Cognition 122 (2012) 454–459

Contents lists available at SciVerse ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Brief article

Linear mapping of numbers onto space requires attention

Giovanni Anobile [a], Guido Marco Cicchini [b], David C. Burr [a,b,*]

[a] Department of Psychology, University of Florence, Via S. Salvi 12, Florence, Italy
[b] CNR Institute of Neurosciences, Pisa, Italy

ARTICLE INFO

Article history:
Received 4 March 2011
Revised 3 November 2011
Accepted 11 November 2011
Available online 10 December 2011

Keywords:
Number
Mental number line
Attention
Central tendency

ABSTRACT

Mapping of number onto space is fundamental to mathematics and measurement. Previous research suggests that while typical adults with mathematical schooling map numbers veridically onto a linear scale, pre-school children and adults without formal mathematics training, as well as individuals with dyscalculia, show strong compressive, logarithmic-like non-linearities when mapping both symbolic and non-symbolic numbers onto the numberline. Here we show that the use of the linear scale is dependent on attentional resources. We asked typical adults to position clouds of dots on a numberline of various lengths. In agreement with previous research, they did so veridically under normal conditions, but when asked to perform a concurrent attentionally-demanding conjunction task, the mapping followed a compressive, non-linear function. We model the non-linearity both by the commonly assumed logarithmic transform, and also with a Bayesian model of central tendency. These results suggest that veridical representation numerosity requires attentional mechanisms.

1. Introduction

Most adult humans can estimate the numerosity of a group of items, as can infants (Xu, Spelke, & Goddard, 2005) – including newborns (Izard, Sann, Spelke, & Streri, 2009) – and many non-human animals, including primates, parrots and even fish (Agrillo, Dadda, Serena, & Bisazza, 2009; Gallistel & Gelman, 1992; Nieder, 2005; Pepperberg, 2006). Numerosity shares many properties with other perceptual attributes, such as obedience of Weber's law (Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004; Ross, 2003) and susceptibility to adaptation (Burr, Anobile, & Turi, 2011; Burr & Ross, 2008). Importantly, the ability to discriminate number, which improves during development (Halberda & Feigenson, 2008; Piazza et al., 2010), is strongly predictive of future mathematical ability (Halberda, Mazzocco, & Feigenson, 2008).

Number and space are intrinsically interconnected. Mapping of numbers onto space plays a fundamental role for many aspects of mathematics, including geometry, Cartesian coordinates and mapping real and complex numbers onto lines or planes (Butterworth, 1999; Dehaene, 1997). Recent work has shown that children's conceptions of how numbers map onto space shifts radically during the early school years (Booth & Siegler, 2006; Siegler & Booth, 2004; Siegler & Opfer, 2003). Kindergarten children can represent numbers in space in a non-random manner, but their representation is compressed, seemingly logarithmic (placing, for example, the number 10 near the midpoint of a 1–100 scale). The compressive non-linearity becomes progressively more linear over the first 3 or 4 years of schooling (Booth & Siegler, 2006; Siegler & Booth, 2004; Siegler & Opfer, 2003), leading some to suggest that the "native" system of representating numbers may be logarithmic, which becomes linearized by schooling (Berteletti, Lucangeli, Piazza, Dehaene, & Zorzi, 2010). Strong support for this idea comes from a recent study of the Mundurucu, an Amazonian indigenous group with a limited number lexicon and little or no formal training:

* Corresponding author at: Department of Psychology, University of Florence, Via S. Salvi 12, Florence, Italy.
E-mail address: dave@in.cnr.it (D.C. Burr).

0010-0277/$ - see front matter
doi:10.1016/j.cognition.2011.11.006

both adults and children of this tribe map numbers and numerical quantities onto space in a logarithmic fashion (Dehaene, Izard, Spelke, & Pica, 2008). This points to both genetic and cultural roots to numerical mapping: the ability to represent numbers in space appears to be innate, but formal mathematical training is required to refine the representation from logarithmic to linear. Interestingly, dyscalculic children from developed societies also show a more logarithmic representation of the numberline than controls (Ashkenazi & Henik, 2010; Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007; Geary, Hoard, Nugent, & Byrd-Craven, 2008).

