UC Irvine
UC Irvine Previously Published Works

Title

Is there really a link between exact-number knowledge and approximate number system acuity in young children?

Permalink

Journal

ISSN

Authors

Negen, James
Sarnecka, Barbara W

Publication Date

DOI

Copyright Information

Peer reviewed

Is there really a link between exact-number knowledge and approximate number system acuity in young children?

James Negen* and Barbara W. Sarnecka

University of California, Irvine, California, USA

Although everyone perceives approximate numerosities, some people make more accurate estimates than others. The accuracy of this estimation is called *approximate number system* (ANS) *acuity*. Recently, several studies have reported that individual differences in young children's ANS acuity are correlated with their knowledge of exact numbers such as the word 'six' (Mussolin *et al.*, 2012, *Trends Neurosci. Educ.*, 1, 21; Shusterman *et al.*, 2011, *Connecting early number word knowledge and approximate number system acuity*; Wagner & Johnson, 2011, *Cognition*, 119, 10; see also Abreu-Mendoza *et al.*, 2013, *Front. Psychol.*, 4, 1). This study argues that this correlation should not be trusted. It seems to be an artefact of the procedure used to assess ANS acuity in children. The correlation arises because (1) some experimental designs inadvertently allow children to answer correctly based on the size (rather than the number) of dots in the display and/or (2) young children with little exact-number knowledge may not understand the phrase 'more dots' to mean *numerically* more. When the task is modified to make sure that children respond on the basis of numerosity, the correlation between ANS acuity and exact-number knowledge in normally developing children disappears.

What does it mean to have a talent for math? Is it possible that seeing (without counting) that one set has more items than another is actually a major component of math ability? Recently, much has been written about the approximate number system (ANS) and its relation to math achievement. At times, the reasoning seems to verge on essentialism (Gelman, 2003), with the ANS imagined as an inner force that causes people to succeed or fail at mathematics.

In this paper, we consider (and ultimately reject) one claim in this literature: The claim that young children's ANS acuity is related to their knowledge of exact-number words (one of the earliest measures of symbolic math learning). We argue that the reported link between ANS acuity and exact-number learning is an artefact of flawed measurement methods.

This argument is supported by two experiments: In the first, children's ANS acuity was measured using a standard task.[1] In the second, the standard task was modified to make sure that children responded on the basis of numerosity. The first experiment replicates the finding from the literature (i.e., that performance on the standard ANS task is linked to

exact-number knowledge); the second experiment shows that when children's ANS acuity is measured more carefully, with age-appropriate controls, the evidence for a link between ANS acuity and exact-number knowledge disappears.

Approximate and exact numbers

Human beings have the ANS that allows us to quickly estimate of the number of items in a set (see, Dehaene, 1997; Feigenson, Dehaene, & Spelke, 2004; Gallistel, 2011, for review). For example, if you are at a wine tasting with 30 people and there are only 20 glasses of wine on the table, the ANS is what allows you to see (without counting) that there are not enough glasses for everyone. In the ANS, the difficulty of telling any two numerosities apart depends on their ratio. The previous example's 2:3 ratio (20 wine glasses/30 people) is fairly easy to discriminate. If there were 120 glasses and 130 people (a ratio of 12:13), it would be much harder to tell that there were not enough glasses. It would be harder even though the absolute difference (i.e., 10 people without glasses) would be the same.

Numeric discrimination becomes more difficult as the ratio of the sets approaches 1 (e.g., 120:130 < 20:30), but there are still individual differences in performance. Some people succeed more often than others at more difficult ratios. The accuracy of a person's discrimination is called their ANS acuity, and individual differences in ANS acuity are measurable even in childhood (e.g., Halberda & Feigenson, 2008).

In addition to the ANS, people in numerate societies have an exact-number system. People start learning this system in early childhood: Children first learn a list of counting words, then a counting procedure, and then the cardinal meanings of the number words. The first few number-word meanings are learned one at a time (first *one*, then *two*, then *three*, sometimes *four*). The higher number-word meanings seem to be learned by a process of inductive reasoning, when children figure out the *cardinality* principle (Gelman & Gallistel, 1978; Schaeffer *et al.* 1974). This principle states that the last word used in counting tells the cardinality of the whole set. It is this principle that connects counting to exact, cardinal numbers: If you understand cardinality, you know the meaning of any number word in any language by its place in the counting list. That is, you know that the eighth word in any counting list means 8, the 15th word means 15, the 294th word means 294, and so on. Children who implicitly understand this key idea are called cardinality-principle knowers, or *CP-knowers* for short.

