Sampling from the Mental Number Line:

How are Approximate Number System Representations Formed?

Matthew Inglis and Camilla Gilmore

Mathematics Education Centre, Loughborough University

Mathematics Education Centre

Loughborough University

Loughborough. LE11 3TU.

United Kingdom

m.j.inglis@lboro.ac.uk

3863 words

Abstract

Nonsymbolic comparison tasks are commonly used to index the acuity of an individual's

approximate number system (ANS), a cognitive mechanism believed to be involved in the

development of number skills. Here we asked whether the time that an individual spends

observing numerical stimuli influences the precision of the resultant ANS representations.

Contrary to standard computational models of the ANS, we found that the longer the stimulus

was displayed, the more precise was the resultant representation. We propose an adaptation

of the standard model, and suggest that this finding has significant methodological

implications for numerical cognition research.

SAMPLING FROM THE MENTAL NUMBER LINE: HOW ARE APPROXIMATE

NUMBER SYSTEM REPRESENTATIONS FORMED?

Fluency with mathematics is essential for day-to-day life. To successfully interact in a modern society it is frequently necessary to interpret, compare and calculate with numerical quantities. Along with a capacity to understand numerical ideas when represented symbolically, humans also have an Approximate Number System (ANS) which can be used to perform arithmetic operations on non-symbolic quantities such as arrays of dots or tones. The ANS is present in very young infants and some non-human animals (for a review see Dehaene, 1997), and recently some theorists have begun to speculate that it serves as the cognitive basis for symbolic mathematics (e.g. Halberda, Mazzocco, & Feigenson, 2008).

The ANS is widely believed to follow Weber's law: the standard model proposes that when we encounter non-symbolic stimuli, a box of n oranges say, the distribution of possible ANS representations follows a normal distribution with mean n and standard deviation wn. Here w is the Weber fraction, a parameter which represents the acuity of an individual's ANS (e.g., Barth, La Mont, Lipton, Dehaene, Kanwisher & Spelke, 2008). Several recent studies have shown that individuals' ANS acuities are correlated with achievement in symbolic mathematics (e.g. Gilmore, McCarthy & Spelke, 2010; Halberda et al., 2008; Libertus, Feigenson & Halberda, 2011; Mazzocco, Feigenson & Halberda, 2011a, 2011b; but see Inglis, Attridge, Batchelor & Gilmore, 2011; Price, Palmer, Battista & Ansari, 2012), lending credence to the suggestion that the ANS is implicated in the development of symbolic mathematics competence.

Although the capabilities of the ANS are now fairly well understood, the process by which the ANS forms representations from visual numerical stimuli is less clear. Several researchers have proposed that a mental 'accumulator' is central to this process (e.g. Dehaene

& Changeux, 1993; Gallistel & Gelman, 2000; Piazza & Izard, 2009; Verguts & Fias, 2004).

Gallistel and Gelman drew an analogy between filling up a beaker with cups of liquid, and

filling up the accumulator with "accumulator units". They suggested that when an array of

objects is observed, the scene is first normalized to remove numerically-irrelevant between-

object differences (color, shape, size etc), then one cupful of 'liquid' is added to the

accumulator per item. The contents of the accumulator are then emptied into memory which

introduces noise proportionate to the accumulator's contents (the "sloshing" of liquid in the

memory beaker, in Gallistel and Gelman's analogy). It is this noise, when the contents of the

memory beaker are read off (converted into a numerical quantity), which causes the

approximate nature of ANS representations.

It is notable that both Barth et al.'s (2008) computational model of the ANS, and

Gallistel and Gelman's (2000) accumulator beaker analogy assume that the duration for

which a numerical stimulus is displayed is irrelevant to the ANS representation that an

individual encodes from it. To date this assumption has not been tested. We see two reasons

for questioning it. First, earlier researchers have reported different Weber fractions in studies

which have used different display times. For example, a dot comparison task with a stimuli

duration of 200ms resulted in less precise ANS representations (mean $w = 0.3$, Halberda et

al., 2008) than one with a display time of 750ms (mean $w = 0.1$, Halberda & Feigenson,

2008). Second, it is well known that performance on many other visual tasks is dependent on

stimuli durations (e.g., visual search: Guest & Lamberts, 2011; McElree & Carrasco, 1999;

absolute identification: Guest, Kent & Adelman, 2010).

Our goal in this paper is to explore whether the precision of an individual's ANS

representation varies with the length of time they spend studying the numerical stimulus. This

question is important for at least two reasons. First because, as discussed above, it sheds light

on the underlying mechanism that the ANS uses to form representations. Second, because

numerical cognition researchers have to date adopted widely varying methods when conducting experimental studies. When presenting numerical comparison tasks (where participants are shown two dot arrays and asked to determine which is more numerous), some researchers have permitted participants to decide how long to study the stimuli before reaching a judgement (e.g., Inglis et al., 2011; Pica et al., 2004), whereas others have displayed the stimuli for a fixed period. Among those who have used fixed stimuli durations, some have displayed stimuli for as little as 200ms (e.g., Halberda et al., 2008) whereas others have used up to 2500ms (e.g., Halberda & Feigenson, 2008), and some researchers have used different stimuli durations for different participants within the same experiment (e.g., Halberda & Feigenson, 2008; Mazzocco, Feigenson & Halberda, 2011b). All these authors have assumed that these methods investigate the same underlying process, but if the formation of ANS representations is time dependent then it is questionable whether results from these studies are comparable.

