap in what is called a general magnitude system (GMS). If a GMS exists, then it is conceivable that nonverbal spatial estimation abilities may correlate with math achievement as well. This study is the first to examine this relation between magnitude estimation, both spatial and numerical, and math achievement in adults at one point in time. Preliminary support was found for a relation between both spatial and numerical acuity and math performance, opening the possibility that number and space have a common processing mechanism relating to math performance.

Magnitude and Mathematics: Number, Space, and Mathematical Achievement

By

Sonia G. Rao

Adviser

Dr. Stella Lourenco

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Department of
Psychology

2010

Acknowledgements

This thesis would not have been possible without the support and enthusiasm from my adviser, Dr. Stella Lourenco, who has been influential not only during this process but also throughout my entire psychology education at Emory. A special thank you to members of the Spatial Cognition Lab for always giving me valuable feedback, helping me with technical difficulties, and agreeing to be guinea pigs while I learned to administer a standardized test. Additionally, I would like to thank my committee members, Dr. Lawrence Barsalou and Dr. Elena Glazov-Corrigan for their support. Finally, I would like to thank my family and friends for reminding me that I would, in fact, finish this paper, and encouraging me every step of the way.

Table of Contents

Introduction……………………………………………………………………………1

Methods…………………………………………………………………………….13

Results………………………………………………………………………………..19

Discussion …………………………………………………………………………26

References …………………………………………………………………………36

Tables ………………………………………………………………………………..40

Figures………………………………………………………………………………47

Magnitude and Mathematics: Number, Space, and Mathematical Achievement

"How is it possible simultaneously to view mathematics as a system of transparent truths and as a tangle of relations between opaque entities?"

-Feigenson, Dehaene, & Spelke (2004)

One of the most puzzling aspects of mathematics is the apparent coexistence of both accessible and enigmatic concepts. Additive principles, for instance, are understood even by infants and nonhuman animals (Wynn, 1992; Rumbaugh, Savage-Rumbaugh, & Hegel, 1987), and it goes without saying that most adults do not find it difficult to understand why the summation of "4" and 8" cannot equal "1." On the other hand, certain domains of mathematics are only understood after excessive schooling and practice; multiplication tables, algebraic formulas, and negative numbers, for instance, are not intuitive concepts. It would not be surprising for us to meet someone with a strong math education who needs to use a tip calculator at a restaurant, nor would it be incomprehensible that someone may forget whether the multiplication of 7 x 8 is equal to 56 or 48. Certain mathematical abilities, then, are taken for granted, and others do not become automatic even after years of training. How does this system of mathematics arise and what facilitates our comprehension of it? Why is it that some people are better at understanding mathematics than others? What factors are related to higher math performance? These questions, while they will not be answered in full by this study, are the impetus for this research, which examines the relation between lower-level magnitude processing (i.e., estimation) and higher-level mathematical achievement (e.g., calculation).

Approximate Number System

The notion of an approximate number system (ANS) may hold some explanatory power in discerning the origins of numerical understanding and the mechanisms that underlie math achievement. The ANS refers to the ability to represent and approximate number nonverbally: that is, without counting (Feigenson, Dehaene, & Spelke, 2004; Dehaene, 1997). This approximate number system is, at least to some extent, abstract, and thus remains consistent regardless of how numerical information is presented, that is, via visual or auditory modalities (Barth, Kanwisher & Spelke, 2003; Dehaene, Dehaene-Lambertz & Cohen, 1998). One of the major hallmarks of the ANS is its imprecision; representations and approximations of number, in accordance with Weber's law, grow increasingly fuzzy as numerical magnitude grows progressively larger (Dehaene, 1997; Dehaene et al. 1998; Feigenson et al. 2004; Pica, Lemer, Izard, & Dehaene, 2004). Accordingly, the representation of the number "5" will be more precise than the number "50," as the ANS' imprecision increases with progressively larger numerical magnitudes. This imprecision is particularly salient when making comparisons, or discriminations, of these numerical representations, as accuracy decreases as the distance between number sets diminishes—this phenomenon is known as the distance effect (Dehaene et al., 1998). For example, in studies where adults are presented with arrays of dots and asked to discriminate which set is greater in number (Halberda, Mazzocco & Feigenson, 2008; Hurewitz, Gelman & Schnitzer, 2005), participants are slower and more inaccurate in making discriminations as the numerical difference between dot arrays decreases (i.e. between 10 and 20 as opposed to 10 and 12). Conversely, they are faster and more accurate in making these discriminations as the numerical difference between dot arrays increases. These distance effects persist when comparing symbolic representations of number (i.e. Arabic numerals). Moyer and Landauer (1967), for example, showed that when adults are asked to compare two Arabic numerals, they are slower and more error-prone in their discrimination when the numbers are closer in magnitude (i.e. 9 and 10) than when they are farther apart (i.e. 9 and 4).

Although distance effects in number discrimination are fairly robust, observed even in infants (Xu & Spelke, 2000; Xu, Spelke, & Goddard, 2005) and animals (Rumbaugh, Savage-Rumbaugh, & Hegel, 1987), there is some evidence for individual differences in ANS acuity (Halberda et al., 2008). In other words, some evidence suggests that certain individuals have a more “fine-tuned” ANS than others in that they are able to reliably make numerical discriminations at very low ratio differences. If presented with various arrays of dots in a way that prevents counting, for instance, they may be able to judge which set is greater in number when sets differ by a ratio of 7:8; others, however, may not be able to make this discrimination reliably and instead may be limited to a easier ratio difference (i.e., 3:4) at which they can make numerical discriminations above chance. Individual differences in ANS acuity may be expressed as a Weber fraction, which indicates the amount of error in numerical discrimination based on the just noticeable difference between two sets of numbers (Cordes, Gelman, Gallistel, & Whalen, 2001; Halberda et al., 2008; Pica et al., 2004). In other words, the lowest difference between numbers, or lowest ratio, that an individual can reliably discriminate allows for an assessment of the precision of one’s ANS. For example, if an individual can reliably discriminate number sets down to a ratio difference of 9:10 and another individual can only reliably discriminate number sets down to a ratio difference of 1:2, then we would conclude the former individual has a higher ANS acuity than the latter.

In addition to its imprecision, the other major hallmark of the ANS is its independence from a learning process (Dehaene, 1997; Dehaene et al., 1998; Feigenson et al., 2004; Pica et al., 2004). Infants as young as six months of age are capable of discriminating numbers that vary by a 1:2 ratio difference and, like adults, have a threshold ratio at which performance on number discrimination tasks diminishes. Six-month-olds, for example, can only make discriminations between numbers that vary by a 1:2 ratio (Xu & Spelke, 2000; Xu, Spelke, & Goddard, 2005), whereas 10-month-olds can discriminate ratios up to 2:3, suggesting that discrimination capabilities become more refined throughout development. As infants have no explicit training in number understanding, mathematics, or counting, their judgments must be based on an internal conception of magnitude that is subject to the same Weber-like restrictions of an adult ANS. It is probable, then, that infants already possess an ANS, albeit a less refined version, that assists them in discriminating numbers.

Similarly, studies with nonhuman animals show that they are capable of making number discriminations in the absence of explicit mathematical training, suggesting that they, too, make numerical judgments based on an internal understanding of magnitude similar to the ANS in humans. Rats, for instance, are capable of demonstrating number discrimination by lever pressing, and, like infants and adults, their precision on these tasks decreases as the ratio difference between number sets decreases (Gallistel & Gelman, 1992; Meck & Church, 1983). Studies of nonhuman primates show that chimpanzees are capable of understanding number and proportion (Woodruff & Premack, 1981) as well as using additive principles to compare magnitudes of food (Rumbaugh, et al., 1987). These demonstrations of numerical understanding in chimpanzees are limited to lower numbers, and the use of additive principles to compare food piles is more difficult when the sums are close together (i.e., 1+1 and 2+1); these restrictions are highly similar to those associated with the human ANS.

Although it is possible to train animals to perform tasks we associate with higher-level mathematical reasoning, such as representing and manipulating numerical values symbolically, such training is extensive and produces limited results; this suggests that animals do not automatically rely on advanced mathematical reasoning for numerical discrimination tasks (Dehaene et al., 1998;

Dehaene 1997). It would seem likely, then, that animals depend upon an ANS-like mechanism, subject to similar Weber-like constraints as the human ANS, to make numerical decisions. The presence of an ANS-like mechanism in animals, particularly its presence in non-human primates, suggests that the human ANS is an ancient, evolutionary-based system that may be biologically determined and perhaps even innate; this further underscores the separation of ANS from learning processes.

If the ANS is a biologically determined, evolutionary by-product that allows us to estimate number through an innate understanding of magnitude, then it is possible that a relationship exists between the ANS and higher-level math processes. More specifically, it is possible that our initial understanding of number resides in an innate ANS, which becomes the foundational basis upon which we build more complex concepts of number and mathematics (Dehaene, 1997; Feigenson et al., 2004; Gallistel & Gelman, 1992) As Gallistel and Gelman (1992) argue, the symbolic representations of number do nothing to help us understand the mathematical operations in which we use them. Nothing about the appearance of the number 4 and the number 1, for instance, gives us any indication that their combination is 5 and could not possibly be 3. What makes mathematical operations meaningful for us, then, is an internal understanding of magnitude associated with each numerical symbol (Gallistel & Gelman, 1992). In this light, our symbolic concept of number, and, by extension, the mathematical operations in which these concepts are used, could be mapped onto our innate understanding of numerical magnitudes, or our ANS (Dehaene, 1997; Feigenson et al., 2004; Gallistel & Gelman, 1992).

Several studies suggest an intersection between processes associated with ANS and higher-level mathematical thinking. Symbolic representation of number, for instance, is thought to be a more advanced mathematical concept as it assigns arbitrary symbols to various levels of numerical magnitudes. The number 3, for example, is an arbitrary Arabic symbol assigned to a given magnitude that can be manipulated in higher-level mathematical operations (i.e. 3+x, 3/x, etc.). As discussed earlier, Moyer and Landauer (1967) showed that the discrimination of numbers presented as Arabic symbols (i.e. 3 and 8) is subject to the same distance effects that appear when discriminating physical representations of number presented in a way that prevents counting (i.e. briefly presented arrays of dots). The fact that distance effects appear in both discriminations relying on the ANS and in those relying on an understanding of symbolic number suggests a common or analogous processing mechanism for approximate and symbolic representations of number, further underscoring the relatedness of the ANS and higher-level mathematical processing.

This overlap between the ANS and higher-level math processes is also suggested by the flexibility adults demonstrate when moving from symbolic to approximate numerical representations (and vice versa). Whalen, Gallistel, and Gellman (1999) found that when asked to press a lever "x" amount of times at a pace too fast to count, adults were able to rely on their ANS to approximate the number of presses. Similarly, if a number of tones were presented at a pace too fast to count, adults were able to provide an estimate of the number of tones using Arabic symbols. Both studies from Whalen et al. (1999) and Moyer and Landauer (1967) suggest some level of overlap between the primitive ANS and higher-level mathematical processing, as constraints associated with ANS manifest themselves even when number is represented symbolically and adults are able to move flexibly from symbolic to approximate numerical representations.

It should be noted that there are several theories accounting for the process by which mathematical concepts such as number symbols are mapped onto a preexisting, nonverbal number sense. These theories remain controversial, as the data remain mixed and not enough is known yet about the relationship between the ANS and higher-level mathematical understanding (see Rips et al., 2008 for a review). This study, as it will become clear in the upcoming pages, is concerned with elucidating the association between the ANS and mathematical reasoning, an important first step in ultimately understanding the mechanisms by which these two may be mapped.

