

Running head: NON-VERBAL NUMBER ACUITY AND SYMBOLIC MATHEMATICS

Non-Verbal Number Acuity Correlates with Symbolic Mathematics Achievement:

But only in Children

Matthew Inglis

Loughborough University

Nina Attridge

Loughborough University

Sophie Batchelor

University of Nottingham

Camilla Gilmore

University of Nottingham

3981 words

Abstract

The process by which adults develop competence in symbolic mathematics tasks is poorly understood. Non-human animals, human infants, and human adults all form non-verbal representations of the approximate numerosity of arrays of dots, and are capable of using these representations to perform basic mathematical operations. Several researchers have speculated that individual differences in the acuity of such non-verbal number representations provide the basis for individual differences in symbolic mathematical competence. Specifically, prior research has found that 14-year-old children's ability to rapidly compare the numerosities of two sets of colored dots is correlated with their mathematics achievements at ages 5-11. Here we demonstrate that although when measured concurrently the same relationship holds in children, it does not hold in adults. We conclude that the association between non-verbal number acuity and mathematics achievement changes with age, and that non-verbal number representations do not hold the key to explaining the wide variety of mathematical performance levels in adults.

NON-VERBAL NUMBER ACUITY CORRELATES WITH SYMBOLIC MATHEMATICS ACHIEVEMENT: BUT ONLY IN CHILDREN

Dealing accurately with numerical quantities is fundamental to success in modern societies. On a daily basis we are asked to make judgments about concepts that are expressed numerically, be it monetary value, weather, temporal duration or spatial distance. But what is the cognitive basis for our ability to engage with numerical quantities such as these? Recently it has been proposed that the answer to this question is an innate and inexact analog system known as the *Approximate Number System (ANS)*. This evolutionarily ancient system enables us, for example, to rapidly decide, without explicitly counting, which of two orange trees has the greatest number of fruit, or which of two herds has the greatest number of gazelle. The ANS supports approximate numerical operations, such as comparison and addition, on both visual and auditory arrays, in adults, children and in non-human animals (Barth, La Mont, Lipton, Dehaene, Kanwisher & Spelke, 2006; Brannon & Terrace, 2000; Cordes, Gelman, Gallistel & Whalen, 2001; Dehaene 1992, 1997; Gallistel & Gelman 1992).

These findings raise the possibility that the ANS is the cognitive basis of everyday numeracy skills such as exact addition, subtraction and multiplication with Arabic numerals. At least four sources of evidence support this possibility. First, the ANS appears to be automatically activated in response to Arabic numerals by adults as well as children (Dehaene, 1997; Moyer & Landauer, 1967). Second, children who have had no formal mathematical instruction seem to harness the ANS when asked to perform approximate *symbolic* arithmetic operations (additions of numerosities represented as Arabic numerals, Gilmore, McCarthy & Spelke, 2007). Third, several different measures of children's ANS proficiency (using both symbolic and nonsymbolic stimuli) have been found to correlate with their performance on tests of early symbolic numeracy skills (e.g., symbolic stimuli: Durand, Hulme, Larkin, & Snowling, 2005; Holloway & Ansari, 2009; non-symbolic stimuli:

Halberda, Mazzocco & Feigenson, 2008; Mundy & Gilmore, 2009). Fourth, one measure of ANS proficiency taken at the start of formal schooling, the symbolic numerical distance effect, has been shown to predict young children's symbolic mathematics competence a year later (De Smedt, Verschaffel & Ghesquière, 2009). However, other measures often assumed to reflect ANS proficiency, notably the non-symbolic numerical distance effect, have been found to not correlate with school-level mathematics achievement (Holloway & Ansari, 2009; Mundy & Gilmore, 2009). It may be that the NDE is a poor measure of ANS proficiency (Gilmore, Attridge, & Inglis, in press).

A natural question arises from this set of findings: If adults exhibit an automatic ANS-based response to symbolic numerosities, and if the ANS is involved in early symbolic numerical operations in children, does the ANS also influence more sophisticated numerical operations of the types conducted by adults in everyday life? In other words, is the ANS a fundamental part of the way humans of all ages engage with numerical quantities? Or, alternatively, is the large variance in mathematical competence possessed by adults unrelated to differences in ANS acuity? In view of the educational potential for harnessing the ANS to develop effective instruction, it is clear that further evidence which speaks to the relationship between the ANS and formal mathematics achievement would be desirable. Our goal in this paper is to explore this relationship.