Here we report two experiments which directly investigated whether the acuity of ANS representations encoded from visual stimuli varies with stimuli duration. In Experiment 1 we demonstrate that individuals' accuracies and Weber fractions are strongly dependent on stimuli duration, in Experiment 2 we show that this is not the result of differing onset-to-decision latencies, and in the general discussion we propose an adaptation of the standard model of the ANS which accounts for these data.

Experiment 1

Method

Participants. Participants were 12 staff or students of Loughborough University with normal or corrected-to-normal vision, who participated in exchange for a small inconvenience allowance. The study took place in a quiet laboratory using a 15" laptop.

Procedure. Each of 400 trials began with a fixation cross which was displayed for 1000ms. This was followed by two dot arrays (a red array on the left of the screen and a blue array on the right) which were displayed for either 16ms (the refresh rate of the monitor), 300ms, 600ms, 1200ms or 2400ms. After the alloted time period the dot arrays were replaced by two red and blue masks and a question mark. Participants were then required to select which of the arrays was the more numerous by pressing either a blue or red key on a response box. No feedback was given to participants. Trials were blocked by display time, and each participant was given the blocks in a random order. Each block consisted of 80 trials (which were identical between blocks) and was preceeded by a practice block of 10 trials. The problems used numerosities in the range 5 to 21, with comparison ratios of approximately 0.5, 0.6, 0.7 and 0.8. Each problem appeared twice, once where the larger numerosity was on the left hand side of the screen, and once when it was on the right. Stimuli were created using Gebuis and Reynvoet's (2011) method. The paradigm is summarised in Figure 1.

Figure 1. An illustration of the procedure used in Experiment 1.
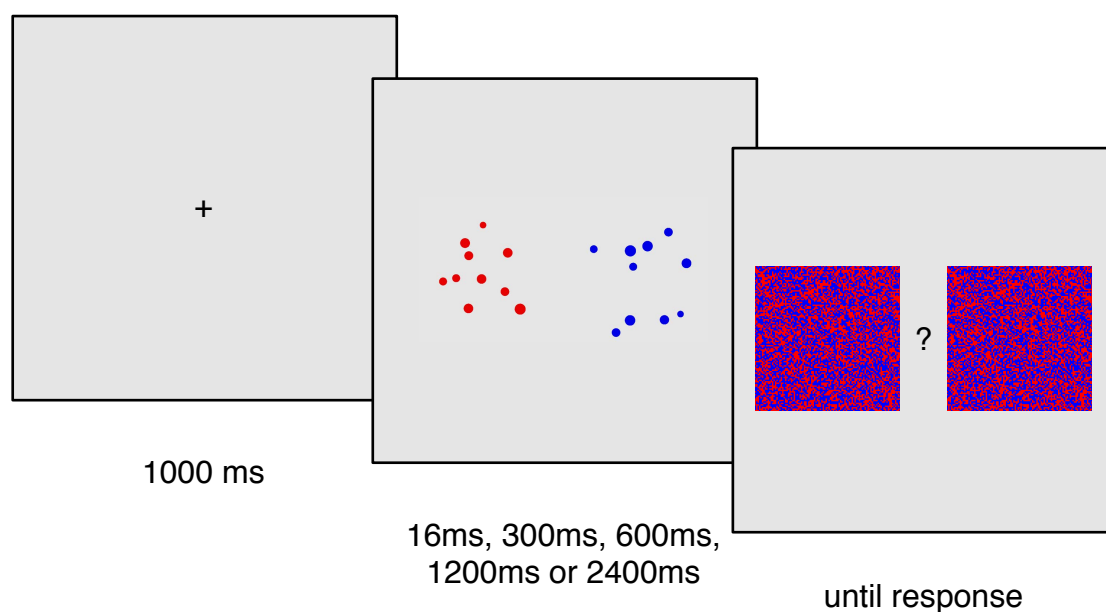
Modelling. As well as calculating each individual's accuracy, we fitted participants'

data to the standard computational model of the ANS (Barth et al., 2008) using the log

likelihood method. According to the standard model, accuracy on a given trial is a function of

the numerosities involved and the individual's Weber fraction:

$$\mathrm{acc}(n_1, n_2; w) = \frac{1}{2} + \frac{1}{2}\,\mathrm{erf}\!\left(\frac{|n_1 - n_2|}{\sqrt{2}\,w\sqrt{n_1^2 + n_2^2}}\right)$$

Where n_1 and n_2 are the to-be-compared numerosities, w is the Weber fraction and erf is the

error function.

Results.

Participants' accuracies varied from .77 to .92 ($M = .85$, SD $= 0.05$), and their overall

Weber fractions varied between 0.18 and 0.39 ($M = 0.27$, SD $= 0.06$). We first calculated

participants' Weber fractions separately for each of the five display durations. The mean ws

were 0.57, 0.29, 0.25, 0.19 and 0.17 for the 16ms, 300ms, 600ms, 1200ms and 2400ms

conditions respectively, $F(1.081, 0.065) = 16.636$, $p = .001$, $\eta_p^2 = .602$ (Greenhouse-Geisser

correction), which represented a significant linear trend, $F(1, 10) = 23.348$, $p = .001$, $\eta_p^2 =$

.680.