Approximate Number System and Mathematical Achievement

Although substantial research has been done on the ANS and its relation to mathematically-related tasks, few studies have examined the relation between the ANS and mathematical achievement. In a recent study, Haldberda, Mazzocco, and Feigenson (2008) examined whether individual differences in ANS acuity related to differences in individuals' level of mathematical achievement as measured by standardized testing. Interestingly, ANS acuity scores for 14-year-olds correlated with mathematical achievement scores dating all the way back to kindergarten. ANS acuity was assessed by presenting a participant with two sets of dots on a computer screen, in this case yellow and blue, and asking him or her to indicate which set had the greater number of dots. These arrays of dots were presented very rapidly at 200 ms, so as to make counting impossible. Differences in number magnitude between the blue and yellow sets differed according to either a 1:2, 3:4, 5:6, or 7:8 ratio. Individual Weber fractions were obtained from accuracy scores on this ANS acuity task and were then correlated with mathematical achievement scores measured by either the Woodcock-Johnson III (WJ-III) Calculation subtest or the Test of Early Math Achievement (TEMA). These findings underscore the relation between ANS and more advanced mathematical thinking that had been suggested previously (i.e. Dehaene, 1997).

One particular question that arises when examining the Halberda et al. (2008) study is the conceptualization of “math achievement.” Correlations were only reported between ANS acuity scores and a few choice subtests from the TEMA or WJ-III, which seems to narrow the operational definition of mathematical achievement. The WJ-III, in particular, has four subtests available to assess mathematical ability; however, in this study only correlations with one subtest, Calculation, were reported. Mathematical understanding involves far more than just calculation skills; it also can refer to a whole host of other abilities including (but not limited to) the ability to represent and manipulate objects geometrically, perform mental math, and apply mathematical knowledge to solve everyday problems. When assessing mathematical competency, then, it seems more applicable to include several measures of mathematical achievement so as to arrive at a more nuanced picture of an individual's competency that reflects ability in a multiplicity of mathematical domains. We can then determine to what extent ANS acuity relates to the several different components of mathematical achievement. The conceptualization of math achievement is an important issue, and is one we will address again in the coming pages dealing with the design of this particular study.

The Case for Space

While we have discussed the relation between the ANS and mathematical achievement, there does seem to be another quality, in addition to the ANS that could relate to mathematical competence: spatial acuity. There is evidence that suggests the ability to make spatial judgments (e.g., individual element size and cumulative surface area), like numerical judgments, is present in infants as young as six months of age (Brannon, Lutz, & Cordes, 2006; Clearfield & Mix, 1999; Feigenson, Carey, & Spelke, 2002; Gao, Levine, & Huttenlocher 2000) and in nonhuman animals (Cheng, 1986, Goutex et al. 2001, Tommasi & Polli, 2004, Vargas et al. 2004).

Animal research has shown that rats (Cheng, 1986), rhesus monkeys (Goutex et al. 2001), domesticated chicks (Tommasi & Polli, 2004), and even goldfish (Vargas et al. 2004) are capable of making spatial judgments. All animals in these studies were placed in a rectangular or rhombus-shaped enclosure and expected to search for food (or, in the case of the goldfish, an exit into a larger aquarium) in a specific corner, which they had been trained to locate. The layout of the enclosure was such that two corners were geometrically identical; that is, in the case of the rectangular enclosure, two corners would have the longer wall on the left side and a shorter wall on the right. Following training, the animal was disoriented and placed back into the enclosure, requiring it to reorient itself and locate the corner containing the food or exit. In all studies, the animal in question searched equally for food or an exit in the geometrically identical corners, suggesting they were attending to spatial stimuli to orient themselves and locate the correct corner (Cheng, 1986; Goutex et al. 2001; Tommasi & Polli, 2004; Vargas et al. 2004). Although it is unknown whether these spatial judgments are subject to any Weber-like constraints (i.e., less accuracy when the length of the walls is more similar), this research does provide support the theory that spatial judgments have evolutionary roots, and, like the ANS, may be innate.

Infant studies have shown both the existence of infant spatial judgments and their adherence to Weber's law. Using a habituation paradigm, Feigenson, Carey, and Spelke (2002) showed that six-month-old infants were capable of detecting a 1:2 difference in cumulative surface area, as indicated by longer looking times in trials that presented a novel image differing from the habituation images by a 1:2 difference in cumulative surface area. Similarly, Clearfield and Mix (1999) found six-month-olds dishabituated to test stimuli that differed from the habituation stimuli in cumulative contour length (i.e., perimeter) by a 1:2 ratio. These spatial discriminations in infants are subject to the same ratio limits as their numerical discriminations, as six-months' discrimination of surface area differences fail with harder ratios (e.g., 2:3), which is the identical ratio limit associated with numerical discrimination in infants of this age range (Brannon et al., 2006).

Though there are similarities between the ANS and spatial processes, including their existence in both infants and animals as well as their adherence to Weber's law, there is considerable evidence to suggest that these processes are not only similar, but overlap. This overlap between spatial and numerical processing is most salient in studies that examine the size congruity effect, a phenomenon in which the physical size of Arabic numbers interferes with one's ability to make numerical discriminations (Banks & Flora, 1977). Henik and Tzelgov (1982) found that, when simultaneously presenting adults with the numbers "3" and "5" the physical size of the numbers had an effect on the time it took adults to determine which number was larger in magnitude. If spatial information was congruent with magnitude differences, i.e. "3" was smaller in physical size than "5," reaction times were very fast. Conversely, adults were much slower to make numerical judgments if the spatial information was incongruent with magnitude difference, i.e. when "3" was larger in physical size than "5." Interestingly, this interference of spatial information on numerical judgments was applicable in the reverse as well; numerical information was also found to interfere with spatial judgments of physical magnitude (Henik & Tzelgov, 1982). For instance, adults were much slower to determine a difference in physical size if the larger number was "3" and the smaller number "5." Such interference of numerical information on spatial judgments and vice versa suggests that they have a common processing mechanism, or, at the very least, that they are separate processes with interfering outputs (Henik & Tzelgov, 1982).

In another study asking adults to compare magnitudes of dot arrays (Hurewitz et al., 2006), spatial information was shown to interfere with adult participants' nonverbal numerical judgments (ANS) and numerical information with nonverbal spatial judgments. Adults were slower, for instance, to determine the set of dots that was larger in number if the smaller number set was larger in cumulative surface area. Similarly, adults were slower to determine which set of dots was larger in cumulative surface area if the larger surface area set was smaller in number. Like Henik and Tzelgov's (1982) study, these results suggest a common processing mechanism or overlapping outputs for numerical and spatial information; however, in this case the numerical processing in question is more specific to ANS than symbolic number understanding.

Additionally, functional Magnetic Resonance Imaging (fMRI) and electroencephalography (EEG) studies have suggested that approximate number and spatial processing have common neurological underpinnings located in the right inferior parietal cortex (Dehaene & Cohen, 1999a; Dehaene & Cohen, 1999b; Dehaene et al. 1999; Walsh, 2003). Transcranial magnetic stimulation (TMS) studies have shown that inhibiting parietal functioning has a detrimental effect on both spatial and numerical processing, suggesting a shared location in the cortex (Bjoertomt, Cowey, & Walsh, 2002; Gobel, Walsh, & Rushworth, 2001). Extending these experimental results further, Walsh (2003) proposed that space and number (along with time) are part of a more general magnitude system dedicated to overall magnitude processing, whether it be spatial, numerical, or temporal magnitudes. A general magnitude system, if it exists, would help explain why spatial and numerical information can confound one another, as well as why spatial and numerical processes seem to be localized in the parietal cortex (Walsh, 2003). Furthermore, if numerical and spatial processes were indeed component parts of a general magnitude system, it is plausible that spatial acuity, like number acuity, would be related to mathematical achievement. In other words, the more refined one's spatial acuity abilities are, the higher level of his or her mathematical achievement. If both spatial acuity and ANS acuity were found to be related to mathematical achievement, this would provide further support for the numerical and spatial processes being interrelated component parts of a general magnitude system.

This Study

The purpose of this study is threefold; I first want to replicate Halberda, Mazzocco, and Feigenson's (2008) finding that there are individual differences in ANS acuity, and relatedly, whether individual differences in ANS acuity correlate with mathematical achievement. Second, I am interested in whether nonverbal spatial acuity, like ANS acuity, is related to mathematical achievement. Third, I am curious to see the extent to which number and spatial acuity correlate with different types of mathematical understanding (e.g., calculation, geometry, estimation, etc.)

This study will be similar in structure to that of Halberda et al. (2008), however with a few significant changes. First, in addition to ANS acuity tasks, this study will also include spatial acuity tasks. Second, a larger battery of standardized tests will be used to assess mathematical achievement. In addition to the Calculation subtest from the WJ-III, three more WJ-III math subtests will be used, as well as three subtests from the KeyMath Diagnostic Assessment. The use of additional WJ-III subtests as well as those from KeyMath provides the opportunity to see if any specific mathematical skills are more or less correlated with spatial acuity, and also allows for a broader operational definition of "mathematical achievement." Third, this study is the first of its kind to examine the relation between magnitude processing and mathematical achievement in adults at one point in time, and therefore will provide a valuable perspective on the nature of this interaction. My hypotheses for this study are as follows:

H1: There will be individual differences in ANS acuity.

$H1_A$: If a relation between ANS acuity and mathematics exists, there will be correlations between ANS acuity levels and the standardized test scores measuring mathematical achievement, consistent with Halberda et al. (2008).

H2. There will be individual differences in spatial acuity.

$H2_A$: If a relation between spatial acuity and mathematics exists, there will be correlations between spatial acuity levels and the standardized test scores measuring mathematical achievement, consistent with Halberda et al. (2008)

H3: If a general magnitude system is related to mathematical achievement, then *both* spatial acuity and ANS acuity levels will correlate with standardized test scores measuring mathematical achievement.

H4: Some subtests of the WJ-III and KeyMath will be more correlated with ANS acuity and/or spatial acuity than others. Although Halberda, Mazzocco, and Feigenson (2008) found WJ-III Calculation subtest scores correlated with ANS acuity, there is evidence suggesting that calculation skills are perhaps more learning-based and rely on other parts of the parietal lobe, specifically those associated with language (Dehaene & Cohen, 1999; Dehaene et al., 1991). In this light, certain subtests such as Mental Computation and Estimation may be more highly correlated with ANS and/or spatial acuity because they require use of estimation skills associated with nonverbal numerical and spatial judgments.

Methods

Participants

Participants were recruited from two introductory psychology courses and one cognitive development course at Emory University. In exchange for participation, they received research credit that contributed to their grade in one of these courses. The majority of the 37 participants in this study were female (62.2%) and psychology majors (51.2%). The next three most common majors were undecided (16.3%), business (7.0%), and economics (7.0%). The average age of participants was 19.27 years (*SD* = 1.08), and the average number of mathematics courses taken at a college level was 1.92 (*SD* = 1.32). Average SAT math scores were relatively high (*M* = 707.10, *SD* = 52.54), as were SAT verbal scores (*M* = 683.55, *SD* = 60.36) and ACT math scores (*M* = 30.0, *SD* = 2.58). All procedures were approved by the local ethics committee.

Procedure

Before beginning the study, participants signed a consent form, which outlined the tasks associated with their participation, as well as a math education background form, which asked participants to list previous math courses taken since high school and provide their SAT and/or ACT scores. Although data from the math education background form was not used in the final analyses, it was collected for pilot work for future research. Once finished with these forms, participants first completed a computerized task that assessed numerical and spatial estimation ability, and then they underwent a battery of standardized tests that assessed mathematical competence. On average, the total testing time for each participant was one hour and 40 minutes.

Computerized tasks: number and space estimation. The purpose of the computerized task was to assess each participant's ability to make numerical and spatial estimations without relying on explicit calculation strategies. To this end, the computerized portion was divided into two tasks: one that involved estimating numerical values (Number Task) and another that involved estimating the cumulative surface area (Space/Area Task). In both the Number and Space/Area Tasks, participants were asked to make ordinal judgments; that is, to choose the larger set with respect to number or cumulative surface area. The order of the Number and Space/Area Tasks was counterbalanced so that half the participants received the Number Task followed by the Space/Area task and half received the Space/Area task followed by the Number Task.