Becoming a CP-knower takes a long time, on the order of a year or more. Children learn the meanings of the first three or four number words one at a time, with weeks or months elapsing before they learn the meaning of the next word. The child's progress on this front is called their *number-knower level* or just *knower-level*, and it is commonly assessed using the Give-N task. In this task, the child is asked to give some number of items to a stuffed animal (e.g., 'Give the anteater *five* bananas'.). A *pre-number knower* is a child who does not succeed reliably at any number. A *one-knower* gives one item when asked for 'one', and more than one when asked for any other number, but does not reliably distinguish among numbers higher than one. A *two-knower* succeeds at 'one' and 'two' but no higher numbers, and the pattern continues through the *three-knower* and *four-knower* levels until children figure out the cardinality principle, becoming CP-knowers. Children at the one-knower through four-knower levels are collectively known as *subset-knowers* (Le Corre & Carey, 2007) because they know the meanings of only a subset of the words in their counting list.

How is ANS acuity related to exact-number knowledge?

Because learning the exact-number system takes a long time, and children learn it at different speeds, any large group of preschoolers includes some one-knowers, some two-knowers, some three-knowers, and so on. Individual differences are also found in children's ANS acuity. Naturally, people have wondered whether the two measures are correlated: Do the children who know more number words also perform better on ANS tasks?

At least four recent studies have reported that they do (Mussolin, Nys, Leybaert, & Content, 2012; Shusterman, Slusser, Halberda, & Odic, 2011; Wagner & Johnson, 2011; see also Abreu-Mendoza, Soto-Alba, & Arias-Trejo, 2013), but we think this correlation may be suspect. The tasks intended to measure young children's ANS acuity might not actually do so.

One problem arises in how area is controlled. Some studies have used stimuli in which the total summed area of every set is kept constant. This means that the side with *more dots* always has *smaller dots* (see Figure 1). Thus, children could always
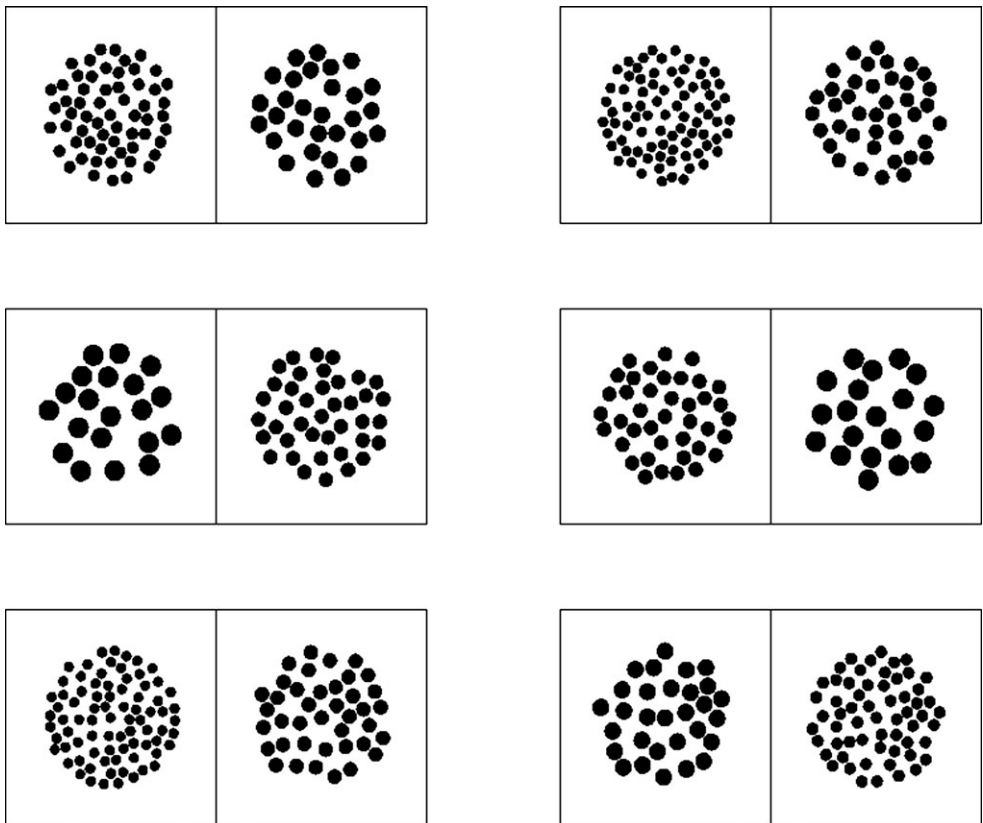

Figure 1. An example of stimuli that are not fully controlled for area. In the displays above, the two dot clouds in each pair have the same area, meaning that the cloud with a greater number of dots always has dots that are smaller in diameter. Thus, participants can always choose the correct answer just by pointing to the side where dots are smaller. To make sure that participants' responses are based on the numerosity of the dots (and not their size), the side with more dots should have larger area and larger dots 50% of the time.

succeed on this task (regardless of their ANS acuity) just by choosing the side with smaller individual dots.