The Approximate Number System

Representations of numerosities within the ANS are noisy, and grow noisier as the magnitude of the to-be-represented numerosity increases. To capture this noise, Barth et al (2006) proposed that ANS-representations of a numerosity n follow a normal distribution with mean n and standard deviation wn . Here w is the internal Weber fraction, which gives a measure of the acuity of an individual's ANS. Thus, on a non-symbolic comparison task,

where participants are asked to select which of two arrays of colored dots are numerically greater, those participants with high w s have less precise representations, and consequently lower accuracy rates.

Halberda et al (2008) gave 14-year-old children a non-symbolic comparison task, calculated each individual's w (henceforth ANS acuity), and related these to standardized mathematics achievement tests which had been taken at ages 5-11. They found strong relationships between these two measures at each testing-point (r^2 s varied from 0.11 to 0.33). These correlations retained significance after controlling for covariates such as IQ and working memory measures. This result seems to suggest that ANS acuity and mathematics achievement are closely related. However, both ANS acuity (Halberda & Feigenson, 2008) and symbolic mathematics achievement are developmental so, because Halberda et al. did not test their participants concurrently on the two tasks, it is possible that their ANS measure (taken at age 14) had been influenced by developmental patterns not reflected in the mathematics achievement tests taken at ages 5-11. For example, it is conceivable that individual differences in early mathematics achievement leads, over several years school experience, to differential levels of exposure to numerical ideas (both in terms of quality and quantity of the exposure), which in turn might lead to differential ANS acuities. Thus measuring ANS acuity some years after mathematical achievement might overstate the relationship between these two constructs. Some support for this possibility comes from Iuculano, Tang, Hall & Butterworth's (2008) finding that the non-symbolic addition performance of 8-9 year old children did not correlate with their exact symbolic addition performance when tested concurrently.

To investigate these issues we conducted Experiment 1. Our aim was to determine whether there is a relationship between children's ANS acuities and formal mathematics achievements when the two constructs are measured concurrently.

Experiment 1

Participants

Participants were 39 children (20 male) aged 7.6-9.4 years ($M = 8.4$) who took part, with parental consent, at school and were rewarded with stickers. Three tasks were administered, as listed below.

Non-symbolic Comparison Task

Participants completed a computer-based non-symbolic comparison task in which they selected the more numerous of two dot arrays, designed based on the version used by Pica, Lemer, Izard, & Dehaene (2004). The two arrays (one red and one blue on a white background) were presented side by side simultaneously on a 15" LCD laptop screen. The ratios between the numerosities of the left and right arrays were 0.5, 0.6, 0.7, 0.8 and their inverses, and the numerosity of the arrays ranged from 5 to 22. The color and side of screen of the correct array were fully counterbalanced. Participants were asked to select, as quickly and accurately as possible, which array was more numerous. Responses were recorded via the leftmost (left bigger) and rightmost (right bigger) buttons on a five button response box.

Each of 128 trials began with a fixation point for 1000ms, followed by the dot arrays for 1500ms. If the participant had not responded within 1500ms, the arrays were followed by a white screen with a black question mark. This allowed participants to still respond whilst preventing them from counting the arrays. Participants rarely exceeded this duration (2.5% of trials) and the mean RT was well within this limit (776ms). The design is summarized in Figure 1. Experimental trials were preceded by a practice block of 8 trials.

To prevent participants reliably using strategies based on continuous quantities correlated with number (dot size, total enclosure area), the stimuli were created following the method adopted by Pica et al. (2004). For each problem two sets of stimuli were created: one

in which the dot size and total enclosure area decreased with numerosity, and one in which the dot size and total enclosure area increased with numerosity.

Woodcock-Johnson III Tests of Achievement

The Calculation subtest of the Woodcock Johnson III Tests of Achievement was administered with the standard procedure (participants had no time limits, and continued until they had answered six questions incorrectly in succession). An example problem is given in Table 1.

Weschler Abbreviated Scale of Intelligence

Participants completed the Matrix Reasoning subtest of the Weschler Abbreviated Scale of Intelligence (WASI), following the standard procedure. The raw scores were converted into T-scores to give an age-standardized measure of non-verbal intelligence.