Stimuli. For both the Number and Space/Area Task, participants were shown images containing both brown and blue circles. The blue and brown colors used were controlled for luminance so that one color was not more salient than the other. Images were presented on a 30 x 37.5 cm computer screen at a visual angle of 1.36. In each image, the brown circles differed from the blue circles in terms of total number or cumulative surface area depending on the task. On the Number Task, one set of circles, either blue or brown, was larger in number than the other set, with equivalent cumulative surface area for each trial. Cumulative surface area varied across trials (Figure 1). In the Space/Area Task, one set of circles, either blue or brown, had a larger cumulative surface area than the other set, with total number of circles held constant across both groups. The total number of circles varied across trials (Figure 2). The controls for each task helped to ensure that participants made numerical or spatial judgments without an influence from the other variables, for instance numerical information interfering with spatial judgments, and vice versa. The color assigned to the larger number or larger surface area was counterbalanced across trials. Differences in either surface area or number between the two color groups were of one of the following ratios: 1:2, 3:4, 5:6 7:8, and 9:10 (see Tables 1 and 2). An equal number of trials, 26, were assigned to each ratio in both the Number and Space/Area Task. Additionally, the spacing of circles on each image was varied; specifically, they could be considered either clustered or spread out. Spacing was counterbalanced across tasks and ratios.

***Design*.** The computerized task was run through E-Prime software (Psychological Software Tools, Pittsburgh, PA). The task was divided into four blocks, two devoted to the Number Task and two devoted to the Space/Area Task. At the beginning of each section, participants were presented with an instructional screen explaining the task and allowing for questions of the experimenter before they began. For the Number Task, participants were asked to indicate which color of circles was largest in number. For the Space/Area Task, participants were asked to indicate which color of circles was greater in cumulative surface area. Participants were then given 10 practice trials followed by 60 test trials for each section, for a total of 40 practice trials and 240 test trials across all four blocks of the two tasks.

For each trial, a blank screen was presented for 250 milliseconds followed by the test image for 250 ms, and then a blank screen that gave participants unlimited time to indicate their answer.

Each image was presented for such a short amount of time to prevent participants from counting or using any other strategy other than instinctual estimation. Responses were indicated using a color-coded keyboard that had one blue and one brown sticker, each covering the left and right keys (e.g. blue on the left and brown on the right). These stickers were switched after completion of each block, so that one section of both the Number and Space/Area Tasks was allotted for the right key to be associated with 'brown' and left key to be associated with 'blue,' and one section of both the Number and Space/Area Task allotted for the reverse. Whether the left key or the right key started with the brown sticker was counterbalanced across participants and experimental conditions so that any participant could be given one of four possible trials: area tasks first with the right key associated with brown first, area tasks first with left key associated with brown first, number tasks first with right key associated with brown first, and number tasks first with left key associated with brown first. The entire computerized task took approximately 20 minutes to complete.

Standardized Testing. Once finishing the computerized portion of the experiment was completed, each participant was given a battery of standardized tests selected from the Woodcock-Johnson III Tests of Achievement (WJ-III) and KeyMath-3 Diagnostic Assessment (KeyMath-3). For each participant, all WJ-III subtests were administered prior to those from KeyMath-3. This was because use of the KeyMath-3 subtests was exploratory, and if unforeseen time constraints arose during the study it was important to ensure that all Woodcock-Johnson III subtests had been completed. Two participants did not have enough time to complete the final KeyMath subtest, Mental Computation/Estimation, and one participant did not have enough time to complete any KeyMath subtests.

Woodcock-Johnson (III) Tests of Achievement. Each participant completed six subtests of the Woodcock-Johnson III Tests of Achievement, four of which pertained to mathematical aptitude and two of which pertained to vocabulary knowledge. Each subtest was administered according to the guidelines set forth by the Woodcock-Johnson III Tests of Achievement manual.

Mathematical subtests. The four mathematical subtests from the Woodcock-Johnson III Tests of Achievement were: Calculation, Math Fluency, Applied Problems, and Quantitative Concepts. The Calculation subtest has a median reliability of .89 and is comprised of 45 problems that assess addition, subtraction, multiplication, and division ability; it also includes some questions on geometric, trigonometric, logarithmic, and calculus operations, though there are fewer questions of these types. (Mather & Woodcock, 2001). The Math Fluency subtest (reliability = .92) is a timed test which measures a participant's ability to solve addition, subtraction, multiplication, and division problems quickly and accurately; each participant had three minutes to complete 160 simple arithmetic problems such as "5x4" or "4+3." The Applied Problems subtest (reliability = .95) assesses participants' ability to solve mathematical story problems, requiring them to attend to the relevant information necessary to solve each problem using mainly addition, subtraction, multiplication, and division principles. The Quantitative Concepts subtest (reliability = .90) was divided into two sections, the first of which assessed a participant's knowledge of mathematical concepts and symbols and the second the participant's ability to find patterns in numerical sequences.

Scores from each mathematical subtest could be combined with others to obtain cluster scores, which provide a more comprehensive picture of mathematical aptitude. For each participant, three cluster scores were obtained using an aggregate of standard scores from various mathematical subtests: Broad Math, Math Calculation Skills, and Math Reasoning. The Broad Math cluster score (reliability = .96) was formed from a combination of Calculation, Math Fluency, and Applied Problems; it is considered a more comprehensive measure of math achievement. The Math Calculation Skills cluster score (reliability = .94), a combination of Calculation and Math Fluency, served as a general measure of computational skills and basic mathematical knowledge. Finally, the Math Reasoning cluster score (reliability = .97) measured a participant's level of math knowledge and reasoning; it was obtained through combining Applied Problems and Quantitative Concepts.

Vocabulary subtests. The two vocabulary subtests in the Woodcock Johnson III Tests of Achievement were Picture Vocabulary and Reading Vocabulary. The Picture Vocabulary subtest (reliability = .90) assessed lexical knowledge through the identification of objects (Mather & Woodcock, 2001). The Reading Vocabulary subtest (reliability = .92) was divided into three sections: Synonyms, Antonyms, and Analogies. The Synonyms section required reading words out loud and providing synonyms, the Antonyms section reading words aloud and providing antonyms, and the Analogies section reading three words of an analogy aloud and providing the fourth word to complete the analogy.

Procedure. The Woodcock-Johnson allows for subtests to be administered in any order provided the experimenter completes each subtest before moving on to the next (Mather & Woodcock, 2001), and in this study all math subtests were administered first followed by vocabulary subtests. Each participant received the Woodcock-Johnson subtests in the following order: first, Calculation, followed by Math Fluency, Applied Problems, Quantitative Concepts, Picture Vocabulary, and Reading Vocabulary. Though the basal and ceiling values varied depending on the subtest, each participant, regardless of the subtest, received one point for every correct answer and zero points for every incorrect or omitted answer to obtain a raw score for each mathematical subtest. These scores were then entered into Compuscore® and Profiles Program, a Woodcock-Johnson software, in order to obtain a standard score for each subtest. Compuscore® was also used to obtain cluster scores using the mathematical subtests.

KeyMath-3 Diagnostic Assessment. In addition the Woodcock-Johnson subtests, participants completed three subtests from KeyMath-3 Diagnostic Assessment Form A (KeyMath-3 DA): Numeration, Geometry, and Mental Computation and Estimation. The Numeration subtest assessed a participant's basic mathematical competence, using an assortment of questions that required an understanding of basic arithmetic, geometry, fractions, numerical estimation, and mathematical terminology. The Geometry subtest assessed a participant's geometry understanding by asking him or her to calculate dimensions of simple shapes, rotate objects mentally, and identify similarities and differences between two or more objects. The Mental Computation and Estimation subtest required the participant to mentally compute mathematical problems to either give an exact or estimate of the answer, and therefore provided an assessment of a participant's ability to perform mental math. When asked to give an estimate answer, participants had to provide an answer within a range provided by KeyMath. For each subtest, the basal was three and the ceiling was four; one point was given for every correct answer, and zero points for every incorrect or omitted answer to compile a raw score. Raw scores were converted to scaled scores using KeyMath-3 DA Normative and Interpretive Tables.

Results

ANS Acuity and Math Achievement

The first hypothesis concerning the ANS was that there would be individual differences in ANS acuity. In accordance with previous studies on the ANS, it was expected that performance on the Number Task would improve as the ratio difference between number sets increased (Barth, Kanwisher & Spelke, 2003; Cordes, Gelman & Gallistel, 2001; Halberda, Mazzocco, & Feigenson, 2008). The data do suggest that performance improved as the ratio difference between numbers increased because when plotting all participant data (Figure 3), the resulting slope significantly differs from zero (t (36) = 7.852, p <.001). Average performance at all ratios was significantly above chance with exception of performance at Ratio 7:8 (M = .49, SD = .11) which not significantly different from chance (t (36) = -.578, p = .567) and Ratio 5:6 (M = .53, SD =. 12) which was also not significantly different from chance (t (36) = 1.38, p = .177). Interestingly, average performance seems to hover around 50-60%, but significantly above chance (t (36) = 2.41, p < .05), for the four lowest ratio differences (M = .54, SD = .04) before jumping to 80% at the highest ratio difference of 1:2 (Figure 3). The average performance at ratio 1:2 (M = .76, SD =.16) was significantly different from the average performance of all other ratio differences (t (36) = -8.08, p < .001). That performance makes a significant jump from just over 50% to 80% only at the highest ratio difference may suggest that these adults are atypically inept at making numerical discriminations; 14-year olds in Halberda et al.’s (2008) study responded at a performance level just above 60% for their hardest ratio difference of 7:8, and we would expect adult acuities to have been at least this high. Another, more likely reason for these adults’ atypically low performance could be that this number discrimination task was particularly difficult for some other reason distinct from an inability to make numerical discriminations (see general discussion).

In order to assess whether there were individual differences in ANS acuity, ANS acuity scores were calculated. An individual’s acuity score was the ratio difference at which he or she could accurately differentiate between number sets 75% of the time. These acuity scores (M=1.79, SD= .44) were found by modeling each participant’s data so as to extract the ratio value that resulted when his accuracy was at 0.75, the same threshold value used by Halberda et al. (2008). Data were either modeled to a linear or logarithmic function to extract the ratio value at 0.75, and, in the case of the ANS data, a linear function was used because it fit the data significantly better than a logarithmic function (t (36) =11.439, p < .001). Six participants were dropped from the analyses because their data could not be fit to a linear model. Ultimately, there were individual differences found in ANS acuity scores, with values ranging from 0.02 to 2.55 (Figure 4).

The second hypothesis concerning the ANS was that there would be a correlation between the individual differences in ANS acuity and mathematical achievement; that is, the lower the ratio an individual could discriminate 75% of the time, the higher his or her math standardized test scores. In order to examine the relation between ANS acuity and mathematical achievement, standardized scores from the WJ-III and standard scores from KeyMath (Table 3) were correlated with ANS acuity scores (Table 4). Performance was relatively uniform for both the WJ-III (Figure 5) and Keymath (Figure 6) subtests. While none of these correlations were statistically significant, the WJ-III Calculation, WJ-III Math Fluency, KeyMath Numeration, KeyMath Geometry, and Keymath Mental Computation/Estimation subtests all had correlations within the range of .117 and .326, which is noteworthy because all correlations in this range were found to be significant in Halberda, et al.'s (2008). As Halberda et al. (2008) tested over 60 participants and this study only included 31 in its ANS analyses, the non-significance of this study's correlations may be attributed to a lack of statistical power.

Partial correlations controlling for more general intelligence factors, assessed by the WJ-III Picture Vocabulary and the WJ-III Reading Vocabulary subtests, were also performed to determine whether the correlation between ANS acuity and math achievement was due to differences in ANS acuity or more general performance factors (Table 5). In most cases, controlling for either Picture Vocabulary or Reading Vocabulary resulted in a decrease in correlations, and this decrease was nearly always greater when Reading Vocabulary was controlled. These decreases in correlation when controlling for general intelligence factors could suggest that general cognitive abilities play, at least some role, in the relationship between ANS acuity and mathematical achievement. Further, the fact that correlations decrease more when Reading Vocabulary is controlled could suggest that the skills assessed in the Reading Vocabulary subtest are more entwined in mathematical processes than the skills assessed in the Picture Vocabulary subtest. It is important to keep in mind, however, that these fluctuations in the data may be the result of lack of power, and should thus be interpreted with caution. Even with correlation decreases when controlling for general intelligence factors, however, it is important to note that the correlations associated with the WJ-III Calculation, WJ-III Math Fluency, KeyMath Numeration, KeyMath Geometry, and Keymath Mental Computation/Estimation all remained within the .117 to .326 range found significant in Halberda et al. (2008).