Another, trickier problem is that subset-knowers may not yet have a clear concept of what an exact, cardinal number is and thus may not understand the question 'Which side has more dots?' (or similar phrasings) as a question about *numerosity*. They may misinterpret the question as being about some other dimension (e.g., the size of individual dots or the total area of the display), or they might simply guess at random. On the other hand, CP-knowers (who clearly do have a concept of what an exact, cardinal number is) are likely to understand the question the way the experimenters intended (i.e., as a question about the *number* of dots).

If subset-knowers do not understand the task, they perform at chance. If CP-knowers do understand the task, they perform better on easier ratios and worse on more difficult ratios (just like adults). In this case, a correlation arises between exact-number knowledge and performance on the ANS task. But this correlation has nothing to do with the children's ANS acuity, because the subset-knowers' ANS acuity has not been measured. To address this problem, a study needs to show that individual children are performing above chance, at least at the very easy ratios, and that subset-knowers are performing above chance as a group.

There is some evidence to support the idea that subset-knowers have trouble identifying numerosity as a perceptual dimension. Slusser and Sarnecka (2011) tested 116 children, aged 30–48 months, on a Match-to-Sample task. In this task, the experimenter showed the child a sample picture of, for example, eight happy, green turtles and said, 'This picture has *eight* turtles'. Then, the experimenter placed two more pictures on the table and said, 'Find another picture with *eight* turtles'. One picture had the same number of turtles as the sample, but the turtles were bigger or smaller than those in the sample picture, so that the total area did not match. The other picture matched the sample in total summed area of the turtles, but not in number.

Only the CP-knowers correctly chose pictures of the same number (rather than the same area) to match the sample. Subset-knowers failed to do this, although they were very good at matching pictures on colour ('Find another picture with *green* turtles') or mood ('Find another picture with *happy* turtles'). In other words, only the CP-knowers recognized *numerosity* as the dimension on which one picture of 'eight turtles' should match every other picture of 'eight turtles'. Based on this result, we might expect that only CP-knowers will correctly interpret the phrase 'more dots' in the ANS task instructions to mean the *more numerous set* of dots. We might expect that subset-knowers will simply guess or that they might misinterpret the instructions to be about summed area, average area of each dot, density, and so on.

In a related finding, Odic, Paul, Hunter, Lidz, and Halberda (2013) showed that children understand and verify sentences about which side of an array has 'more dots' (numeric) versus 'more goo' (surface area) at around 3.3 years old on average. That study did not assess children's exact-number knowledge, but 3.3 years is close to the average age when children become CP-knowers in higher-SES populations. It is plausible that the success by children over 3.3 years actually reflects success by CP-knowers and failure by subset-knowers (who of course tend to be younger, on average).

Potential issues with previous research

Although several studies have reported a correlation between exact-number knowledge and ANS performance in young children, none have fully solved the two problems

outlined above. One of the first studies to report this correlation (Wagner & Johnson, 2011) had both the problems described above. First, area was equated in the stimuli, so that a child could answer correctly by always picking the set with smaller individual items. Second, the study reported that children as a group performed above chance in the ANS task, but no such data were reported for individual children or for subset-knowers as a group. In fact, the data from that study show children at lower knower-levels performing under 50% correct, consistent with the possibility that those children did not understand the task. (For example, if the children had always chosen the display with *larger* individual dots, their performance would be significantly below chance.)

A similar problem with chance performance can be seen in a study published the following year (Mussolin *et al.*, 2012). In that study, the data can be seen as two clusters. One group, children who scored low on the symbolic battery (testing knowledge of spoken number words and written Arabic numerals), scored around 50% correct on the ANS task. Several of these individual children scored below 50% on the task. A second group, children who scored higher on the symbolic battery, scored above chance on the ANS task, with some individual children scoring over 90%. These findings are consistent with the possibility that children with less knowledge of exact numbers did not base their responses on numerosity in the ANS task. (In other words, their ANS acuity was not actually measured.)

Preliminary data presented by Shusterman *et al.* (2011) fit the same pattern. Here, there was a large jump in children's ANS performance when the children changed from subset-knowers to CP-knowers. There was no evidence of understanding of the task (i.e., above-chance performance on easy ratios) either for individual subset-knowers, or for subset-knowers as a group.