Results

Those participants who appeared to be using strategies based on continuous quantities correlated with number (i.e. those who were not using their ANS) on a majority of trials were eliminated from the sample (i.e. those participants whose accuracy rates on the two sets of stimuli created by Pica et al.'s (2004) method differed by more than 0.5, $N = 10$). In addition we removed those participants whose performance was not above chance ($N = 5$). This left 24 participants for the main analysis.

Accuracy rates varied from 0.59 to 0.81, with a mean of 0.69 ($SD = 0.06$), and were subjected to an Analysis of Variance (ANOVA) with ratio as a within-subjects factor. There was a significant effect of ratio, $F(3,69) = 11.76$, $p < .001$, and a significant linear trend, $F(1,23) = 50.32$, $p < .001$. As is characteristic of the ANS, accuracy rates were highest at the 0.5 ratio, and lowest at the 0.8 ratio. These data are shown in Figure 2.

Using the log-likelihood method, each participant's accuracy data were individually fitted to the model proposed by Barth et al. (2006)¹. Values of the w parameter varied from

0.34 to 1.17, with a mean of 0.65 (SD = 0.23).² The relationship between participants' non-symbolic comparison accuracy and their Woodcock Johnson Calculation subtest scores is shown in Figure 3. ANS acuity, as measured by w parameters, was found to negatively correlate with Woodcock Johnson Calculation subtest scores, after controlling for age-standardized WASI matrix reasoning scores and age, $pr = -.548$, $p = .008$. In other words, high ANS acuities (w parameters close to zero) were related to high scores on the Woodcock-Johnson Calculation subtest.

Discussion

Our goal in Experiment 1 was to determine whether the relationship found by Halberda et al. (2008) between ANS acuity at age 14 and mathematics achievement at ages 5-11, could be replicated when the two measures are taken concurrently. We measured the ANS acuity and mathematics achievement of 7-9 year old children and found a strong relationship between the two constructs. Those children with high ANS acuity tended to have high mathematics achievement scores. As both ANS acuity and mathematics ability are developmental, it is natural to ask whether these two constructs co-develop into adulthood. In other words, do children develop their mathematics ability and ANS acuity together in a mutually reinforcing cycle? To explore this issue we conducted a second experiment with adult participants.

Experiment 2

The primary goal of Experiment 2 was to determine whether the relationship between ANS acuity and mathematics achievement we found with 7-9 year old children in Experiment 1 also holds with adults experienced in everyday mathematics. Along with a measure of achievement focused on numerical calculation (used by Halberda et al. (2008) and in Experiment 1), we took several other measures of achievement that, taken together, better

reflect the broad nature of mathematics. In particular along with various measures of numerical skill (calculation, calculation fluency etc), we also took measures of non-numerical mathematical skills related to logical inference and geometry.

Participants

Participants were 101 adults (50 male, aged 18-48, $M = 23$) recruited from the participant panel of the University of Nottingham's Learning Sciences Research Institute; each was paid £20 for taking part. Testing was conducted individually. Computer tasks were presented on a 17" Philips 170B LCD. As well as the tasks reported in this paper, participants tackled a number of other numerical cognition tasks (non-symbolic addition, subitizing etc.) not discussed here. The order of tasks was counterbalanced between participants, except that all tasks with symbolic stimuli were presented after tasks with non-symbolic stimuli, so as to avoid cuing counting strategies (cf. Gilmore, et al., in press).

Non-symbolic Comparison Task

Participants completed the non-symbolic comparison task from Experiment 1, with minor changes to the stimuli characteristics and procedure (pilot testing revealed that using identical stimuli to those given to the children may have led to ceiling effects). The numerical size of the arrays ranged from 9-70, and the pairs differed by the ratios 0.625, 0.714 and 0.833 (5:8, 5:7 and 5:6) and their inverses. Again, the stimuli were created following the method adopted by Pica et al. (2004).

Each of 120 trials began with a fixation point for 1000ms, followed by the dot arrays until response. Participants were asked to select, as quickly as possible, which array was more numerous. Responses were recorded via the leftmost (left bigger) and rightmost (right bigger) buttons on a five-button response box.

There was a response time limit of 1249ms to prevent ceiling effects. The limit was determined by taking the mean plus one standard deviation of the reaction times found in

pilot testing. This was enforced with a “Please speed up” message, followed by the next trial. Few trials led to the display of this message (approximately 3.6%), and these were recorded as incorrect responses. The experimental trials were preceded by a block of 10 practice trials. Participants were prompted to take breaks after every 20 trials.