There were some instances where correlations increased when controlling for general performance factors, as was the case for the WJ-III Math Fluency subtest and WJ-III Broad Math and WJ-III Math Calculation Skills cluster scores (Table 5). Correlations for WJ-III Broad Math and WJ-III Math Calculation Skills showed particularly large changes when controlling for language measures moving from correlations that were close to 0 (r (29) = .078, p = .648 and r (29) =.009, p = .957, respectively) to values within the range of those found to be significant in Halberda et al. (2008). These increases in correlations when more general intelligence factors are controlled support the idea that the relation between ANS acuity and mathematical achievement is due to individual differences in ANS acuity rather than general cognitive performance factors. In all three cases, correlations were lower when Reading Vocabulary was controlled, although for WJ-III Math Fluency this difference was only 0.01. Again, that the increase in correlations is less when Reading

Vocabulary is controlled as opposed to Picture Vocabulary is consistent with the notion that Reading Vocabulary assesses skills more deeply connected to mathematical processing; however, as indicated above, fluctuations in the data may be due to lack of statistical power and thus should be interpreted cautiously.

Spatial Acuity and Math Achievement

The first hypothesis concerning spatial discrimination was that there would be individual differences in spatial acuity. As in the Number Task, it was expected that performance on the Space/Area task would increase as the ratio difference between sets increased, in accordance with Weber's law and previous studies on spatial discrimination (Brannon et al. 2006). For nearly all participants, performance on the Space/Area Task increased as the ratio difference of surface area increased (Figure 6). Performance leveled off as ratio differences approached the highest level of 2, which is consistent with standard psychophysical patterns of performance (Shepard, 1987; Stevens & Marks, 1965). Since overall performance on the Space/Area task was significantly better than the average performance on the Number Task (t (36) = 5.03, p <.001), it could be argued that participants were inherently better at making spatial discriminations. Another, perhaps more likely, possibility is that the spatial discrimination task was markedly easier for some other reason distinct from an inherent ability to make spatial discriminations more easily than numerical discriminations (see general discussion).

To assess whether there were individual differences in spatial acuity, spatial acuity scores were calculated. An individual's spatial acuity score was the ratio difference at which a he or she could accurately differentiate cumulative surface area between sets 75% of the time. In this case, a participant's data was fit to a logarithmic model so as to extract the ratio value at 0.75 accuracy, since a logarithmic model fit the data significantly better than a linear model (t (36)=8.949, p<.

001). Nine participants were dropped from the analyses because their data could not be fit to a logarithmic model. There were individual differences found in spatial acuity scores, (*M*=1.65, *SD*=.44*)*, ranging from 0.32 to 7.80 (Figure 7).

The second hypothesis concerning spatial discrimination was that individual differences in spatial acuity would correlate with math performance; more specifically, that the lower the ratio difference an individual could discriminate 75% of the time, the higher his or her scores on standardized math tests. To assess this hypothesis, spatial acuity scores were correlated with standardized scores from the WJ-III and standard scores from KeyMath (Table 6). Although none of these correlations were statistically significant (at $p < .05$), several subtests reached levels within the range found to be significant in Halberda et al. (2008), including WJ-III Calculation, WJ-III Math Fluency, WJ-III Quantitative Concepts, WJ-III Math Calculation Skills, KeyMath Geometry, and KeyMath Mental Computation/Estimation. Only one cluster score, WJ-III Math Calculation Skills, had a correlation within the range found significant in Halberda et al. (2008). As previously discussed, the non-significance of the correlations in this study could be attributed to a lack of power, given the low number of participants in our study compared to that of Halberda et al. (2008).

Partial correlations controlling for more general performance factors, assessed by the WJ-III Picture Vocabulary and the WJ-III Reading Vocabulary subtests, were also performed to determine whether the correlation between spatial acuity and math achievement was due to differences in spatial acuity or more general performance factors (Table 7). In many cases, several correlations decreased when controlling for general performance factors in comparison to those correlations found when intelligence factors were not controlled. It should be noted, however, that decreases in correlations found to be in the significant range for Halberda et al. (2008) remained well within that range, even with this decrease. There was also a substantial number of correlation increases when controlling for general performance factors, although no increases showed any marked difference that we saw in the ANS acuity data; all correlation increases happened to subtests that had higher correlations to begin with. Interestingly, as in the partial correlations with ANS acuity and mathematical achievement, decreases in correlation were larger when Reading Vocabulary was controlled and increases in correlation were larger when Picture Vocabulary was controlled. This pattern remains consistent with the possibility of skills assessed in the Reading Vocabulary subtest being more intimately connected with mathematical competence (see general discussion). However, as discussed earlier, these fluctuations in correlation values should be interpreted cautiously given that they do not reach statistical significance.

Correlations with Specific Mathematical Subtests

Our final hypothesis posited that some mathematical subtests would be more or less correlated with acuity scores than others, and the data are consistent with this possibility. The subtests WJ-III Calculation, WJ-III Math Fluency, KeyMath Geometry, and KeyMath Mental Computation/Estimation had consistently higher correlations with both spatial and numerical acuity scores. The cluster scores WJ-III Broad Math, WJ-III Math Calculation Skills and WJ-III Math Reasoning Skills also reached higher correlations with both spatial and numerical acuity scores. On the other hand the WJ-III Applied Problems was a subtest with consistently low correlations with both spatial and numerical acuity scores. Certain subtests, such as the WJ-III Quantitative Concepts and KeyMath Numeration, showed some differentiation in relation to ANS and spatial acuity scores in that high correlations existed in one magnitude relationship but not the other. For example, the WJ-III Quantitative Concepts had a higher correlation with spatial acuity scores (r (26)= -.167, p = .422), however its correlation with ANS acuity scores was much lower (r (29) = -.092, p = .643). Similarly, KeyMath Numeration showed a much higher correlation with ANS acuity scores (r (29) =

-.272, p = .161) than with spatial acuity scores (r (26) = .004, p = .984). While these patterns are being drawn from nonsignificant correlations and thus must be interpreted carefully, it is interesting to note that, at least preliminarily, certain biases may exist in mathematics that favor a relation with either spatial or numerical processes but not both. It is important to note, however, that the majority of these subtests were approaching significant relations with both spatial and numerical acuity scores, suggesting that special relationships between certain math abilities and certain magnitude acuities is more of a nuance than a rule.

Discussion

Although none of our results were statistically significant, trends in these data provide preliminary support for the existence of individual differences in both ANS and spatial acuity, as well as correlations between these individual differences and mathematical achievement. Importantly, specific mathematical subtests (e.g., calculation) were more related to acuity scores than others (e.g., applied problems), suggesting that magnitude estimation (whether number or space) may vary in its relation to specific mathematical abilities. These findings collectively support the idea of a general magnitude system (GMS) relating to several facets of mathematical achievement, and implications and theoretical questions surrounding these ideas will be discussed in greater detail below. Before arriving at these points of interest, however, it is necessary to discuss some less central results that were particularly puzzling.

To begin, average performance on the Space/Area Task was considerably higher than performance on the Number Task. As briefly discussed above, there are two potential explanations for this discrepancy: either these participants were simply less adept at making numerical discriminations, or the Number Task was more difficult than the Space/Area Task. That these particular adults are unable to make numerical discriminations seems unlikely; all were students at a selective university with high average SAT scores (*M* =683.55, *SD* = 60.36) and ACT scores (*M* = 30.0, *SD* = 2.58). Although some literature suggests numerical discriminations are easier than spatial (Brannon, Abbott, & Lutz, 2004; Cordes & Brannon, 2008), these studies are in infants. Some adult studies (Hurewitz, Gelman, & Schnitzer, 2006) suggest that processing spatial information is far easier and more automatic than processing numerical information, however we must ultimately consider that the average accuracy on the Number Task (*M* = .58, *SD* =. 10) was far lower than the average accuracy (about .76) for the 14-year olds in Halberda et al. (2008). Even if number information is more difficult to discriminate, we would expect adults to have at least the same, if not higher, average performance levels as teenagers.

Perhaps the more likely reason for this difference in performance is that the Number Task was more difficult than the Space/Area Task, and with some scrutiny of the stimuli, it seems this may have been the case. As discussed in the Methods section, cumulative surface area was equal between the two sets of circles in the Number Task; this was to prevent surface area from facilitating numerical judgments, as the larger number of circles would necessarily have the larger cumulative surface area without this control. As a result, however, controlling for cumulative surface area may have created interference from individual circle sizes, making this task more difficult. Because each set of circles needed to have the same cumulative surface area, the set with the larger number of circles necessarily had smaller individual circles on average. For example, if one set of circles totaled 5 in number, the other totaled 6 in number, and both sets needed to have a cumulative surface area of 2.237 (see Figure 1), then the set with 5 circles had an average circle size of .238 cm^2, while the set with 6 circles had an average diameter of .639 cm^2. Spatial information in the form of circle size, thus, may have conflicted with numerical judgments because the larger set of circles in number was also the smallest in average size of circles. As we know from previous studies on spatial and numerical judgments, when spatial information conflicts with numerical information, accuracy suffers (Henik & Tzelgov, 1982; Hurewitz, et al., 2006).

Another puzzling finding less central to our main hypotheses was the overall decrease in partial correlations when reading vocabulary was controlled as opposed to picture vocabulary. As discussed earlier, these fluctuations should be interpreted cautiously because the findings to not meet statistical significance; however, this trend does coincide with some literature on mathematical achievement, particularly studies that assess math aptitude in patients with brain lesions. Dehaene and Cohen (1997) found that a patient with damage to parts of the brain associated with verbal processing had difficulty in tasks that required him to use knowledge encoded verbally (i.e. mathematical times tables), suggesting that there is a verbal processing component to mathematical achievement. Albeit speculative, the skills assessed in Reading Vocabulary could account for some of the verbal components necessary for math achievement. More participants would need to be added to the data to see if this trend persists.

Theoretical Questions

The first major theoretical question to consider is the nature of a general magnitude system (GMS). What exactly does this GMS look like? Walsh (2003) makes two suggestions for the layout of a GMS, the first of which suggests space, number, (and time) are all processed individually and then compare and communicate information afterward. In this model, different dimensions of magnitude are all processed uniquely and the overlap between them occurs from communication after they are processed. The second model suggests that space, number, and time involve shared neural mechanisms, thus the overlap between them occurs during initial processing. There is also the possibility that these models are not mutually exclusive, and either elected for use separately or simultaneously depending on the stimuli being processed or perhaps an individual's developmental age. For example, number and space could exist as overlapping systems with more common processing mechanisms early in development, allowing for numerical cues to be extracted from spatial information, such as cumulative surface area. As individuals learn more about number and numerical concepts, the number processing mechanisms could become more differentiated from the spatial, connecting instead to areas associated with more explicit mathematical reasoning (i.e. language). As a result, connections between spatial and numerical processing occur as communication after the initial information has been processed.

Although this study cannot speak directly to neural mechanisms, the data seem more consistent with the idea that number and space share common processing mechanisms. Initially, it would seem that because this study assessed acuity levels separately for spatial and numerical discrimination, it necessarily treated space and number processes as separate entities potentially encompassed in a larger GMS. However, my results seem to show more support for a common processing mechanism for space and number. For example, in the Number Task, I have suggested that specific spatial cues (i.e. individual circle size) affected number acuity, perhaps leading to worse performance because of the inconsistent spatial information (i.e., the greater number set was smaller, on average, with respect to circle size). If spatial and numerical information is truly processed independently, then we would not expect any conflicting spatial information to be detrimental to numerical judgments; however, this did not seem to be the case. In fact, that our experimental controls were necessary at all (for surface area in the Number Task and number in the Space/Area Task) might be considered indirect evidence for an overlapping processing mechanism that must be manipulated in order to arrive at purely spatial or numerical judgments.