Mean accuracies for each of the comparison ratios are shown in Figure 2, and were

analysed using a 4 (comparison ratio: 0.5, 0.6, 0.7, 0.8) × 5 (stimuli duration: 16ms, 300ms,

600ms, 1200ms, 2400ms) within-subjects analysis of variance (ANOVA). As is characteristic

of ANS tasks, the main effect of ratio was significant, $F(3, 33) = 143.076$, $p < .001$, $\eta_p^2 =$

.929, and also showed a significant linear trend, $F(1, 11) = 408.307$, $p < .001$, $\eta_p^2 = .974$.

Critically, there was also a significant effect of stimuli duration, $F(4, 44) = 28.638$, $p < .001$,

$\eta_p^2 = .722$, which also showed a significant linear trend, $F(1, 11) = 49.075$, $p < .001$, $\eta_p^2 =$

.817. The longer the stimuli duration, the more accurate participants were. The interaction

effect did not approach significance, $F(12, 132) = 1.550$, $p = .114$.

Figure 2. Participants' mean accuracies in each of the five stimuli duration conditions in

Experiment 1, by comparison ratio. Error bars show ±1 SE of the mean.
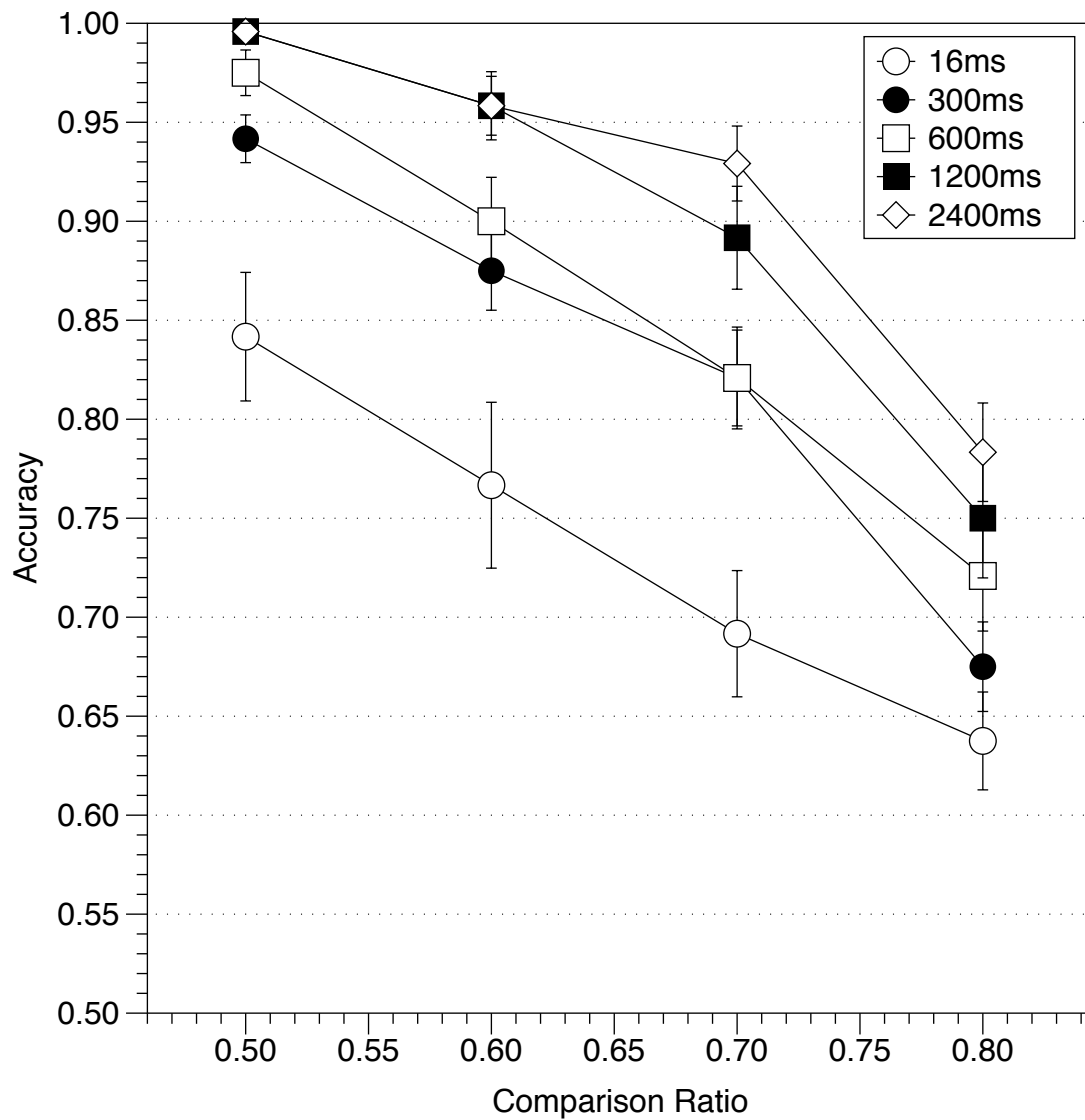
Next we used one-sample *t*-tests to assess each mean accuracy figure to determine whether participants were performing at above chance levels. Every mean was well above 50%, all *p*s < .001, suggesting that participants were able to perform comparisons of numerosities with ratios 4:5 in only 16ms. To check whether participants' relatively high accuracies in the 16ms condition was the result of cross-block learning, we considered the performance of the three participants who received the 16ms block first. Each performed at above chance levels (74%, 66% and 63%, all *p*s < .02) and were well within the range of other participants (59% to 88%), suggesting that the ANS is capable of processing numerosities under extreme time pressure. No participant reported failing to see the stimuli in the 16ms condition, perhaps because we did not use a forward mask; it therefore remains to be seen whether the ANS can process numerical concepts subliminally.