Woodcock-Johnson III Tests of Achievement

Again, the calculation subtest of the Woodcock Johnson III Tests of Achievement was administered with the standard procedure. However unlike in Experiment 1, we also administered the Math Fluency, Applied Problems, Quantitative Concepts and Number Series subtests of the Woodcock Johnson, again with the standard procedure. Descriptions of each of these subtests are given in Table 1.

Non-numerical tasks

Participants answered 32 paper-based conditional inference problems following the design used by Evans and Handley (1999). In addition, they were given 20 minutes to complete a reduced version of the paper-based van Hiele Level Geometry Test (Usiskin, 1982), and also completed the matrix reasoning subtest of the Weschler Abbreviated Scale of Intelligence (WASI), following the standard procedure.

Results

As in Experiment 1, those participants who appeared to be using non-numeric cues for the majority of trials ($N = 25$), or whose performance was not above chance ($N = 11$), were eliminated from the analysis. This left 64 participants (an additional one participant did not complete all the tasks in the study).

Accuracy rates varied from 0.58 to 0.85, with a mean of 0.73 ($SD = 0.06$), and were subjected to a one-way ANOVA with ratio as a within-subjects factor. Again, responses

showed the ratio effect characteristic of the ANS, $F(2,126) = 107.6, p < .001$, and a significant linear trend, $F(1,63) = 219.0, p < .001$. These data are shown in Figure 4.

Values of the w parameter varied from 0.22 to 0.95, with a mean of 0.39 (SD = 0.14). The relationship between participants' non-symbolic comparison accuracy and their Woodcock Johnson Calculation subtest scores is shown in Figure 5. Unlike in Experiment 1, ANS acuity, as measured by w parameters, was not found to correlate with Woodcock-Johnson calculation subtest scores, after controlling for age-standardized matrix reasoning scores and age, $pr = +.161, p = .211$. This non-significant positive correlation was found to be significantly different to the significant negative correlation found in Experiment 1, Fisher's r -to- z transformation, $z = 3.07, p = .001$.³ In other words, as well as being not significantly different to zero, the correlation between ANS acuity and mathematical achievement found in adults was significantly different to that found in children.

In addition, having controlled for age-standardized matrix reasoning scores, no significant correlation was found between ANS acuity and either the Math Fluency, $pr = +.184, p = .156$, Applied Problems, $pr = -.098, p = .453$, Quantitative Concepts, $pr = -.110, p = .401$, or Number Series, $pr = +.081, p = .535$, subtests of the Woodcock Johnson III Tests of Achievement. Nor were there significant correlations between ANS acuity and overall scores on the Conditional Inference task, $pr = -.110, p = .400$, or the van Hiele Levels Geometry Test, $pr = -.164, p = .208$.

Summary

In Experiment 2 we found no significant relationship between adults' ANS acuity and any measure of mathematical achievement. We asked participants to answer a wide variety of mathematical tasks, including calculation, speeded calculation, conditional reasoning, and applied problems, and found no relationships between these scores and participants' ANS

acuties. In particular, the correlation between ANS acuity and calculation achievement for adults was significantly different to that for children.

General Discussion

Halberda et al. (2008) found a relationship between individuals' ANS acuity, tested at age 14, and their mathematics achievement at ages 5-11. They speculated that this may be because the ANS plays a causal role in individual differences in symbolic mathematical competency. Since adults also have access to the ANS, which appears to be automatically activated when participants view Arabic numerals (Moyer & Landauer, 1967), it is natural to hypothesise that a similar relationship holds in adults. Here we confirmed that in 7-9 year old children there is a strong relationship between ANS acuity and numerical calculation achievement (when tested concurrently), but demonstrated that the same relationship does not hold with adults. This finding rules out the possibility that ANS acuity is directly implicated in the large individual differences found in adults' numerical calculation achievement. Together these findings suggest that, along with ANS acuity and mathematical achievement changing with age, the strength of the association between these constructs does as well.

One speculative hypothesis that would account for this set of data is to suppose that the ANS plays a bootstrapping role in the development of whole number understanding. For example, children may come to understand whole numbers by assigning verbal and symbolic names to visual and auditory stimuli that give rise to similar ANS representations. Thus for young children we might expect that their fluency with symbolic numbers would be associated with their ANS acuity as their symbolic numbers would be nothing more than tags for ANS representations. However, once children had reached a certain sophistication with numerical concepts, other factors (working memory capacity, strategy choice, teaching effectiveness, etc) may come to dominate individual differences in mathematical

performance, leading to a decline in the relationship with ANS acuity. While speculative, this hypothesis does suggest that a detailed microgenetic study of how ANS acuity, and the relationship between ANS acuity and mathematics achievement, develops through formal schooling would be a valuable contribution to our understanding of how numerical concepts are formed.