With respect to the layout of the GMS that relates to mathematical ability, our results again support a common processing mechanism. Only two of the eight subtests with correlations in the range found significant by Halberda et al. (2008) were only related to either spatial or numerical acuity; the remaining six showed high correlations with both spatial and numerical acuity. If the GMS consisted of disparate space and number processing systems, then we would likely see different relations between space, number, and the various mathematical subtests. It is important to note here that I have not completely ruled out general intelligence factors here, as I have only used two vocabulary subtests, so it is possible that commonalities in the spatial and ANS data reflect more general performance factors. However, the fact that, at least preliminarily, correlations between math and spatial acuity and math and ANS look fairly similar suggests more commonalities between spatial and numerical processing systems, which is more in line with Walsh's (2003) conception of a GMS existing as a common processing mechanism between space, number and time. Again, it is important to further note that none of these results are significant, and an increase in the number of participants may alter this pattern even to the point of supporting the model of a GMS with separate number and space processing centers.

In addition to discussing the possible formats of magnitude systems, it is also important to discuss the consistency of the other major concept in this study: mathematical achievement. This study broadened the operational definition of mathematical achievement to include more subtests than were reported in Halberda et al. (2008), including three additional WJ-III subtests and three subtests from KeyMath. I had initially predicted that some subtests would be more or less correlated with spatial/ANS acuity scores, and this seemed to be the case. While several of the subtests were correlated with both ANS/spatial acuity scores, one score, in particular, was not correlated with either acuity score: Applied Problems. This supports the idea that mathematics can be defined as several component parts spanning variety of abilities, as opposed to simply the calculation abilities reported in Halberda et al.'s (2008) paper, and further, that basic estimation processes employed during numerical and spatial processing may not relate equally to all aspects of math. Additionally, these results also support the notion that magnitude systems can correlate with a variety of mathematical abilities rather than simply the ones we may consider to be the most obvious, for example, ANS with calculation and spatial acuity with geometry. It would be critical to determine whether increasing the power of this data set would result in stronger correlations for all subtests, or if certain subtests would emerge as more significant than others. Additionally, it would be informative to assess whether increased power would reveal more biases of subtests; for instance, that some subtests would only have significant correlations with ANS acuity as opposed to spatial acuity.

The final major theoretical question that arises here is a question of directionality: is it the approximate number/spatial systems (or perhaps GMS) that impact mathematical achievement, or is it mathematical achievement impacting the precision of the approximate number/spatial systems? In other words, are people who have higher spatial/ANS acuity better at math because these estimation abilities serve as a foundation for higher-level mathematical reasoning, or is it because better math skills lead to a higher accuracy in estimation skills? Although this study cannot decisively answer questions of directionality, it is interesting to consider the plausibility and implications of each scenario.

Several researchers (Dehaene, 1997; Feigenson et al., 2004; Gallistel & Gelman, 1992; Haberda et al., 2008) are proponents of the notion that magnitude systems impact later mathematical achievement. The primacy of an approximation system, for instance, suggests that it is the founding point for all of mathematical thought; we begin with an innate mechanism for understanding quantity, and this mechanism serves as the basis for which we apprehend symbolic representations of number and the mathematical operations in which they are involved. This perspective provides a nativist account on mathematical achievement; if math performance is based on the acuity of an innate magnitude system, then we are either endowed with the tools for mathematical success or we are not. Mathematical success, then, can be attributed to a highly refined magnitude approximation mechanism; those who are good at math may find success because they are able to represent numerical (and perhaps spatial) values more precisely, which comes in handy when it is time to manipulate these numbers mathematically. As the mathematician Wim Klein says (in Dehaene, 1997), “Numbers are friends to me, more or less. It doesn’t mean the same for you, does it, 3,844? For you it’s just a three and an eight and a four and a four. But I say: ‘Hi, 62 squared!’”

Support for this nativist account comes in part from studies that attempt to change the precision of magnitude systems. For example, Dehaene (1997) attempted to eradicate the distance effect by having adults extensively practice a comparison task. Adults were asked to indicate using a keyboard whether the digit they were presented with was smaller or larger than “5,” pressing the left-hand key if the digit was smaller, and the right-hand key if the digit was larger. The task was fairly simple and straightforward in that only numbers used in the task were 1, 4, 6, and 9, so the task was essentially to push the left key for “1” and “4,” and push the right key for “6” and “9.” Surprisingly, even after 1,600 training trials, adults remained slower in indicating an answer if the numbers presented were closer to “5” (i.e. 4 or 6) than if they were farther away (i.e. 1 or 9). The fact that such a salient property of the ANS remains unaltered in spite of a direct attempt to change it makes the notion of mathematical experiences refining or changing the composition of the ANS or other magnitude systems seem unlikely. Further cementing the unlikelihood of math achievement impacting magnitude systems is the fact that distance effects appear even when professional mathematicians make these numerical discriminations (Dehaene, 1997), suggesting that neither directed practice nor mathematical experience is capable of altering properties of magnitude systems, namely the distance effect. In this light, then, directionality of this relation between magnitude approximation and math achievement goes from basic approximation/estimation abilities (perhaps in the GMS) to knowledge of formal mathematics and not the other way around.

There is evidence, however, that mathematical achievement can in fact impact lower-level magnitude systems. More specifically, that more mathematical experience and/or achievement can lead to a refinement of the ANS. The most cogent evidence for this relation is found in cultural studies of individuals who live without knowledge of even basic mathematics. Peter Gordon's work (2004) with members of the Pirahã tribe in Brazil provides a particularly salient example of the effect a lack of mathematical experience can have on the ANS. The Pirahã have an extremely limited counting system, with only words for 1 and 2, referring to the rest of quantities with a word that translates to "many." Number words, even 1 and 2, are conceptualized with a degree of fuzziness; the word for "2", for example, can also be used to refer to 3 or 4, and the number 4 itself represented by anywhere from 3 to 5 fingers. Given this highly imprecise system of number, it is no exaggeration to say members of the Pirahã do not have the extensive experience with formal mathematics and number as is seen in our culture and other more technologically-advanced societies.

When Gordon (2004) asked members of the Pirahã to perform tasks that assessed nonverbal numerical reasoning skills, he found their numerical estimation abilities were poor; they were only capable of representing numbers fewer than 3, and when the task required greater cognitive effort, even the ability to represent these small numbers was compromised. Gordon's (2004) findings suggest that mathematical input can have an effect on the magnitude systems that support basic estimation processes; without exposure to a mathematical language that quantifies these magnitudes, and by extension to mathematical procedures and concepts, the Pirahã's ANS may be less refined. The likely answer, then, for directionality of the magnitude/math relationship is that it is bidirectional; some properties of our mathematical understanding may be founded in innate brain mechanisms, with basic estimation abilities serving as the foundation for later learning of various mathematical operations. In the opposite direction, formal training in number and mathematics can impact estimation abilities by refining them, and thus environmental experience molds our inherent conceptions of number.

Further Directions and Conclusions

Future directions for studies on this magnitude/mathematics relation should consider correlating performance on duration tasks with mathematical achievement to see if the final link in the GMS according to Walsh (2003), time, is also correlated with math performance. If support were found for this relation, then these results would further cement the idea that a GMS is related to math performance. If no support is found, then either a GMS is not related to math or perhaps the landscape of the GMS shifts with math ability; in other words, number and space have common processing mechanisms that relate to math ability but are still disparate from temporal processes. In this instance, support for both models of the GMS would be found; some aspects of it are overlapping, and others remain separate processes. In any case, it would be interesting to see whether temporal processes have any relation to mathematical achievement

Further research may examine the directionality of this magnitude/mathematics relation. Taking more upper-level mathematics courses, for instance, may refine approximation systems; although Dehaene (1997) found that qualities of these approximation systems were static, only a handful of ratios were tested, and perhaps improvements could have been detected if more ratios were used. Additionally, Dehaene (1997) only used Arabic numbers in training individuals, so whether practicing numerical estimation abilities using non-symbolic stimuli (i.e., dot arrays) would lead to increased precision in numerical estimation is certainly another empirical question. If it were possible to hone one's estimation abilities through practice with dot arrays, for example, then it would be critical to assess whether this refinement in number discrimination affected mathematical understanding.   Findings from these studies could have enormous educational consequences, as understanding what mechanisms can be used to improve math understanding can lead to more effective math instruction in the classroom (i.e., focusing more on the magnitude properties of number rather than the symbolic).

In sum, this study provides preliminary support for both spatial and ANS acuity, potentially having common processing mechanisms in a GMS, relating to several facets of mathematical achievement. This research is in agreement with Halberda et al. (2008) research on the ANS and mathematical achievement, and is the first to demonstrate this relation in adults. Future research should consider the final aspect of the GMS, temporal processing, in determining whether it is larger GMS or simply spatial and numerical processes alone that are relating to math ability. Additionally, this research contributes to the foundational knowledge regarding the relation between magnitude estimation and math achievement that is necessary to determine the directionality of this relation, which will undoubtedly provide greater insight into the nature of mathematical understanding and achievement.

References

Banks, W.P., & Flora, J. (1977). Semantic and perceptual processes in symbolic comparisons. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 278-290.

Barth, H. Kanwisher, N. & Spelke, E. (2003). The construction of large number representations in adults. *Cognition, 86,* 201-221.

Brannon, E.M., Abbott, S., & Lutz, D.J. (2004). Number bias for the discrimination of large visual sets in infancy. *Cognition, 93,* B59-B68.

Brannon, E.M., Lutz, D., & Cordes, S. (2006). The development of area discrimination and its implications for numerical discrimination in infancy. *Developmental Science, 9*, F59-F64.

Bjoertomt, O., Cowey, A., & Walsh, V. (2002). Spatial neglect in near and far space investigated by repetitive transcranial magnetic stimulation. *Brain, 125,* 2012-2022.

Clearfield, M.W., & Mix, K. S. (1999). Number versus controug length in infants' discrimination of small visual sets. *Psychological Science, 10,* 408-411.

Cordes, S. & Brannon, E.M. (2008). Difficulties of representing continuous extent in infancy: Using number is just easier. *Child Development, 79*, 476-489.

Cordes, S., Gelman, R., & Gallistel, R. (2001). Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychonomic Bulletin & Review, 8,* 698-707.

Cheng, K. (1986). A purely geometric module in the rat's spatial representation. *Cognition, 23,* 149-178.

Dehaene, S. (1997). *The number sense*: *how the mind creates mathematics.* New York: Oxford University Press.

Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends in Neuroscience, 21*, 355-361.

Dehaene, S. & Cohen, L. (1997). Two mental calculation systems: A case study of severe acalculia with preserved approximation. *Neuropsychologia, 29,* 1045-1074.

Dehaene, S. & Cohen, L. (1997). Cerebral pathways for calculation: Double dissociation between rote verbal and quantitative knowledge of arithmetic. *Cortex, 3,* 219-250.

Dehaene, S., Spelke, E., Pinel, P. Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain imaging evidence. *Science,* 284, 970-974.

Feigenson, L., Carey, S., & Spelke, E. (2002). Infants' discrimination of number vs. continuous extent. *Cognitive Psychology, 44,* 33-66.

Feigenson, L., Dehaene, S. & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8,* 307-314.

Gallistel, C.R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition, 44*, 43-74.

Gao, F., Levine, S.C., & Huttenlocher, J. (2000). What do infants know about continuous quantity? *Journal of Experimental Child Psychology*, *77,* 20-29.

Gobel, S., Walsh, V. & Rushworth, M.F.S. (2001). The mental number line and the human angular gyrus. *Neuroimage,* 14, 1278-1289.

Goutex, S., Thinus-Blanc, C., Vauclair, J. (2001). Rhesus monkeys use geometric and nongeometric information during a reorientation task. *Journal of Experimental Psychology, 130,* 57-59.

Halberda, J. Mazzocco, M., & Feigenson, L. (2008). Individual differences in nonverbal number acuity predict maths achievement. *Nature, 455*, 665-668.

Henik, A. & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & Cognition, 10*, 389-395.

Hurewitz, F., Gelman, R., & Schnitzer, B. (2005). Sometimes area counts more than number. *Proceedings of the National Academy of Sciences 103*, 19599-19604.

Meck, W.H., & Church, R.M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*, 320-334.

Moyer, R. S. & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature, 215*, 1519-1520.

Pica, P., Lemer, C., Izard, V., Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science, 5695,* 499-503.