However, the notion that number is encoded in a true logarithmic fashion has been challenged from many fronts (Gallistel & Gelman, 1992; Karolis, Iuculano, & Butterworth, 2011). Alternate explanations have also been put forward for the non-linearities in the numberline representation, such as proportion judgments relative to the ends and centers of the numberline (Barth & Paladino, 2011). Another possibility, which we advance in this study, is that the non-linearity is an example of the well known "central tendency of judgment", reported by Hollingworth back in 1910: "judgments of time, weight, force, brightness, extent of movement, length, area, size of angles all show the same tendency to gravitate toward a mean magnitude" (Hollingworth, 1910). This old concept has recently been relaunched in the context of Bayesian analysis to model interval reproduction judgments (Cicchini, Arrighi, Cecchetti, Giusti, & Burr, 2011; Jazayeri & Shadlen, 2010).

Attention has been shown to play an important role in number perception. Attentional-training (through video-game playing) increases the subitizing range (Green & Bavelier, 2003). Although subitizing has often been considered to be "pre-attentive", several studies have shown that it in fact highly attentional-dependent, suffering considerably when attention is diverted with dual-task or attentional-blink paradigms (Burr, Turi, & Anobile, 2010; Railo, Koivisto, Revonsuo, & Hannula, 2008; Vetter, Butterworth, & Bahrami, 2008; Xu & Liu, 2008). Under dual-task attentional conditions, number discrimination (measured by Weber fraction) in the subitizing range falls to the same level as the estimation range (Burr et al., 2010). Attention also affects adaptation to numerosity (Burr et al., 2011).

In this study we ask whether spatial mapping of numbers depends on attention. Adult observers positioned dot-stimuli on a numberline, with and without a concurrent demanding color-conjunction task. With the attention-demanding task, the spatial representation of number, linear under normal viewing, shows clear non-linear compression. One interpretation of the results is that the native system of number representation is logarithmic, even in typical adults with normal mathematical ability, and that the linearization of this representation requires attention. However, we also explore the possibility that the compressive non-linearity results from a central tendency like that described by Hollingworth (1910) for many sensory judgments, which we model quantitatively within the Bayesian context.

2. Methods

2.1. Participants

Four subjects with normal or corrected-to-normal vision participated in this study, one author and three naïve to the goals of the study. All subjects were graduate students, two with previous experience in numerosity judgment tasks (three female, one male; mean age 26).

2.2. Stimuli and procedure

The stimuli were generated and presented under Matlab 7.6 using PsychToolbox routines (Brainard, 1997). They were displayed in a dimly lit room on a 13-in. Macintosh monitor with 1440×900 resolution at 60 Hz refresh rate, mean luminance 60 cd/m^2, viewed binocularly from 57 cm. The stimulus sequence is illustrated in Fig. 1a. Each

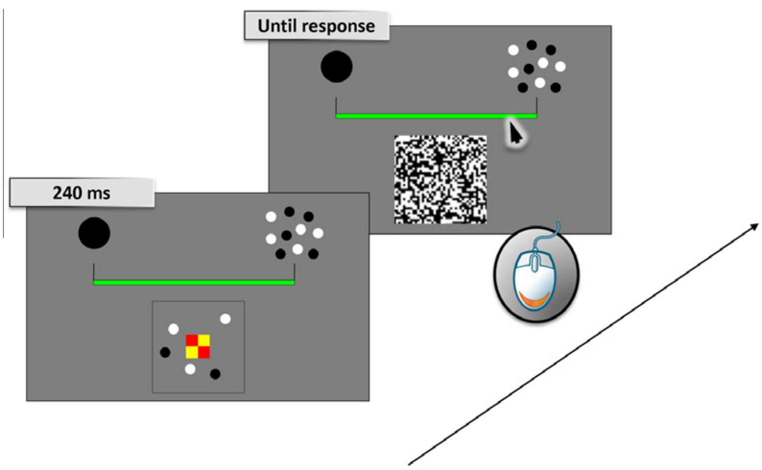

Fig. 1. Illustration of stimuli sequence. At onset of each trial observers view the number line, marked at each end with a single dot to the left and 10, 30 or 100 dots to the right. On key press, the dot stimulus appears, together with four colored squares in the center of the dot cloud. After 240 ms a binary pixel random-noise mask was displayed until subjects respond. Subjects respond first to the color conjunction task (in the dual-task condition, otherwise they ignore it), then mouse-click the numberline at the position they think the dot cloud should occupy.