There is now growing consensus that our abilities to deal with complex symbolic numerical concepts on a day-to-day basis is related in some way to the Approximate Number System, an innate analog system which supports rapid approximate numerical judgments. However, the exact nature of this relationship remains unclear. The finding that the correlation between ANS acuity and mathematics achievement that exists in childhood is not present in adulthood indicates that there is no simplistic relationship between the ANS and symbolic mathematics achievement. Studying the pattern of decline in the relationship between ANS acuity and mathematics achievement as participants gain in maturity and mathematical experience may ultimately shed further light on the cognitive basis of the wide range of numerical operations that we each perform during everyday life.

References

- Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. (2006). Nonsymbolic arithmetic in adults and young children. *Cognition*, 98, 199-222.
- Brannon, E.M. & Terrace, H.S. (2000). Representation of the numerosities 1-9 by rhesus macaques (*macca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*, 26, 31-49.
- Cordes, S., Gelman, R., Gallistel, C.R., & Whalen, J. (2001). Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychonomic Bulletin and Review*, 8, 698-707.
- Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, 44, 1-42.
- Dehaene, S. (1997). *The number sense*. Oxford, UK: Oxford University Press.
- De Smedt, B., Verschaffel, L., & Ghesquière, P. (2009). The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *Journal of Experimental Child Psychology*, 103, 469–479.
- Durand, M., Hulme, C., Larkin, R., & Snowling, M. (2005). The cognitive foundations of reading and arithmetic skills in 7- to 10-year-olds. *Journal of Experimental Child Psychology*, 91, 113–136.
- Evans, J.St.B.T. & Handley, S.J. (1999). The role of negation in conditional inference. *Quarterly Journal of Experimental Psychology*, 52A, 739-769.
- Gallistel, C.R. & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, 44, 43-74.
- Gilmore, C.K., McCarthy, S., & Spelke, E.S. (2007). Symbolic arithmetic knowledge without instruction. *Nature*, 447, 589-591.
- Gilmore, C.K., Attridge, N., & Inglis, M. (in press). Measuring the approximate number system. *Quarterly Journal of Experimental Psychology*.

- Green, D.M. & Swets, J.A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Halberda, J., Mazocco, M.M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, 455, 665-668.
- Halberda, J. & Feigenson, L. (2008). Developmental change in the acuity of the "number sense": The Approximate Number System in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, 44, 1457-1465.
- Holloway, I.D. & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology*, 103, 17-29.
- Iuculano, T., Tang, J., Hall, C.W.B., & Butterworth, B. (2008). Core information processing deficits in developmental dyscalculia and low numeracy. *Developmental Science*, 11, 669-680.
- Izard, V. & Dehaene, S. (2008). Calibrating the mental number line. *Cognition*, 106, 1221-1247.
- Moyer, R.S. & Landauer, T.K. (1967). Time required for judgements of numerical inequality. *Nature*, 215, 1519-1520.
- Mundy, E. & Gilmore, C.K. (2009). Children's mapping between symbolic and nonsymbolic representations of number. *Journal of Experimental Child Psychology*, 103, 490-502.
- Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, 306, 499-503.
- Usiskin, Z. (1982). Van Hiele levels and achievement in secondary school geometry (Final report of the Cognitive Development and Achievement in Secondary School Geometry Project). Chicago: University of Chicago, Department of Education. (ERIC Document Reproduction Service No. ED 220 288).

Author Note

This research was supported by a grant from the Economic and Social Research Council (C.G. & M.I.), a British Academy Postdoctoral Fellowship (C.G.), and a Royal Society Worshipful Company of Actuaries Educational Research Fellowship (M.I.).

Notes

1. Following Green and Swets's (1966) discussion of Weber's Law, Barth et al. (2006) proposed that the accuracy of a participant when comparing n_1 with n_2 is given by

$$\text{acc}(n_1, n_2; w) = \frac{1}{2} + \frac{1}{2} \text{erf} \left(\frac{|n_1 - n_2|}{\sqrt{2}w\sqrt{n_1^2 + n_2^2}} \right), \text{ where erf is the Gaussian error function,}$$

and w is the participants' ANS acuity.