Abreu-Mendoza *et al.* (2013) found a relation, but only in one condition where area was a potential extra cue. They had three conditions where they controlled three different non-numeric aspects of the stimuli: Density, total filled area, and correlated/anticorrelated area (the set with greater number is larger on exactly half of trials). Children's understanding of cardinality was correlated with their performance only in the 'total filled area' condition, where children could succeed just by picking the side with smaller items (see Figure 1). Consistent with the present argument, Abreu-Mendoza *et al.* found no relation between children's exact-number knowledge and their performance in the other two conditions. (Interestingly, they did not appear to have an issue with subset-knowers failing to understand the instructions, which were given in Spanish. It is possible that this indicates the construction '¿*Dónde hay más*?' is somehow more transparent at an earlier age than similar English constructions.)

The present study

This study re-examines the question of whether preschoolers' exact-number knowledge is related to their ANS acuity, with particular attention to the problem of how ANS acuity can be measured in such young children. In Experiment 1, we show how an illusory correlation between ANS acuity and exact-number knowledge arises when steps are not taken to ensure that children attend to numerosity. We test children on the Give-N task and the standard ANS task and replicate the correlation reported in the literature. In Experiment 2, we show how the ANS task can be modified to ensure that children are in fact basing their responses on numerosity. When children are tested using this modified task, we no longer find a correlation between their exact-number knowledge and ANS performance.

EXPERIMENT 1

Method

Participants

Participants included 46 monolingual, English-speaking preschoolers (25 male) with a mean age of 4 years, 3 months (range 2:8–5:6). Children were recruited at private preschools in the USA. Participants were not asked to declare their racial/ethnic background, income, or education levels. However, children were presumably representative of the community from which they were recruited. This community is majority white (45.1%) and Asian (39.0%), the median household income is $85,615 (38% higher than the median income for the state), and 66% of adult residents have at least a college degree. Families received a small prize (e.g., a rubber duck) when they signed up for the study; no prizes were given at the time of testing.

Give-N task

This was our measure of exact-number knowledge. Specifically, its purpose was to determine what number-word meanings each child knew (i.e., the child's number-knower level). The experimenter began the game by bringing out a stuffed animal (e.g., a lion), a plate, and a bowl of 15 small identical rubber toys (e.g., toy bananas). The experimenter said to the child, 'In this game, you're going to give something to the lion, like this [experimenter pantomimes putting an item on the plate and sliding it over to the lion]. I'm going to tell you what to give him'. Instructions were of the form, 'Can you give the lion *two* bananas?' After the child slid the plate over, a follow-up was asked in the form, 'Is that *two*?' If the child said yes, the answer was accepted. If not, the items were put back into the bowl and the trial was restarted.

All children were first asked for one item, then three items. Further requests depended on the child's earlier responses. When a child responded correctly to a request for N items, the next request was for $N + 1$. When the child responded incorrectly, the next request was for $N-1$. The requests continued until the child had at least two successes at a given N (unless the child had no successes, in which case she was classified as a pre-number knower) and at least two failures at $N + 1$ (unless the child had no failures, in which case she was classified as a CP-knower). The highest number requested was six.

A child was credited with knowing the meaning of a given number word if she had at least twice as many successes as failures for that number word. Failures included either giving the wrong number of items for a particular word N, or giving N items when any other number was requested. Each child's knower-level corresponds to the highest number she reliably generated. (For example, children who succeeded at one and two, but failed at three, were called two-knowers.) Children who had at least twice as many successes as failures for trials of five and six were called CP-knowers.

Standard ANS task

Children were shown two sets of dots (Figure 2) and were asked, 'Which side has more dots?' No feedback was given. The first four trials were considered warm-up trials and those data were not counted. There were 10 different numeric ratios tested: 1:2, 2:3, 3:4, 4:5, 5:6, and 6:7. There were 12–50 dots on each side. Area (both average area of individual dots and summed area of the total array) in half the trials was correlated with number; in
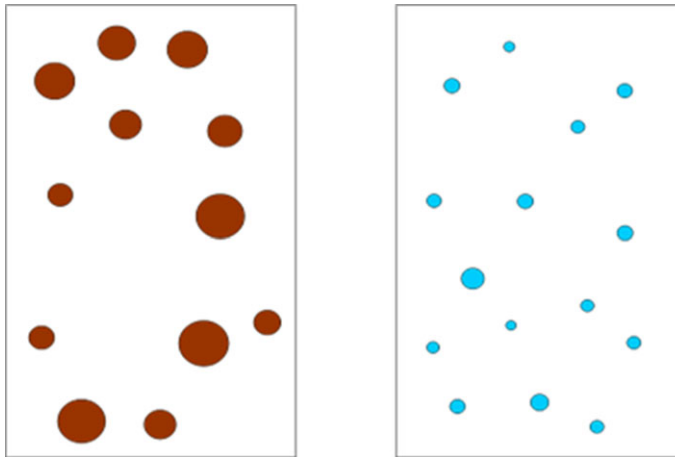
Figure 2. An example of stimuli in the standard approximate number system (ANS) task for Experiment 1.

the other trials, area and number were anticorrelated. Figure 2 is an example of an anticorrelated trial, where the numerically greater array has smaller dots and less summed area.