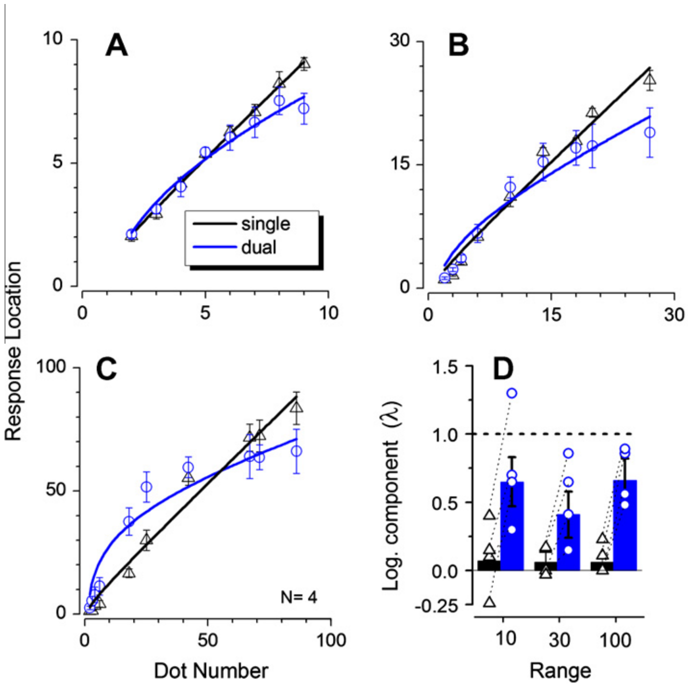

Fig. 2. (A–C) Average response location (pooling over all four subjects), plotted against actual dot number, for the three different ranges tested. Single-task judgments in black, dual-task in blue (error bars (±1 s.e.m.). The curves are best fits of Eq. (1). (D) Values of the logarithmic component (λ of Eq. (1)) of the best fits, for the averaged data of Figs. 2A–C as bars, and also for the individual subjects (connected symbols). The error bars indicate the standard errors of the fit to the averaged data.

trial started with subjects viewing a 20 cm "number-line" with sample dot-clouds representing the extremes: one dot on the left of the numberline, and either 10, 30 or 100 dots on the right. The numberline and samples remained on throughout the trial. On subject initiation, the stimuli were presented for 240 ms, followed by a random-noise mask that remained on until the subject responded. Two stimuli were presented simultaneously: a cloud of non-overlapping dots (the stimulus to be positioned on the numberline and four colored squares (the distractor stimulus). The dots were half-white, half-black at 90% contrast on a gray background, falling inside a circle of 8° diameter, and sparing the central 1°. The four distractor squares were positioned centrally (covering 1° of visual angle), and arranged in color combinations that either did or did not constitute a target (defined as both the top-right and bottom-left squares green, or both top-left and bottom-right squares yellow).

In all trials, subjects were required to position and click a mouse pointer on the position of the number line corresponding to the estimated numerosity. In the *dual-task* paradigm, subjects performed a color-orientation conjunction task on the central squares *before* making the number-line judgment, with a left mouse-click for a target, right-click for not. In the single-task condition subjects ignored the central task (that was always presented). Numberline data were recorded both for correct and incorrect responses for the distractor task, with errors running around 10% for all subjects. The average duration of each trial was 5.8 (±1.1) s for the dual-task condition, and 3.4 (±1.5) s in the single-task condition.

Each block measured one of the six conditions (three ranges, single and dual-task), presenting eight to ten test stimuli of different numerosity were presented in random order once. Three blocks were run for each condition, order randomized between observers. Following Siegler and Opfer (2003), the numerosities used for the three ranges were: 1–10: 2, 3, 4, 5, 6, 7, 8, 9; 1–30: 2, 3, 4, 6, 10, 14, 18, 20, 22; 1–100: 2, 3, 4, 5, 6, 18, 25, 42, 67, 71, 86. To discourage observers using strategies other than numerosity (such as texture density), on each block we kept constant either the *total covered area* (varying individual dot size) or individual dot size (varying total area covered), alternatively trial-by-trial. Thus on average, neither dot size nor total covered area correlated with numerosity.