Rips, L.J., Bloomfield, A., & Asmuth, J. (2008). From numerical concepts to concepts of number. *Behavioral and Brain Sciences, 31,* 623-687.

Rumbaugh, D.M., Savage-Rumbaugh, S. & Hegel, M.T. (1987). Summation in the chimpanzee (Pan troglodytes). *Journal of Experimental Psychology Animal Behavior Processes, 13,* 107-115.

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science, 237,* 1317-1323.

Stevens, J. C., & Marks, L.E. (1965). Cross-modality matching of brightness and loudness. *Proceedings of the National Academy of Sciences*, *54,* 407-411.

Tommasi, L. & Polli, C. (2004). Representation of two geometric features of the environment in the domestic chick (*Gallus gallus). Animal Cognition, 7,* 53-59.

Vargas, J.P., Lopez, J.C., Salas, C., & Thinus-Blanc, C. (2004). Encoding of geometric and featural spatial information by goldfish (*Carassius auratus). Journal of Comparative*

Psychology, 118, 206-216.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space, and quantity. *Trends in Cognitive Sciences, 7,* 483-488.

Whalen, J., Gallistel, C.R., & Gelman, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science, 10,* 130-137.

Woodruff, G. & Premack, D. (1981). Primative (sic) mathematical concepts in the chimpanzee: proportionality and numerosity. *Nature, 293,* 568-570.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358,* 749-750.

Xu, F. & Spelke, E. (2000). Large number discrimination in 6-month old infants. *Cognition, 86,* 201-221.

Xu, F., Spelke, E.S., & Goddard, S. (2005). Number sense in human infants. *Developmental Science, 8,* 88-101.

Table 1

Stimulus-related values used in the Number Task

	Large Total Number of Circles			Small Total Number of Circles		
Ratio	Set 1 Total	Set 2 Total	Surface Area	Set 1 Total	Set 2 Total	Surface Area
1:2	9	18	5.157	4	8	3.493
3:4	12	16	4.058	6	8	2.794
5:6	15	18	2.896	5	6	2.237
7:8	14	16	3.289	7	8	2.433
9:10	19	20	4.003	9	10	2.621

Note: All measurements are in cm. Within each ratio, one set was created using a large total number of circles on the screen, and one set was created using a small total number of circles on the screen. Set 1 Total gives the total number of circles in set 1. Set 2 Total gives the total number of circles in Set 2. Surface Area gives the cumulative surface area for circles in Set 1 and Set 2, as surface area was equivalent for each set of circles in the Number Task.

Table 2

Stimulus-related values used in the Space/Area Task

	Large Total Number of Circles			Small Total Number of Circles		
Ratio	Set 1 Area	Set 2 Area	Number	Set 1 Area	Set 2 Area	Number
1:2	3.595	7.180	14	2.779	5.558	4
3:4	3.673	4.898	12	2.488	3.317	5
5:6	4.035	4.842	18	3.171	3.806	7
7:8	4.106	4.692	20	3.627	4.144	8
9:10	4.176	4.640	16	2.826	3.120	9

Note: All measurements are in cm. Within each ratio, one set was created using a large total number of circles on the screen, and one set was created using a small total number of circles on the screen. Set 1 Area gives the cumulative surface area of the circles in Set 1. Set 2 Area gives the cumulative surface area of the circles in Set 2. Number gives the total number of circles in Set 1 and Set 2, as number was equivalent for each set of circles in the Space/Area task.

Table 3

Scores from the WJ-III and KeyMath Subtests

Subtest	*M*	*SD*
WJ-III Calculation	121.30	11.00
WJ-III Math Fluency	111.03	13.87
WJ-III Applied Problems	110.54	15.18
WJ-III Quantitative Concepts	114.32	9.36
WJ-III Picture Vocabulary	99.60	11.73
WJ-III Reading Vocabulary	103.84	7.49
KeyMath Numeration	13.14	2.26
KeyMath Geometry	11.860	2.38
KeyMath Mental Computation & Estimation	12.0	2.68
WJ-III CS Broad Math	117.27	13.14
WJ-III CS Math Calculation Skills	119.97	11.31
WJ-III CS Math Reasoning Skills	113.87	12.35

Notes: Scores from the WJ-III are standardized and scores from KeyMath are scaled.

Table 4

Correlations between scores from WJ-III and KeyMath subtests ANS acuity scores

Subtest	*r*	*p*
WJ-III Calculation	**-.299**	.103
WJ-III Math Fluency	**-.234**	.206
WJ-III Applied Problems	-.088	.636
WJ-III Quantitative Concepts	-.107	.566
WJ-III Reading Vocabulary	-.100	.594
WJ-III Picture Vocabulary	-.020	.916
KeyMath Numeration	**-.293**	.116
KeyMath Geometry	**-.261**	.163
KeyMath Mental Computation and Estimation	**-.346**	.066
WJ-III CS Broad Math	.078	.648
WJ-III CS Math Calculation Skills	.009	.957
WJ-III CS Math Reasoning Skills	**-.211**	.210

Note: *r* values in bold indicate magnitudes within the range found to be significant in Halberda, Mazzocco, & Feigenson's (2008) study. Negative values indicate that the higher performance on estimation tasks, the higher scores on mathematical tests.

Table 5

ANS Partial Correlations Controlling for Reading Vocabulary or Picture Vocabulary

	Picture Vocabulary		Reading Vocabulary	
	r	p	r	P
WJ-III Calculation	**-.279**	.151	**-.272**	.161
WJ-III Math Fluency	**-.247**	.204	**-.246**	.207
WJ-III Applied Problems	-.088	.655	-.056	.779
WJ-III Quantitative Concepts	-.092	.643	-.058	.771
KeyMath Numeration	**-.272**	.161	**-.255**	.191
KeyMath Geometry	**-.288**	.137	**-.261**	.180
KeyMath Mental Comp/Est	**-.347**	.071	**-.350**	.068
WJ-III CS Broad Math	**-.213**	.258	**-.196**	.300
WJ-III CS Math Calculation Skills	**-.325**	.079	**-.310**	.095
WJ-III CS Math Reasoning	**-.119**	.531	-.087	.646

Note: *r* values in bold indicate magnitudes within the range found to be significant in Halberda, Mazzocco, & Feigenson's (2008) study, and their non-significance may be due to lack of power.

Table 6

Correlations Between Subtests and Spatial Acuity Scores

Subtest	*r*	*P*
WJ-III Calculation	**-.227**	.246
WJ-III Math Fluency	**-.159**	.419
WJ-III Applied Problems	-.072	.715
WJ-III Quantitative Concepts	**-.159**	.419
WJ-III Reading Vocabulary	**-.215**	.273
WJ-III Picture Vocabulary	.043	.826
KeyMath Numeration	-.009	.966
KeyMath Geometry	**-.217**	.277
KeyMath Mental Computation and Estimation	**-.209**	.316
WJ-III CS Broad Math	**-.157**	.424
WJ-III CS Math Calculation Skills	**-.236**	.227
WJ-III CS Math Reasoning Skills	**-.123**	.532

Note: *r* values in bold indicate magnitudes within the range found to be significant in Halberda, Mazzocco, & Feigenson's (2008) study, and their non-significance may be due to lack of power.

Table 7

Spatial Partial Correlations Controlling for Reading Vocabulary or Picture Vocabulary

Subtest	Picture Vocabulary		Reading Vocabulary	
	r	p	r	p
WJ-III Calculation	**-.215**	.314	**-.205**	.161
WJ-III Math Fluency	**-.143**	.505	**-.182**	.395
WJ-III Applied Problems	-.095	.658	-.023	.913
WJ-III Quantitative Concepts	**-.167**	.437	-.102	.635
KeyMath Numeration	.004	.984	.102	.635
KeyMath Geometry	**-.249**	.242	**-.146**	.496
KeyMath Mental Comp/Est	**-.204**	.339	**-.151**	.481
WJ-III CS Broad Math	**-.161**	.422	**-.123**	.540
WJ-III CS Math Calculation Skills	**-.232**	.244	**-.217**	.277
WJ-III CS Math Reasoning	**-.137**	.496	-.070	.729

Note: *r* values in bold indicate magnitudes within the range found to be significant in Halberda, Mazzocco, & Feigenson's (2008) study,and their non-significance may be due to lack of power.

Figure 1

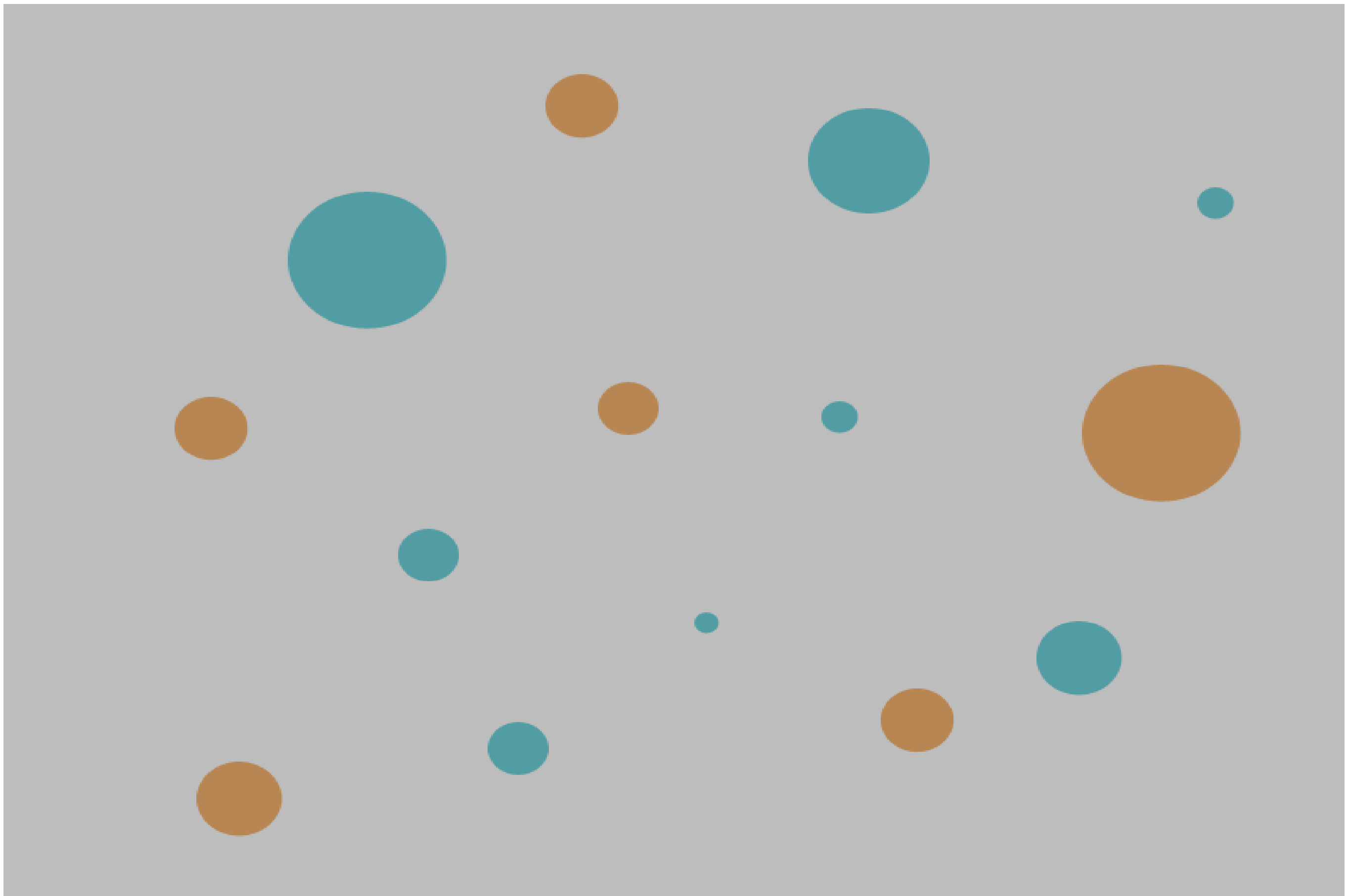

Figure 1. Example of stimuli used in the Number Task. In this example, the numerical difference varies by a ratio of 3:4, as there are 6 brown circles and 8 blue circles. Cumulative surface area is equivalent (2.794 cm^2) for both sets of circles. The spacing of circles on the screen is spread out rather than clustered.