Within each side of each trial, the size of dots varied randomly. The dots were spread out to cover most of the rectangle. All the dots in a set were of the same colour; the two sets appearing in each trial were of two different colours.

Trials were organized into 10 blocks of six trials each, presented in a preset pseudo-random order. Each block contained one trial of each of the six ratios, three of them congruent with area and the other three incongruent. Stimuli were balanced so that the correct answer was on the left exactly 50% of the time, and each dot colour appeared on a more numerous (i.e., correct answer) set exactly 50% of the time. After each block of trials, the child was given the option to continue or stop playing, up to a maximum of ten blocks. About half of the children (54%) continued through all ten blocks of trials; the mean number of completed trials per child was 50 ($SD = 11$).

Results and discussion

Fourteen children were classified as subset-knowers and 32 as CP-knowers. As in previous studies, children's exact-number knowledge was strongly correlated with their performance on the ANS task (Figure 3), $r(44) = .63, p < .001$. However, if we look only at the subset-knowers (i.e., if we exclude the children who understand cardinality), performance on the numeric comparison task drops to 51.76% correct, which is not significantly different from chance. At the level of individual children, only two of the subset-knowers performed above chance at the $p = .10$ level, which is roughly the expected number of type-I errors for that many tests. In contrast, CP-knowers as a group performed well above chance, $t(31) = 10.13, p < .001$, and 29 of 32 performed above chance individually at the $p = .10$ level. These results suggest that subset-knowers did not understand the ANS comparison task, in the sense that they failed to base their responses on numerosity. Chance performance by the subset-knowers was combined with above-chance performance by the CP-knowers to create the illusion of a strong positive correlation between exact-number knowledge and ANS acuity.
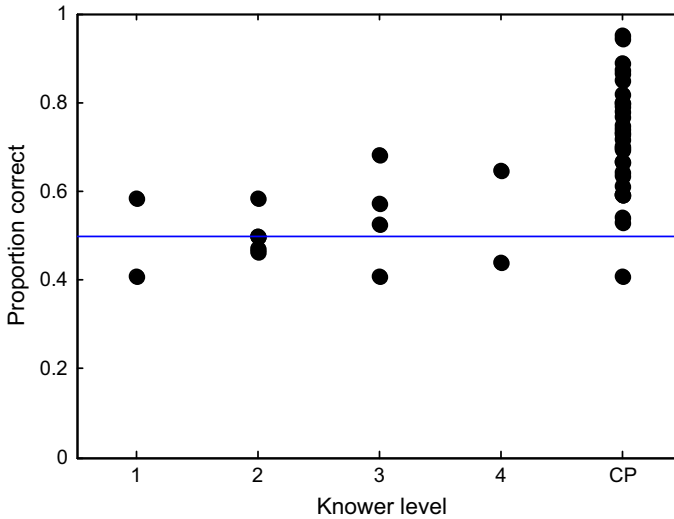
Figure 3. Relation between performance on the standard approximate number system (ANS) task and knower-level in Experiment 1. The solid line reflects chance performance.

Secondary analyses suggest that area was indeed a distracting factor for subset-knowers, at least on the group level. Subset-knowers chose the side with larger dots on 354/647 = 54.7% of trials, p = .0074. Because of this, subset-knowers performed above chance specifically on the area-correlated trials, 181/324 = 55.86%, p = .0151, and marginally below chance on area-anticorrelated trials, 150/323 = 46.44%, p = .1104. CP-knowers, in contrast, performed above chance on both trial types, 584/808 = 72.28% and 590/814 = 72.48%, ps < .0001. CP-knowers also did not pick the side with larger dots more than chance overall, 808/1,622 = 49.82%, p = .45.

EXPERIMENT 2

The difference between CP-knowers and subset-knowers with regard to numerosity is one of attention and salience. Subset-knowers, like humans of all ages and non-human animals, surely have an ANS and can perceive differences in numerosity. How can experimenters get subset-knowers to focus on numerosity for the ANS task? How can we communicate to very young children that *number* is what they should be looking for in this task?

Experiment 2 attempts to do this by modifying the ANS task in several ways. First and most important, we introduced a set of training trials using the ratio of 1:3. The purpose of these trials was to ensure that children based their responses on numerosity, rather than area or some other variable. Because the ratio of 1:3 is very easy to discriminate, we reasoned that children who answered incorrectly on the training trials simply were not attending to numerosity. Children were not allowed to begin the test trials until they had demonstrated (on the training trials) that they interpreted 'more dots' to mean the *more numerous* dot cloud.