Discussion

Two notable findings emerged from Experiment 1. First, participants were able to successfully encode ANS representations from stimuli displayed for as little as 16ms, suggesting that the process of forming ANS representations is automatic, or at least extremely rapid. Second, although participants were reliably able to complete the comparison task in 16ms, they were significantly more accurate with longer display times.

These findings indicate that forming ANS representations and using them to generate behavior is a time-dependent process. An obvious question concerns where in Gallistel and Gelman's (2000) analogy a time dependency could occur. We see three possibilities, which we discuss in turn. First, it could be that the input to the ANS becomes more precise with more time. In other words, that when participants view arrays of dots the precision of their initial visual processing increases with time. Below we suggest that there are theoretical and empirical reasons for doubting this account. A second possibility is that the output of the ANS is processed differently under different time constraints. Perhaps representations from

the ANS are generated extremely rapidly upon stimuli onset, but the process used to compare the two representations – the last stage in Gallistel and Gelman's analogy – is time dependent. We rule out this possibility in Experiment 2. Finally, it could be that the precision of participants' ANS representations themselves increase with time. We discuss a possible mechanism for this latter account in the general discussion at the end of the paper.

The first account suggests that participants' initial visual processing is more precise in the slower conditions, and that therefore the input to the ANS is more precise as well. We see three reasons to doubt this account. First, the initial stages of Gallistel and Gelman's (2000) account can be seen as equivalent to the initial preattentive stage of accounts of visual search behavior (e.g. Triesman & Gelade, 1980; Wolfe, 1994). Since the two dot arrays on typical comparison tasks differ on only one salient feature (typically color), the initial visual processing on each trial is directly analogous to that on single feature visual search tasks (i.e. where participants are asked to, for example, find a blue A among red As). According to the Guided Search account (Wolfe, 1994), when faced with blue- and red-colored stimuli of the type presented in our dot comparison tasks, a color 'feature map' is preattentively and automatically constructed. This map contains tagged activation levels corresponding to each blue or red item in the visual scene. It is only when different feature maps need to be integrated that slow serial visual processing is recruited (e.g. where an individual is searching for a blue A among red As and blue Hs then information from both the color and the shape feature maps needs to be integrated). We suggest that it is the number of blue and red activation locations in this preattentive feature map which provides an input to the ANS. If this account is correct, then the initial visual processing would not be time dependent as the feature map is constructed automatically in parallel.

Second, if the initial visual processing of the dots were a time-dependent serial process, we would expect the number of dots in each trial to be predictive of the trial's

difficulty and, moreover, that the number of dots would interact with the stimuli duration (on short trials we would expect that the difficulty difference between problems with large numbers of dots and those with small numbers of dots would be lower than on long trials). To investigate this we categorized each trial as being either 'large' or 'small' (based on a median split of the total number of dots on each trial), and analyzed participants' accuracy data using a 5 (duration) × 2 (size) within-subjects ANOVA. We found neither an interaction effect, $F(4,44) = 1.015$, $p = .410$, nor a main effect of size, $F(1,44) = 3.680$, $p = .081$. Both these findings are consistent with the suggestion that the initial visual stage of processing in dot comparison tasks takes place in parallel.

Finally, if the initial stage of processing were, in part at least, driving individuals' ANS acuities, then we would expect individual differences in visual processing speeds to be related to ANS acuities. But Simms, Clayton, Cragg, Gilmore, Marlow, & Johnson (2013) found no relationship between performance on Anderson, Reid and Nelson's (2001) measure of visual processing speed, and a typical ANS dot comparison task, $r = .001$, $p = .991$.

The second and third accounts of where the time dependency could occur in Gallistel and Gelman's (2000) analogy are both consistent with the data reported in Experiment 1. In particular, in Experiment 1 stimuli duration was confounded with the latency between stimuli onset and participants' decision points. In other words, it might be possible that participants' increased accuracy in conditions with longer display times was not due to more precise ANS representations (as proposed by the third account), but rather to having longer to perform the arithmetical comparison operation upon representations which were automatically generated at the stimuli onset (as proposed by the second account). This is especially plausible given Guest et al.'s (2010) finding that in absolute identification tasks it is the onset-to-decision latency that influences accuracy, not stimuli display time. We conducted a second experiment to disentangle these factors.

Experiment 2

The primary goal of Experiment 2 was to determine whether the effect found in Experiment 1 was due to stimuli duration, or the onset-to-decision latency. Here we held stimuli duration constant and varied the onset-to-decision latencies.