Figure 2

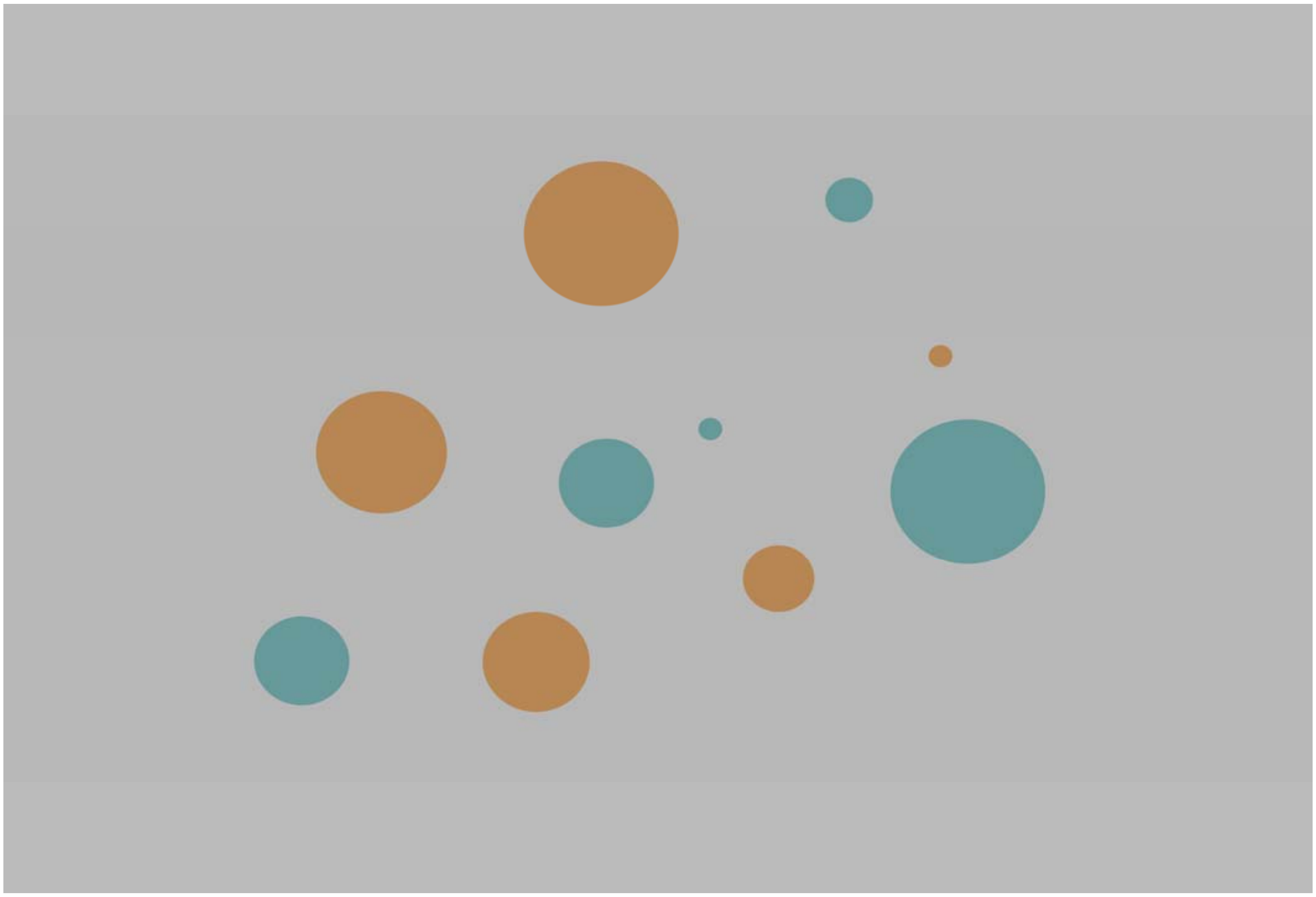

Figure 2. Example of stimuli for the Space/Area Task. In this example, the difference in cumulative surface area between sets varies by a ratio of 3:4, as the brown set has a cumulative surface area of 3.317 cm^2, and the blue set has a cumulative surface area of 2.488 cm^2. The total number of circles is equivalent (5) for both sets of circles. Spacing is clustered rather than spread out.

Figure 3

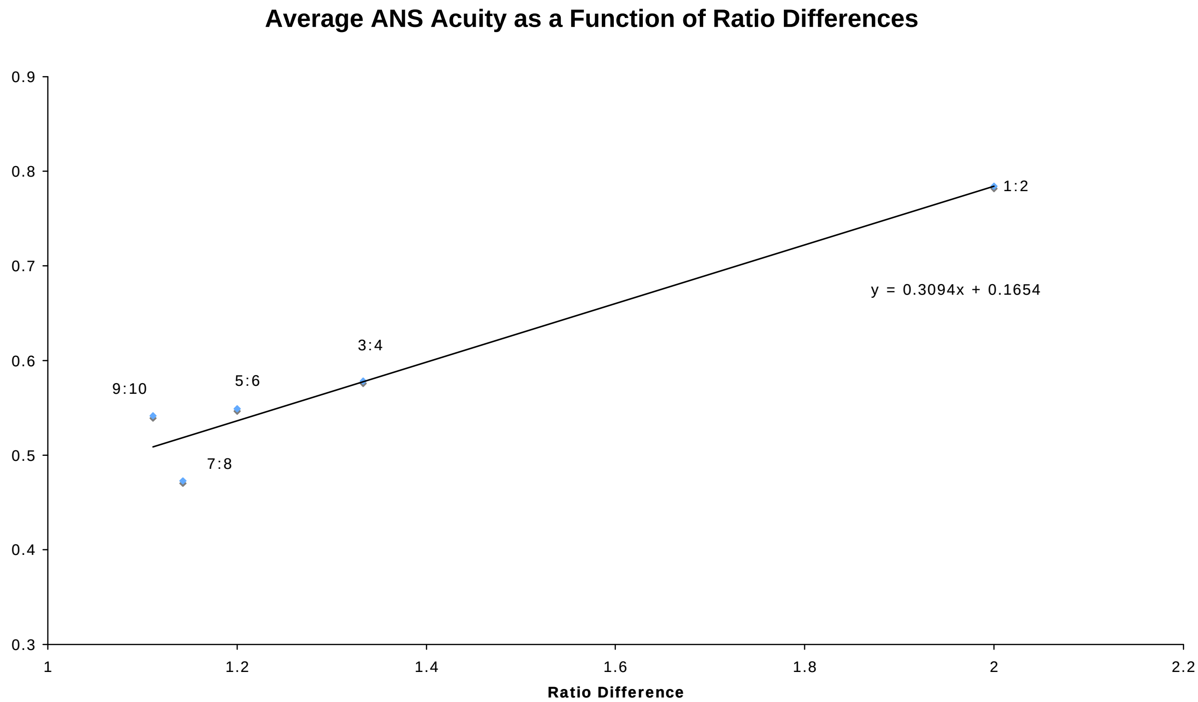

Figure 3. Mean accuracy (proportion) on the Number Task for all participants at each ratio. Average accuracy hovers between 50 and 60 percent before jumping to nearly 80% at Ratio 1:2.

Figure 4

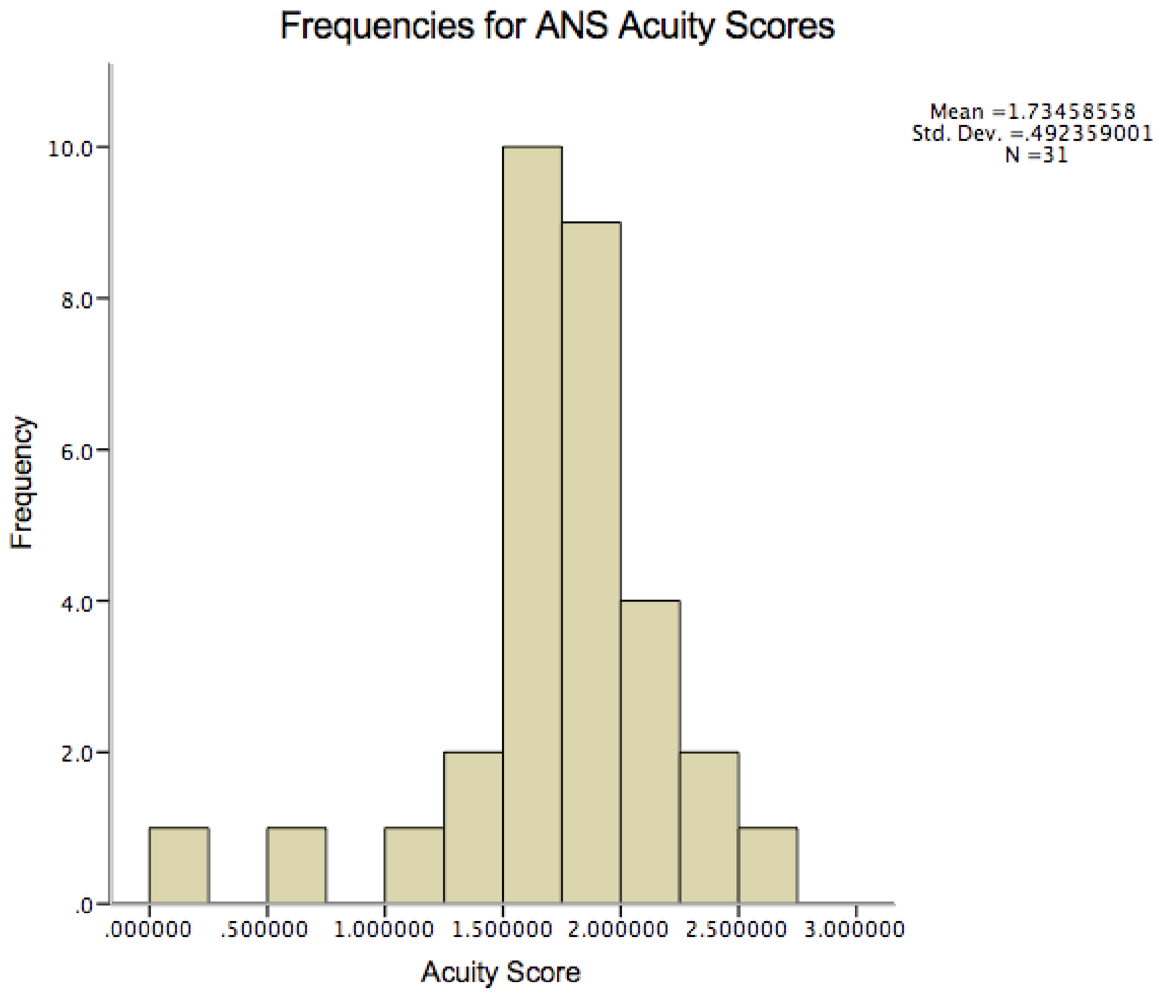

Figure 4. Frequency of ANS acuity scores for all participants whose data could be modeled to a linear function. The higher an individual acuity score, the lower the lower the ability to make numerical discriminations.