Subset-knowers might also perform at chance in the standard task if they felt bewildered or discouraged by the more difficult trials, and/or if the lack of feedback (i.e., the fact that the experimenter seemed happy with whatever answer they gave) left them with the impression that there were no wrong answers and any response was fine. To
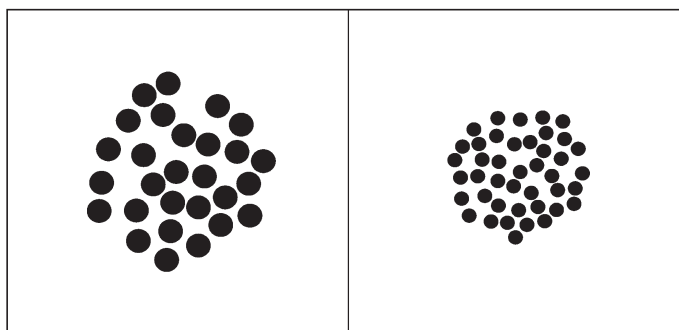
Figure 4. Example of a trial from the modified approximate number system (ANS) task in Experiment 2.

address this, we made three more modifications to the task. First, all trials in a given block tested the same ratio. Second, the blocks were presented in order of increasing difficulty. Third, children received feedback after every trial.

The number of dots was increased to between 20 and 100 dots per side. We also changed several features of the dot clouds to make the two clouds more distinct. These changes included placing the dots randomly around a circle (equal to twice the area of all the dots) rather than spreading them across the page, making all the dots in each cloud a uniform size, and making all the dots black on a white background (see Figure 4).

Method[2]

Participants
Participants included 86 monolingual, English-speaking preschoolers (42 male) with a mean age of 4 years, 4 months (range 2:5–6:1). Children were recruited the same way as in Experiment 1.

Give-N task
This task was completed the same way as in Experiment 1.

Modified ANS task
Children were first given training trials with a ratio of 1:3 and detailed feedback. In particular, if they picked the side with fewer dots but greater total area, they were told, 'Well, *those* dots are *bigger*, but *this* side has *more* dots. They're *smaller*, but there's *more* of them'. There were a total of 24 training trials, but children did not have to do them all. Once a child got eight consecutive training trials correct, they moved on to the test trials. If a child completed all 24 training trials without getting eight consecutive trials correct, the experimenter cycled through the 24 training trials again. This continued until the child answered eight consecutive trials correctly, or the child gave a total of 10 incorrect answers (not necessarily all in a row), or the child became uncooperative.

After the training, blocks progressed from easy to difficult ratios in the following sequence: 1:2, 7:12, 2:3, 17:24, 3:4, 4:5, 5:6, 7:8, and 9:10. (This includes more trials at the easier ratios where, based on the results of Experiment 1, more variation was expected.) Each block had eight trials, counterbalanced for area-correlation (50% correlated; 50% anticorrelated) and for the position of the correct answer (on the left in 50% of trials; on the right in the other 50%).

Results and discussion

Fifteen children were excluded in total. Thirteen children failed to reach criterion on the training trials and did not continue on to the test trials. A logistic regression analysis suggests that the rate of failure to complete training was not related to knower-level when controlling for age, $p = .1$, but was related to age when controlling for knower-level, $p < .01$, suggesting that the task was too difficult for some of the younger children. Additionally, two participants were excluded for scoring extremely low on the numeric comparison task (below 56%, which is the point where performance is significantly different from 50% at the $p = .10$ level), suggesting that they stopped paying attention after the training trials. After these exclusions, 25 of the 71 remaining children were subset-knowers.

Partialling for age, the relationship between knower-level and ANS performance did not reach significance, $r(68) = .16, p = .17$ (Figure 5). In contrast, there was a significant effect of age on ANS performance, $r(69) = .35, p < .01$, even when partialling for knower-level, $r(68) = .23, p < .05$. In other words, children performed better on the ANS task as they got older, but that performance was not closely tied to their knower-level.