Method

Participants were 11 staff or students of Loughborough University. The procedure and stimuli were identical to Experiment 1 except that the two dot arrays were displayed for 48ms on each trial. After the 48ms had elapsed the red-blue masks were displayed and then, after either 0ms, 252ms, 552ms, 1152ms or 2352ms, a question mark appeared to signal that participants should select which array was the more numerous. This paradigm is summarized in Figure 3. If the main factor driving the findings in Experiment 1 were the onset-to-decision latency, we would expect a similar set of data in Experiment 2.

Figure 3. An illustration of the procedure used in Experiment 2.
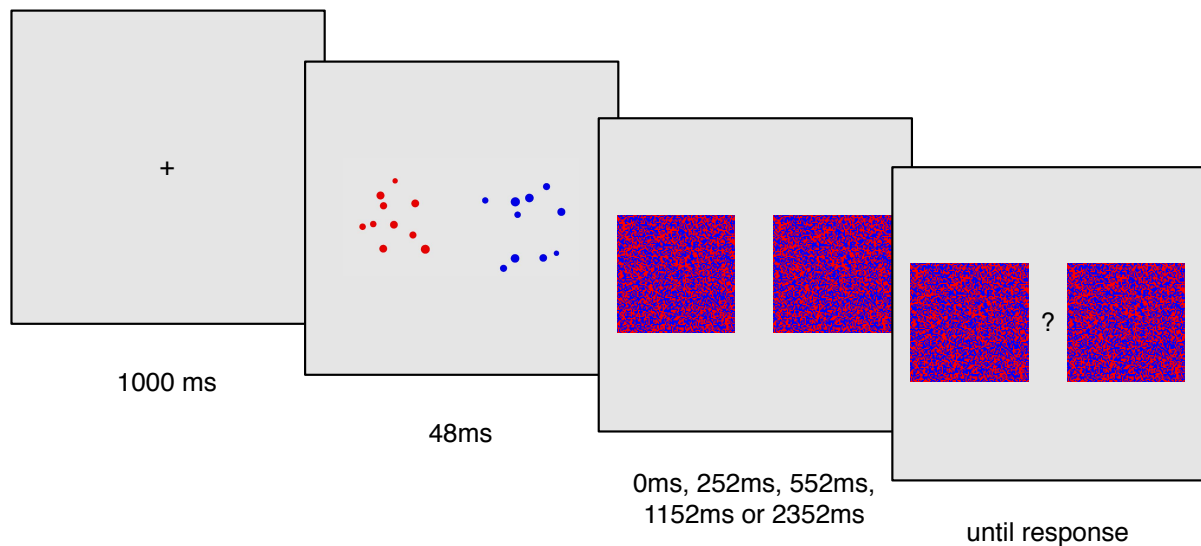
Results and Discussion

Participants' *w* parameters were calculated for each of the five delay times. The mean

*w*s from the five conditions (0.49, 0.54, 0.42, 0.44 and 0.39) were not significantly different,

$F(2.154, 21.537) = 2.623$, *p* =.092 (Greenhouse-Geisser correction). The accuracy data are

shown in Figure 4. These data were analyzed using a 4 (comparison ratio: 0.5, 0.6, 0.7, 0.8) ×

5 (onset-to-decision latency: 48ms, 300ms, 600ms, 1200ms, 2400ms) within-subjects

analysis of variance (ANOVA). As before, the main effect of ratio was significant, $F(3, 30) =$

38.724, *p* < .001, $\eta_p^2 = .795$, and also showed a significant linear trend, $F(1, 10) = 70.239$, *p* <

.001, $\eta_p^2 = .875$. However, critically, the main effect of onset-to-decision latency did not

approach significance, $F(4, 40) = 1.478$, *p* = .227, suggesting that there were no systematic

accuracy differences between the five different conditions. Consequently we can rule out the

possibility that the accuracy differences between display times observed in Experiment 1

were due to differences in the onset-to-decision latencies of the different conditions.