Figure 5

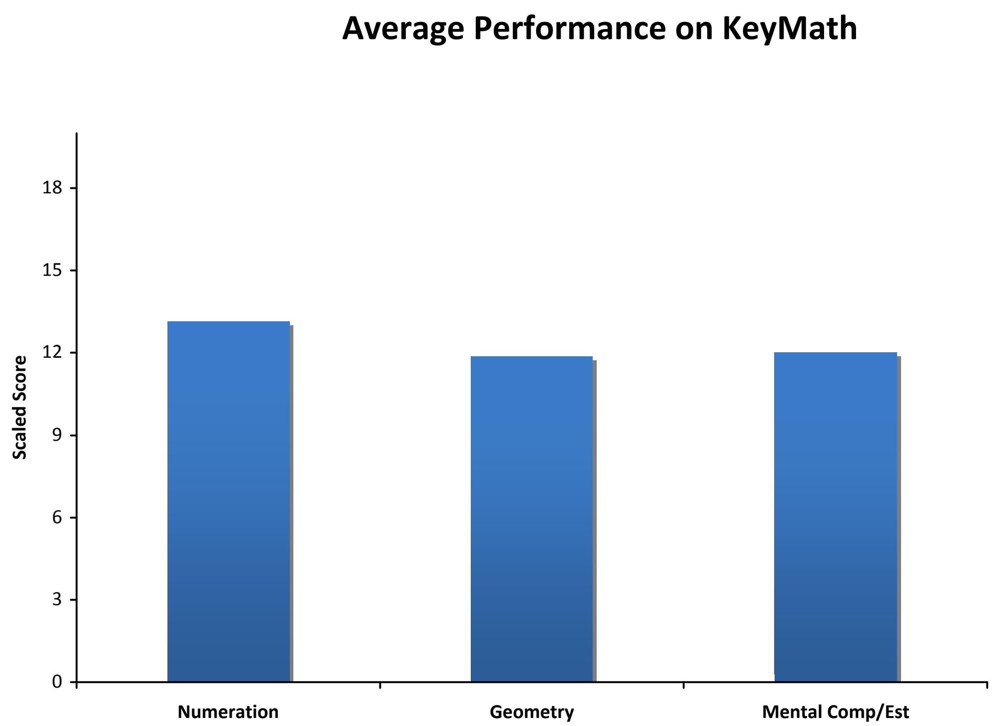

Figure 5. Average performance for all participants on KeyMath Subtests. Scores are scaled, not standardized, meaning they have a *M* =10 and *SD* = 3.

Figure 6

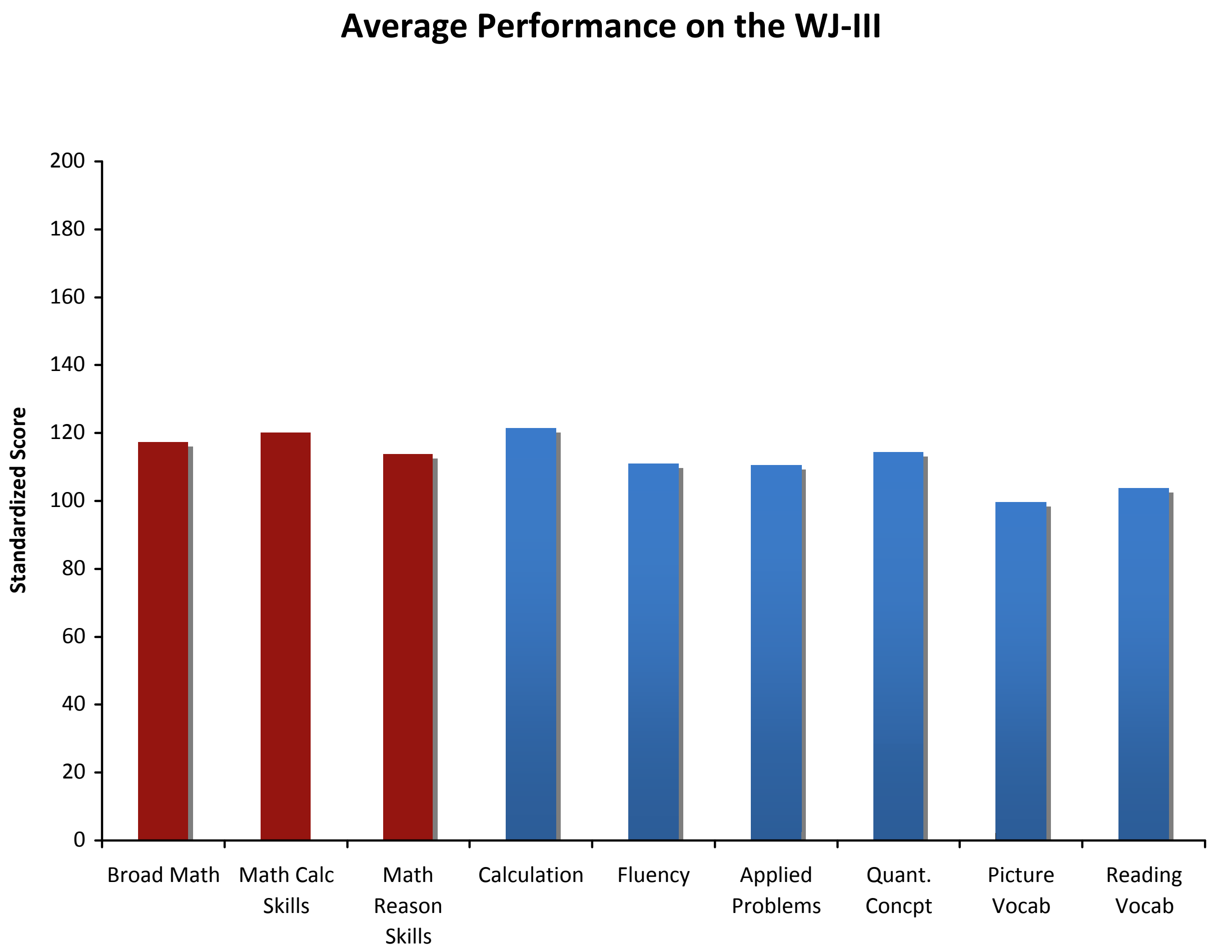

Figure 6. Average performance for all participants on WJ-III subtests and cluster scores. Red bars indicate cluster scores; blue bars indicate subtests.

Figure 7

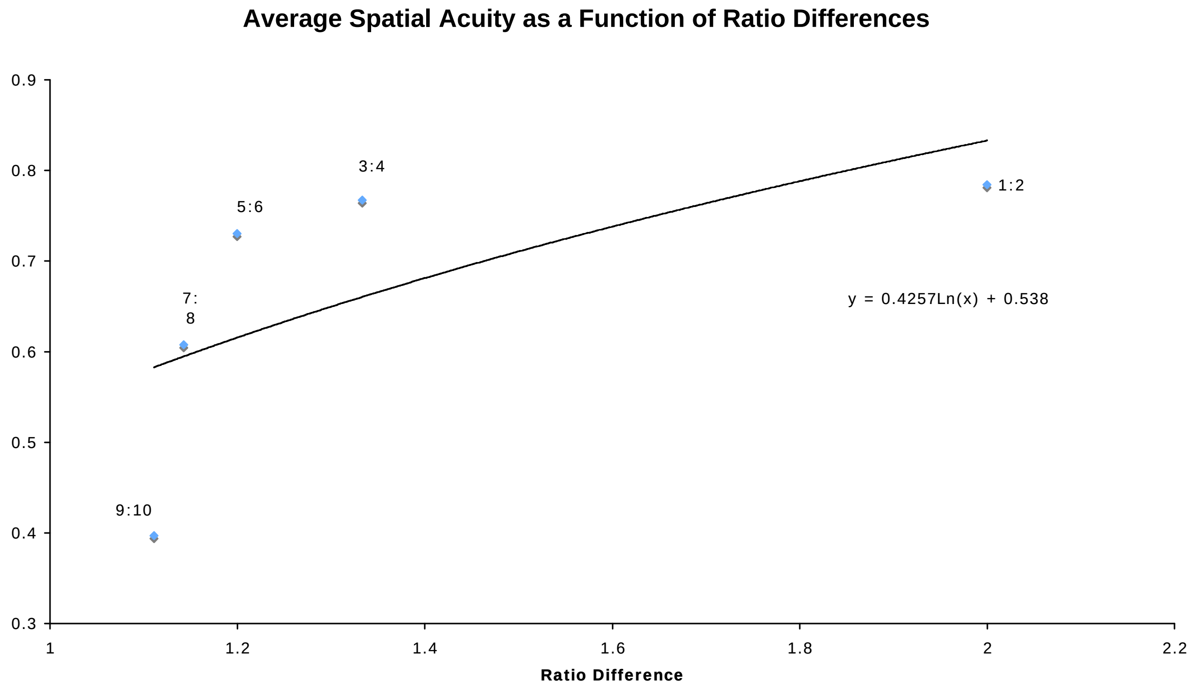

Figure 7. Mean accuracy (proportion) on the Space/Area Task for all participants at each ratio. Average accuracy gradually increases as the ratio difference between sets increases.

Figure 8

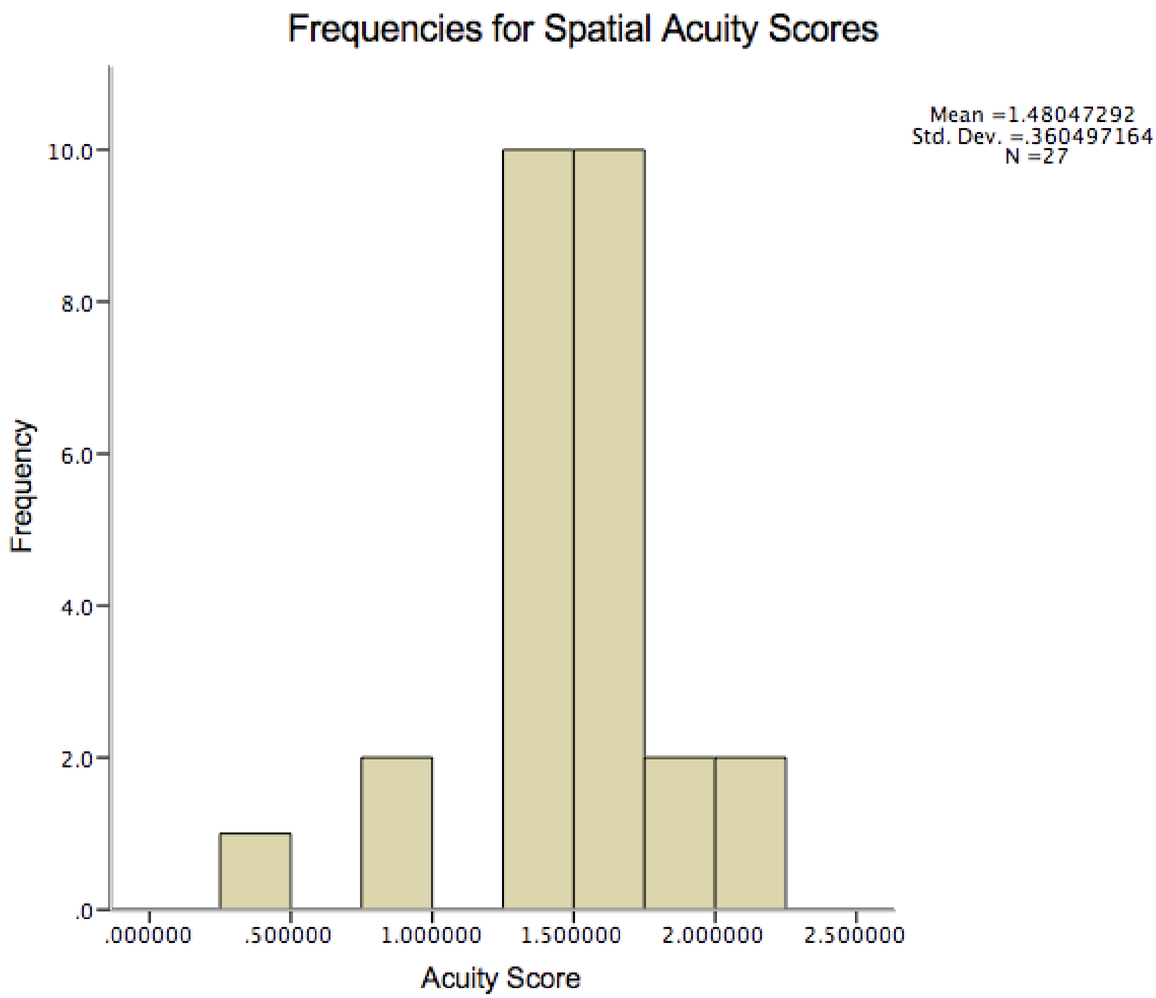

Figure 8. Frequencies of spatial acuity scores for all participants whose data could be modeled to a logarithmic function. The higher an individual acuity score, the lower the ability to make number discriminations.