Two additional analyses were performed. First, we fit the psychophysics model from Halberda and Feigenson (2008) to these data. Appendix describes the model and how it is fit. The model has a single variance parameter for each child. This parameter can be interpreted as a measure of ANS acuity. There was no significant relation between this parameter and knower-level when partialling for age, $r(60) = -.15, p = .23$. Second, we
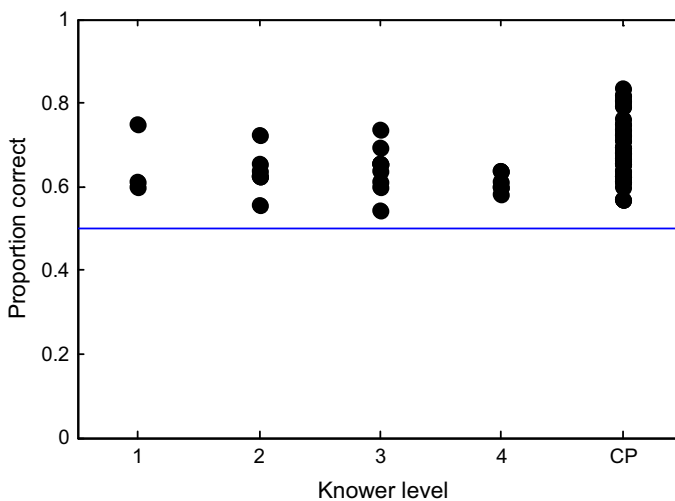

Figure 5. Performance on the modified approximate number system (ANS) task and knower-level in Experiment 2 (n.s.).

checked for a ratio effect, which was strong and significant, $F(8, 5,102) = 30.94$, $p < .001$. (The absence of a ratio effect would have been a strong sign that children were not using the ANS in our task.)

GENERAL DISCUSSION

Taken together, the results from these two experiments suggest that measuring ANS acuity in young children requires special care. Unlike older children and adults, young children (specifically, subset-knowers) may be unlikely to interpret the word 'more' as meaning *more numerous* when asked, 'Which side has more dots?' or 'Are there more blue dots or more yellow dots?' Consequently, subset-knowers may perform at chance when other variables (such as area) are controlled.

On the other hand, it is relatively easy to draw CP-knowers' attention to numerosity, and they seem to have no trouble interpreting 'more' in the way required for the task. This pattern of chance performance by subset-knowers and above-chance performance by CP-knowers on the ANS task can create the impression of a correlation between exact-number knowledge and ANS acuity, when in fact the data cannot be interpreted that way.

In Experiment 1 of the present study, we measured children's exact-number knowledge using the Give-N task and also used a standard ANS task. The results demonstrate how an illusory correlation between exact-number knowledge and ANS acuity can arise. Although exact-number knowledge and ANS performance were indeed highly correlated, children at the lower knower-levels performed at chance, indicating that they may not have understood the task and that we probably did not succeed in measuring their ANS acuity. Earlier papers reporting this correlation have not provided analyses to counter this explanation (Mussolin *et al.*, 2012; Shusterman *et al.*, 2011; Wagner & Johnson, 2011).

In Experiment 2, we modified the ANS task to ensure the children answered on the basis of numerosity. We did this by including only those children who successfully compared sets with a 1:3 ratio. Because a ratio of 1:3 is very easy to distinguish, any normally developing child who is responding based on numerosity should succeed at these trials.

When children were tested using the modified ANS task, no relation was found between exact-number knowledge and ANS acuity. This finding converges with that of Abreu-Mendoza *et al.* (2013), who also found no relation when they used an area-correlated and anticorrelated control. (Those authors did not necessarily interpret this result in terms of ANS acuity, but they did report the same null result between exact-number knowledge and number-comparison performance.)

Our reading of the literature, along with the findings from the two experiments presented here, leads us to conclude that there is no compelling evidence for a relation between exact-number knowledge and ANS acuity in normally developing preschoolers. Of course, one cannot prove a negative. There may be a relation between these two variables, but the existing literature does not show it, nor do the present findings.

As with any negative finding, we must be sure that the lack of correlation we found was not simply due to insufficient statistical power. We address this concern in two ways. First, our Experiment 1 replicated the effect reported in the literature. Experiment 2 (where the effect was not shown) had over 50% more children, including more subset-knowers. Second, our data set was large enough to find the expected ratio effect and the expected effect of age.

Of course, our results here are limited to normally developing young children. A child with a profound deficit in ANS acuity, unable to differentiate a 1:3 ratio, would not make it through the training trials in our Experiment 2 and would not be included in our data. This is a possible explanation for why we found no relation in this age range despite other studies finding a relation at later ages when all children are included (e.g., Halberda *et al.*, 2008). Newer results suggest that ANS acuity differentiates ninth-graders with mathematical learning disabilities from other achievement groups, but not low- from typical-achieving groups (Mazzocco, Feigenson, & Halberda, 2011).

In short, we have shown in Experiment 1 how an illusory relation between exact-number knowledge and ANS performance may arise, if subset-knowers do not understand the ANS task and thus perform at chance. Experiment 2 provided the most rigorous test of this hypothesis to date, by analysing data only from those children who demonstrably understood the task. Under these conditions, the reported correlation between ANS acuity and exact-number knowledge disappears.