Figure 4. Participants' mean accuracies in each of the five onset-to-decision conditions in Experiment 2, by comparison ratio. Error bars show ±1 SE of the mean.
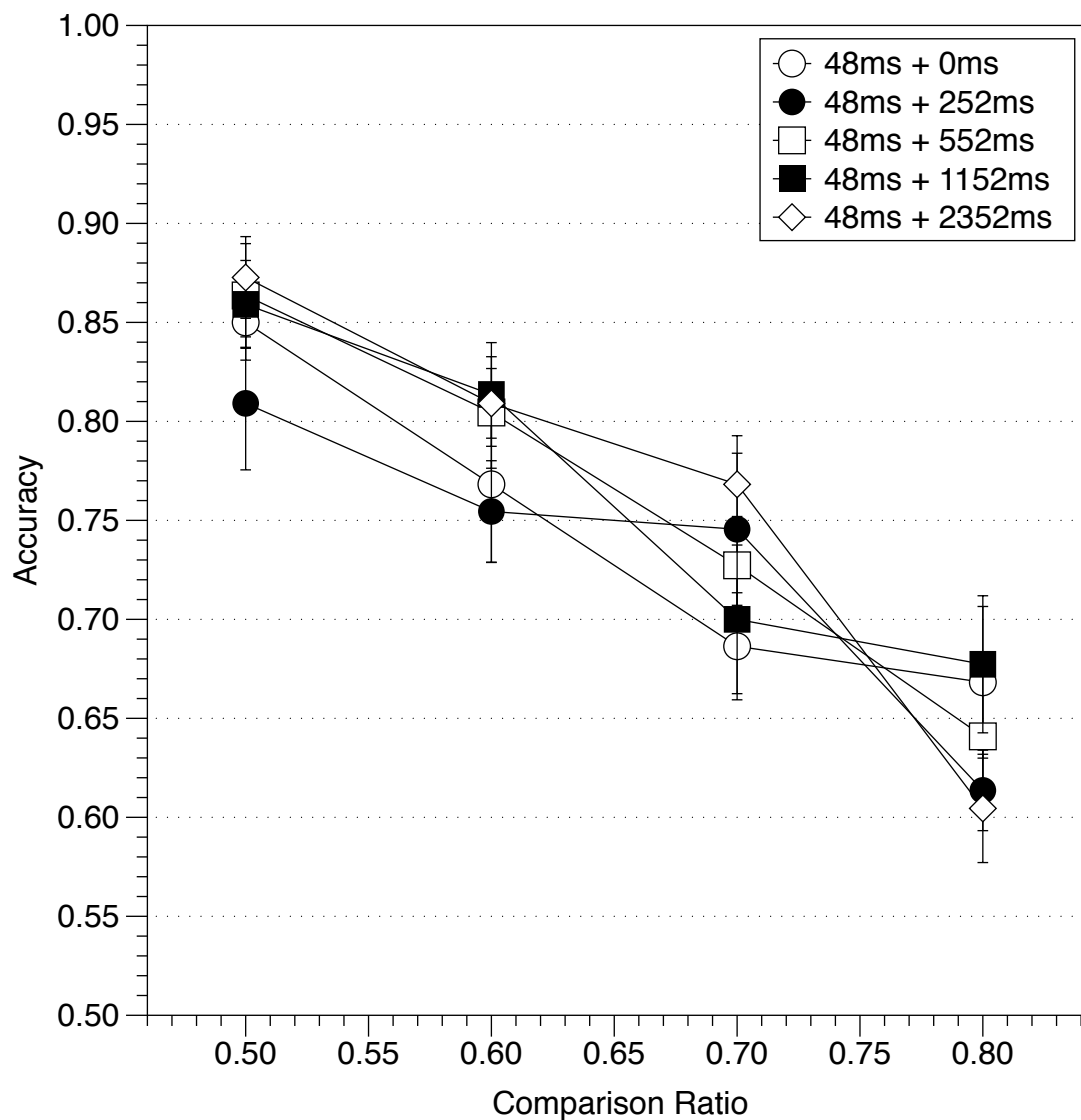
General Discussion

Summary of Main Findings

Our primary goal was to investigate how the duration of numerical stimuli influences the acuity of resultant ANS representations. Noting that earlier researchers have used dramatically different display times to estimate the acuity of individuals' ANSs, in

Experiment 1 we systematically varied stimuli display times on a dot comparison task. We found that participants were able to perform at above chance levels when stimuli were displayed for only 16ms, but that, contrary to the assumption of the standard model, when stimuli were displayed for longer participants responded more accurately. In Experiment 2 we rejected the suggestion that this effect could be due to different onset-to-decision latencies rather than stimuli display times. It therefore seems plausible to suppose that, instead of the higher accuracies in longer conditions found in Experiment 1 being the result of more successful manipulation of similarly precise ANS representations, it was the precision of the ANS representations themselves which varied between conditions.

Taking multiple samples from visual numerical stimuli

In this section of the paper we propose a modification of the standard computational model of the ANS which takes account of our results. Recall that an individual's ANS representation for a numerosity n is traditionally said to follow a normal distribution with mean n and standard deviation wn, in other words $N \sim \mathrm{N}\left(n, (wn)^2\right)$. We propose that when an individual observes a numerical stimuli, rather than taking a single sample from this distribution, they actually take many (the number determined by a function of the display time) and use the mean as the resultant ANS representation. In other words, we suggest that participants go through the first stages of Gallistel and Gelman's (2000) analogy multiple times (perceptual encoding, normalizing the visual scene, filling up the accumulator with liquid, transferring the liquid to the 'memory beaker' and taking a reading), before using the average of their multiple samples – or 'beaker readings' – as the final ANS representation. Assuming that the individual takes $f(t)$ samples, then the resultant ANS representation will follow the distribution of sample means from N: $\overline{N} \sim \mathrm{N}\left(n, \dfrac{(wn)^2}{f(t)}\right)$.

A natural question concerns the identity of the function $f(t)$. Information accumulation models typically assume that stimuli onset is accompanied by rapid information accumulation, the rate of which gradually decreases towards some asymptotic limit (e.g. McElree & Carrasco, 1999). In contrast, in the case of taking samples from the N distribution, there appears to be no *a priori* reason to suppose that there would be a theoretical maximum number of samples that an individual could take. Consequently we suggest that a reasonable candidate function is $f(t) = \alpha \sqrt[k]{t}$, where α and k are parameters which determine the rate of information accumulation and which vary between individuals. Notice that the standard model of the ANS, which assumes a single sample is taken, is a restriction of this proposal, as $f(t) = 1$ when $\alpha = 1$ and $k = \infty$.