2. These w parameters are somewhat higher than those that have been previously reported in the literature for similar ages (e.g. Halberda & Feigenson, 2008). One possible reason for this discrepancy is the shorter durations (participants' mean RT was 776ms) and lack of feedback given in the current study compared to earlier studies (Halberda and Feigenson, for example, gave feedback and had trials which lasted 1200ms with their 4-, 5- and 6-year old participants). We are not aware of any research that has directly investigated the relationship between stimuli display time or feedback and ANS acuity.
3. We also ran a regression including both participants' accuracy and their mean RT as predictors of mathematics achievement. The model did not reach significance, $F(2,61) = 0.596, p = .554$.

Table 1. Details about the various subtests of the Woodcock Johnson III Tests of Achievement used in Experiments 1 and 2.

Name of Subtest	Description	Example Problem
Calculation	Arithmetic computation with paper and pencil	12% of 6.0 =
Math Fluency	Number of simple calculations performed in three minutes (up to 160)	$7 \times 9 =$
Applied Problems	Oral word problems solved with paper and pencil	If 60 feet of wire weighs 80 pounds and you had 150 feet of the wire, how many pounds of wire would you have?
Quantitative Concepts	Oral questions about factual mathematical information	What does this symbol mean? π
Number Series	Determining a numerical sequence	___ 13 20 26

Figure 1. The experimental paradigm used in Experiment 1. The two dot arrays were colored red and blue.

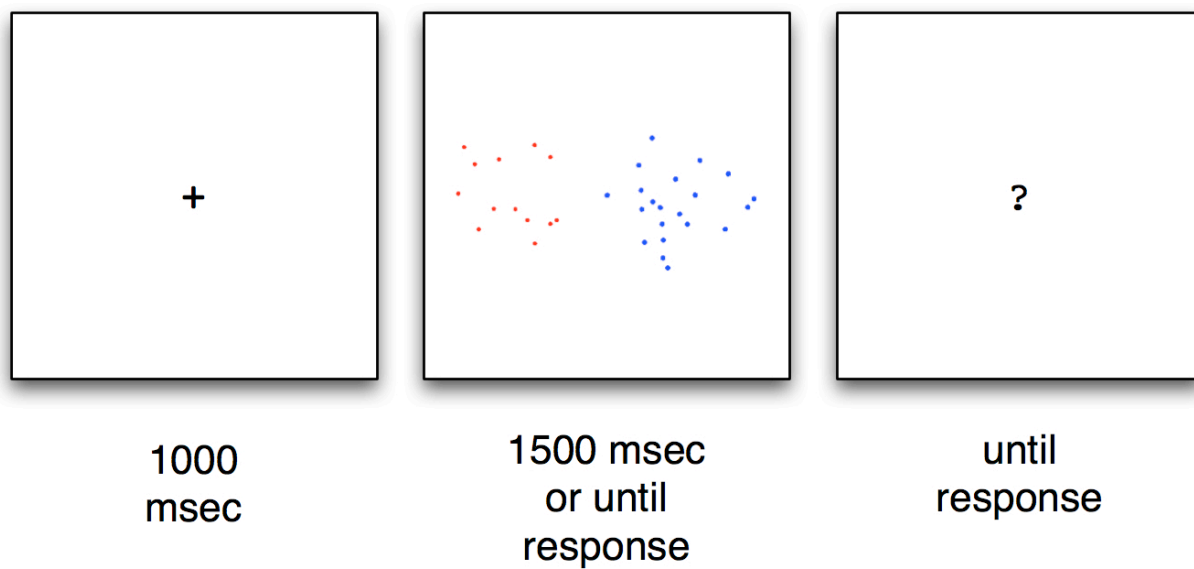


Figure 2. Accuracy rates by problem ratio in Experiment 1. Error bars show ± 1 SE of the mean.