These results are convergent with a previous study (Abreu-Mendoza *et al.*, 2013). Thus, studies from two different labs have now independently tried and failed to find a relation between exact-number knowledge and ANS acuity. Of course, one cannot prove a negative, and such a relation may yet be found. But for now, it seems that there is no credible evidence of a relation between ANS acuity and exact-number knowledge in normally developing preschoolers.

Acknowledgements

References

Abreu-Mendoza, R. A., Soto-Alba, E. E., & Arias-Trejo, N. (2013). Area vs. density: Influence of visual variables and cardinality knowledge in early number comparison. *Frontiers in Psychology*, *4*, 1–10. doi:10.3389/fpsyg.2013.00805

Dehaene, S. (1997). *The number sense*. New York, NY: Oxford University Press.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307–314. doi:10.1016/j.tics.2004.05.002

Gallistel, C. R. (2011). Mental magnitudes. In S. Dehaene & L. Brannon (Eds.), *Space, time and number in the brain: Searching for the foundations of mathematical thought* (pp. 3–12). New York, NY: Elsevier.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, *44*, 43–74. doi:10.1016/0010-0277(92)90050-R

Gelman, S. A. (2003). *The essential child: origins of essentialism in everyday thought*. New York, NY: Oxford University Press.

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

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. doi:10.1037/a0012682

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*, 665–668. doi:10.1038/nature07246

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition*, *105*, 395–438. doi:10.1016/j.cognition.2006.10.005

Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011). Impaired acuity of the approximate number system underlies mathematical learning disability (dyscalculia). *Child Development*, *82*, 1224–1237. doi:10.1111/j.1467-8624.2011.01608.x

Mussolin, C., Nys, J., Leybaert, J., & Content, A. (2012). Relationships between approximate number system acuity and early symbolic number abilities. *Trends in Neuroscience and Education*, *1*, 21–31. doi:10.1016/j.tine.2012.09.003

Odic, D., Paul, P., Hunter, T., Lidz, J., & Halberda, J. (2013). Young children's understanding of "more" and discrimination of number and surface area. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 451–461. doi:10.1037/a0028874

Schaeffer, B., Eggleston, H., & Scott, J. (1974). Number development in young children. *Cognitive Psychology*, *6*, 357–379. doi:10.1016/0010-0285(74)90017-6

Shusterman, A., Slusser, E., Halberda, J., & Odic, D. (2011). *Connecting early number word knowledge and approximate number system acuity*. Poster presented at the Biennial Meeting of the Society for Research Development in Child Development, Montreal, Quebec.

Slusser, E. B., & Sarnecka, B. W. (2011). Find the picture of eight turtles: A link between children's counting and their knowledge of number word semantics. *Journal of Experimental Child Psychology*, *110*, 38–51. doi:10.1016/j.jecp.2011.03.006

Wagner, J. B., & Johnson, S. C. (2011). An association between understanding cardinality and analog magnitude representations in preschoolers. *Cognition*, *119*, 10–22. doi:10.1016/j.cognition.2010.11.014

Appendix : Model-based approach to ANS task analysis

In Experiment 2, we employ a model-based approach to analysing data from the ANS task, following Halberda and Feigenson (2008). Because the details of this model may be unfamiliar to some readers, we describe it in detail below.

The model has three principle assumptions. First, the child approaches the comparison task by making an estimate of the number of items in each of the two arrays and then seeing which estimate is larger. Second, the child's estimates are normally distributed with a centre at the correct answer. Third, the standard deviation of these estimates is the number of items times a variable w, which is constant within each child but can vary between children.

The child's estimate of N items is described by a normal distribution with a mean of N and a standard deviation of wN. The probability that a child will say X items are more than Y items is the probability that a draw from Normal (X, wX) is higher than a draw from Normal (Y, wY). The formula for this probability is $\Phi\big((X - Y)/\sqrt{(wX)^2 + (wY)^2}\big)$, where is the normal cumulative distribution function.

This model has several appealing properties. First of all, it is very simple. It has only one parameter with a very clear interpretation: Lower w means less noise in the child's numeric estimates. Second, it can be fit with a simple numeric search (described below). Third, it reflects both scalar variability (e.g., Gallistel & Gelman, 1992), because the standard deviation of estimates is proportional to the mean, and also Weber's Law, because the chance of getting a correct answer is dependent on the ratio of the number of items being compared.

The single parameter w was fit for each child by starting with $w = .01$ and calculating the full log-likelihood of their data, then incrementing w by .01 until a maximum log-likelihood was reached. Children were excluded from the model-based analysis if they were best-fit by a value of w over 1, which seems psychologically implausible.