Given this proposal we would expect participants' accuracies on a given trial to be a function of the to-be-compared numerosities n_1 and n_2, the display duration t, their Weber fraction w, α and k:

$$\text{acc}(n_1, n_2, t; w, \alpha, k) = \frac{1}{2} + \frac{1}{2} \text{erf}\left(\frac{\sqrt[2k]{t}\sqrt{\alpha}\,|n_1 - n_2|}{\sqrt{2}w\sqrt{n_1^2 + n_2^2}} \right)$$

We fitted each participants' data from Experiment 1 to this model, treating $\frac{w}{\sqrt{\alpha}}$ as a single parameter. Values of $\frac{w}{\sqrt{\alpha}}$ ranged from 0.26 to 3.90 ($M = 1.23$, $SD = 1.08$), and values of k ranged from 0.66 to 8.07 ($M = 1.88$, $SD = 2.04$), indicating that there were substantial individual differences in the rate at which samples were taken. Overall the time-dependent model proved to have a significantly better fit to the data than the standard model, likelihood ratio test, $\chi^2(1) = 238.3$, $p < .001$.

Given the large individual differences in the k parameter, we investigated the relationship between individuals' acuity parameters from the two models (w from the

standard model and $\dfrac{w}{\sqrt{\alpha}}$ from the time-based model). Although these parameters were

correlated, $r = .71$, $p = .009$, this relationship was far from exact, as shown in Figure 5. This

observation is important, because if our analysis is correct then what earlier researchers have

reported as w parameters have actually been figures for $\dfrac{w}{\sqrt{\alpha}\,\sqrt[2k]{t}}$, suggesting that comparing

Weber fractions between studies which have used different display times is flawed. We

expand upon this remark in the remaining section of the paper.

Figure 5. Participants' w parameters (derived from the standard model) plotted against their

$\dfrac{w}{\sqrt{\alpha}}$ parameters (derived from the time-based model).
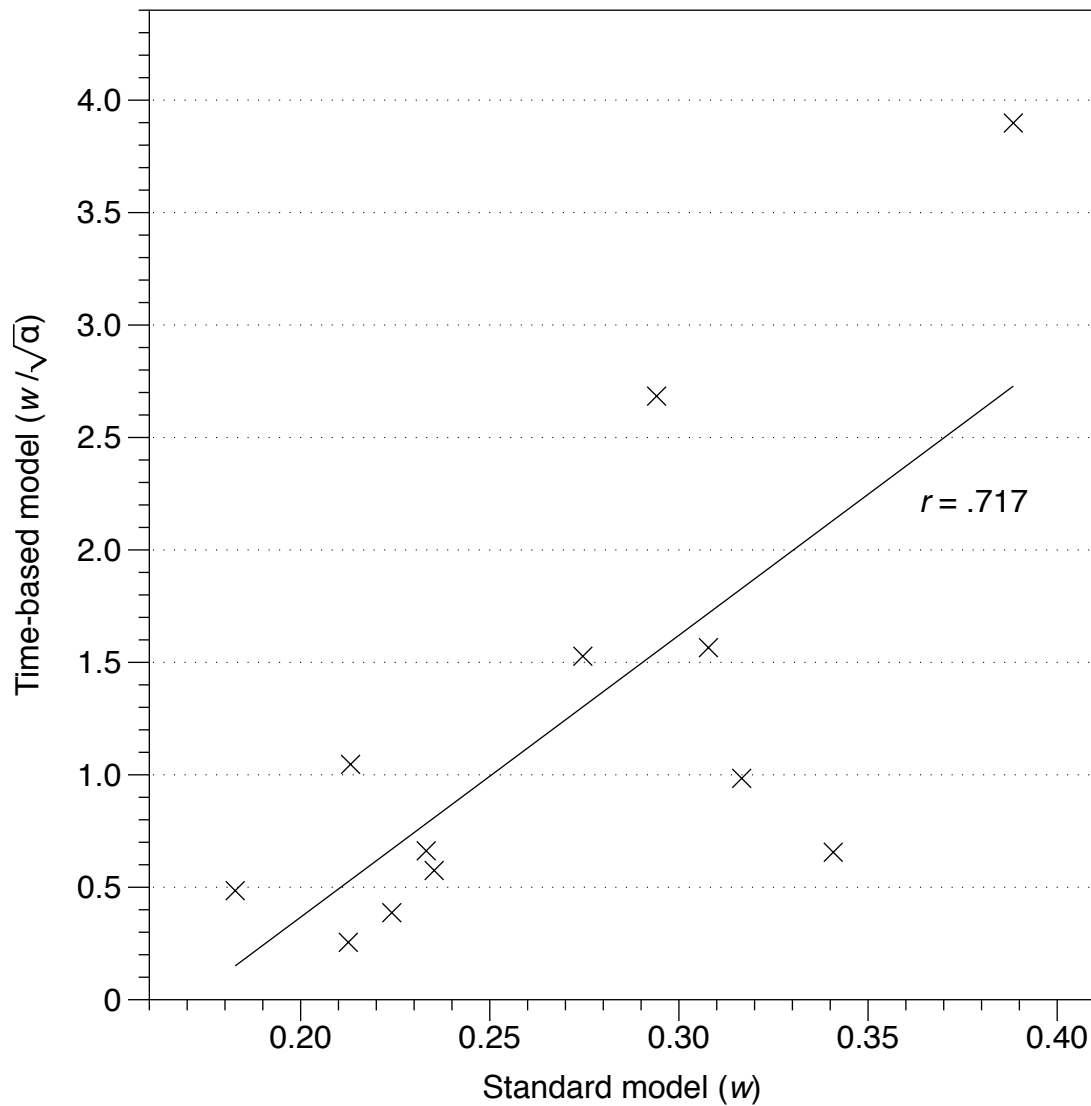
Methodological implications

       Our finding that stimuli durations influence the acuity of ANS representations has

important methodological implications for numerical cognition researchers. We conclude the

paper by highlighting two.