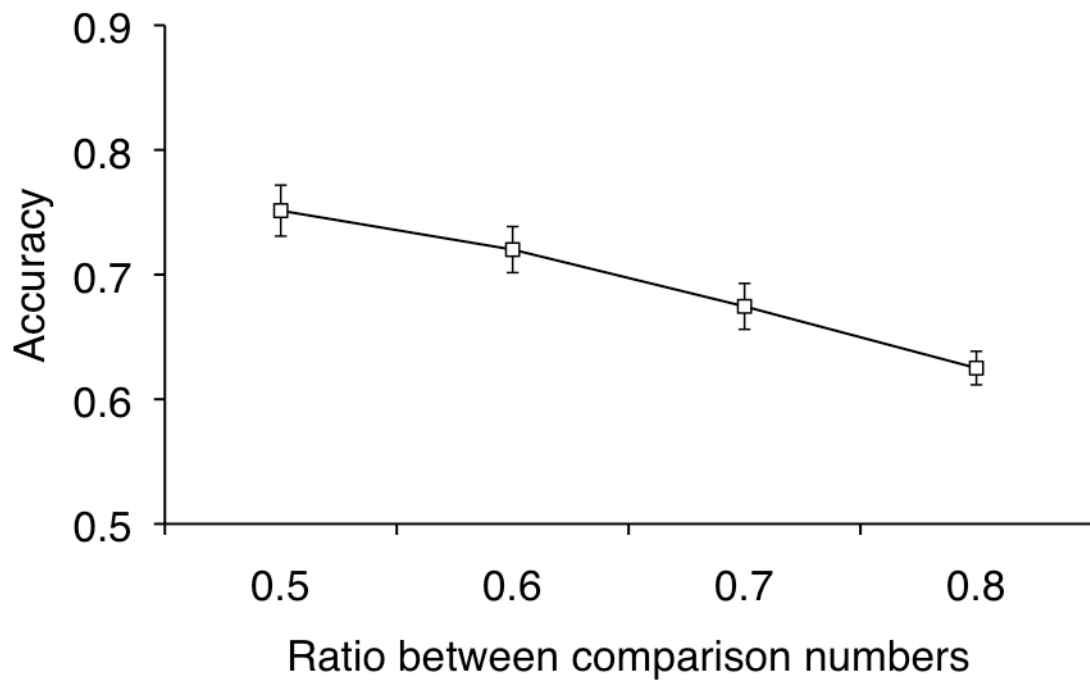


Figure 3. The relationship found in Experiment 1 between standardized residuals (controlling for age-standardized matrix reasoning scores) for ANS Acuity (non-symbolic comparison internal Weber fraction) and the Woodcock Johnson calculation subtest.