       First, we believe that the comparison of Weber fractions between tasks which have

used different methods is problematic. For example, Gilmore, Attridge & Inglis (2011) gave

participants a nonsymbolic comparison task and a nonsymbolic addition task (where

participants were asked to determine the larger of n_1+n_2 and n_3), and surprisingly found that

the ws derived from each did not correlate. They argued that this called into question the suggestion that a single system, the ANS, was used to complete these tasks. Our findings here suggest an alternative account. As is common, Gilmore et al. presented their comparison task concurrently (i.e. both n_1 and n_2 were onscreen at the same time), and allowed participants to respond at their own pace; but on their addition task, they presented n_1, n_2 and n_3 consecutively, each for 500ms. It may be that this discrepancy in stimuli duration was the cause of the lack of correlations between performance observed on the two tasks.

Second, researchers who have investigated how the ANS develops through childhood have typically used different display times with different aged children. For example, Halberda and Feigenson (2008) used displays of 2500ms for 3-year-olds, 1200ms for 4-, 5- and 6- year olds, and 750ms for adults, and Mazzocco et al. (2011b) used display times of 1200ms and 2500ms for the two age groups in their study, combining the data into a single analysis. Our findings indicate that this analysis strategy may be flawed, and that Weber fractions derived from tasks with different display times are not comparable. In future, researchers interested in individual differences in ANS acuities should pay attention to how their stimuli are displayed.

References

Anderson, M., Reid, C., & Nelson, J. (2001). Developmental changes in inspection time: what a difference a year makes. *Intelligence, 29,* 475-486.

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.

Dehaene, S. (1997). *The number sense*. Oxford, UK: Oxford University Press.

Dehaene, S. & Changeux, J.-P. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience, 5,* 390-407.

Gebuis, T. & Reynvoet, B. (2011). Generating nonsymbolic number stimuli. *Behavior Research Methods, 43,* 981-986.

Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: from reals to integers. *Trends in Cognitive Science, 4,* 59-65.

Gilmore, C., Attridge, N., & Inglis, M. (2011). Measuring the approximate number system. *Quarterly Journal of Experimental Psychology, 64*, 2099-2109.

Gilmore, C.K., McCarthy, S.E., & Spelke, E.S. (2010). Non-symbolic arithmetic abilities and achievement in the first year of formal schooling in mathematics. *Cognition, 115,* 394-406.

Guest, D., Kent, C., & Adelman, J. S. (2010). Why additional presentations help identify a stimulus. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 1609–1630.

Guest, D. & Lamberts, K. 2011. The time course of similarity effects in visual search. *Journal of Experimental Psychology: Human Perception and Performance, 37,* 1667–1688.

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.

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

Inglis, M., Attridge, N., Batchelor, S., & Gilmore, C. (2011). Non-verbal number acuity correlates with symbolic mathematics achievement: But only in children. *Psychonomic Bulletin & Review, 18,* 1222-1229.

Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science, 14*, 1292–1300.

Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011a). Impaired acuity of the approximate number system underlies mathematical learning disability (Dyscalculia). *Child Development, 82*, 1224–1237.

Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011b). Preschoolers' precision of the approximate number system predicts later school mathematics performance. *PloS One*, 6(9), e23749.

McElree, B. & Carrasco, M. (1999). The temporal dynamics of visual search: Evidence for parallel processing in feature and conjunction searches. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1517-1539.

Piazza, M. & Izard, V. (2009). How humans count: Numerosity and parietal cortex. *The Neuroscientist, 15,* 261-273.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science, 306,* 499-503.

Price, G. R., Palmer, D., Battista, C., & Ansari, D. (2012). Nonsymbolic numerical

magnitude comparison: Reliability and validity of different task variants and outcome

measures, and their relationship to arithmetic achievement in adults. *Acta*

Psychologica, 140, 50-57.

Simms, V., Clayton, S., Cragg, L., Gilmore, C., Marlow, N. & Johnson, S. (2013, April).

Counting and Executive Functions, Not Basic Numerical Representations, Predict

Mathematical Achievement in 8-to-10-year-olds. Society for Research in Child

Development Biennial Conference, Seattle, USA.

Triesman, A., & Gelade, C. (1980). A feature-integration theory of attention. *Cognitive*

Psychology, 12, 97-136.

Verguts, T., & Fias, W. (2004). Representation of number in animals and humans: a neural

model. *Journal of Cognitive Neuroscience, 16*, 1493–1504.

Wolfe. J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic*

Bulletin & Review, 1, 202-238.

Author Note