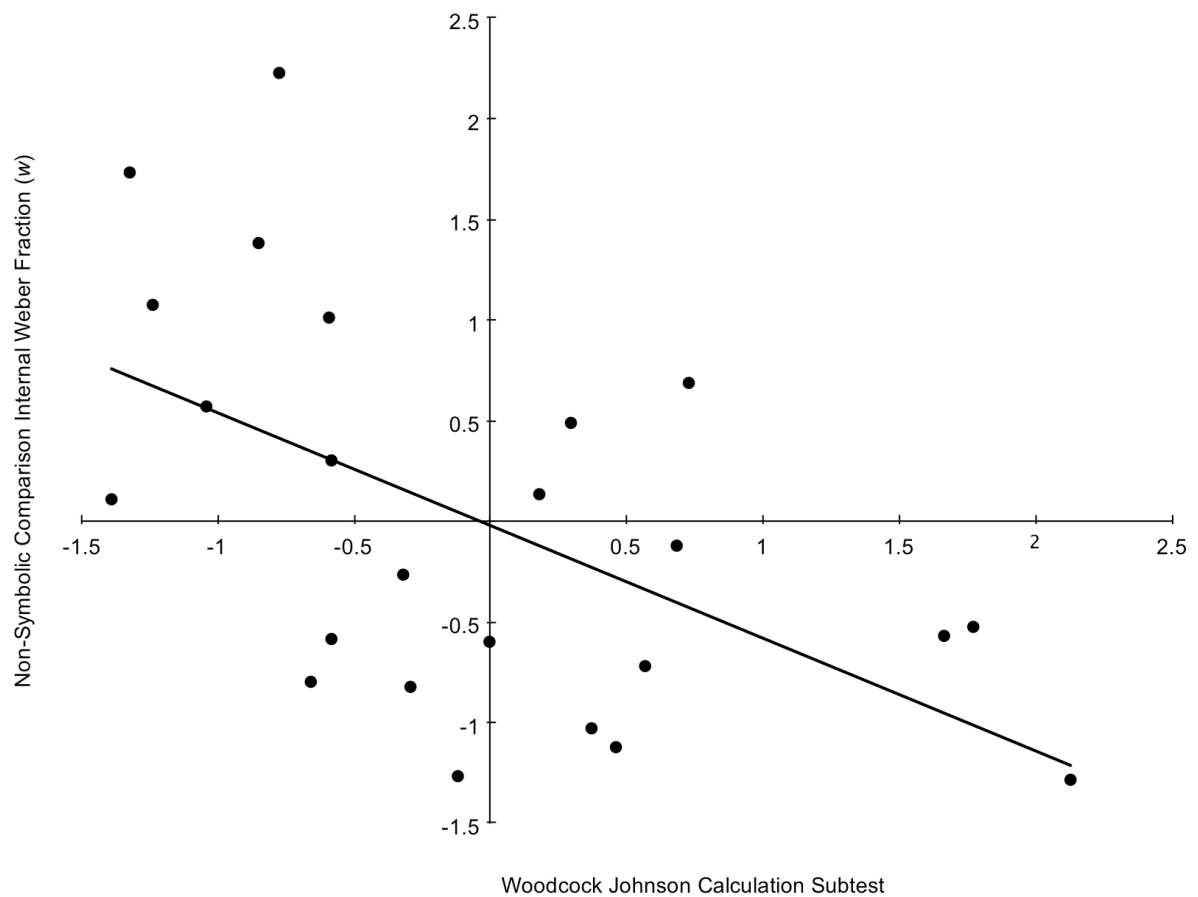


Figure 4. Accuracy rates by problem ratio in Experiment 2. Error bars show ± 1 SE of the mean.

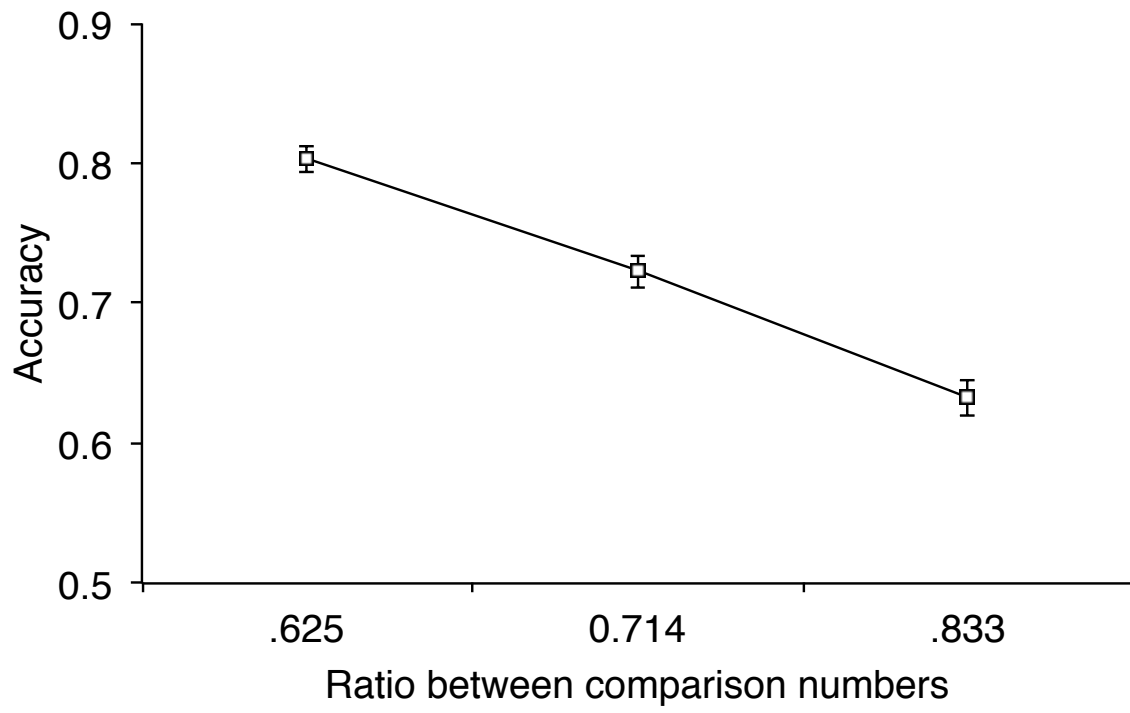


Figure 5. The relationship found in Experiment 2 between standardized residuals (controlling for age-standardized matrix reasoning scores) for ANS Acuity (non-symbolic comparison internal Weber fraction) and the Woodcock Johnson calculation subtest.