2.3. Data analysis

We performed two types of analysis, a linear-logarithmic fit, and a Bayesian model of central tendency.

2.3.1. Linear-logarithmic fits

In this model we assume that the data can be described as the sum of a linear and a logarithmic component, given by the following equation:

$$L = a\left((1-\lambda)N + \lambda \frac{N_{max}}{\ln N_{max}} \ln N\right) \quad (1)$$

where N is the number of stimulus dots, L the average response location for that stimulus, N_{max} the set size (10, 30 or 100), and a a scaling factor. The main free parameter of the equation is λ, which determines the logarithmic component of the fit ($\lambda = 0$ defines a linear function, $\lambda = 1$ logarithmic). Both individual data and pooled data were well fit by this equation, with $0.84 \leqslant R^2 \leqslant 0.99$, and no tendency for the goodness of fit to vary with the degree of linearity.

2.3.2. Bayesian model

We also model the data with a Bayesian model, illustrated in Fig. 3. Bayes rule, in this context, is

$$p(L|N) \propto p(L)p(N|L) \quad (2)$$

where L represents the mapped location and N the actual number of dots. $p(L)$ is the *prior*, which in our model represents the central tendency, the *apriori* likelihood that the response is at a given point near line center. We model the prior with a Gaussian function, with center $\overline{P}$, and standard deviation σ_P determined by best fit of the data. The term $p(N|L)$ represents the sensory likelihood, which is set by the observer's precision. We assume that Weber's law applies over the range tested (Ross, 2003), so precision thresholds will be given by wN, where w is the Weber fraction (arbitrarily set to be 0.3). Again we model the sensory likelihood by a Gaussian probability density function, centered at the physical numerosity N with standard deviation wN. The *posterior* – $p(L|N)$ – will also be a Gaussian pdf,

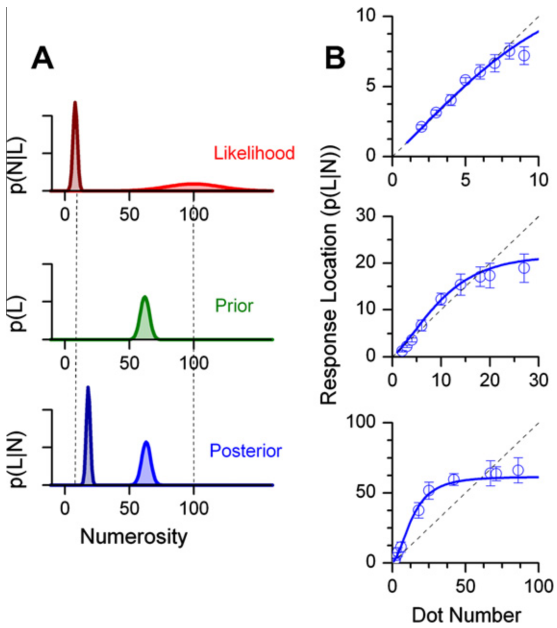

Fig. 3. Illustration of the central-tendency model of non-linear mapping. (A) Probability density functions for *likelihood*, *prior* and *posterior* (Eq. (2)), for two physical displays of 10 or 100 dots to be mapped onto a 1–100 numberline. For all three numberlines, the prior is a Gaussian pdf centered at 60% on the number line with standard deviation of 5 (determined by best fit to data). The likelihood was also Gaussian, centered at the physical number L, with standard deviation equal to wL, where w is the Weber fraction (set at 0.3). The posterior is the product of the sensory likelihood and the prior. (B) Data from Fig. 2, with the simulations shown by continuous curves. Again the fit to the data was good, with $0.95 \leqslant R^2 \leqslant 0.98$ (pooled data).

whose mean will be between the sensory estimate and the central prior. The extent to which the prior draws the results towards the mean depends on the relative widths of the prior and sensory likelihood functions. As the width of the sensory pdf is proportional to N, the effect will be stronger for large than for small numerosities, resulting in a compressive function.

There were only two free parameters to optimize, $\overline{P}$ and σ_P, the mean and standard deviation of the prior. The choice of the Weber fraction was somewhat arbitrary (but reasonable), but as it is the *ratio* of the Weber fraction to prior width that determines the effect, it makes no difference to the fit what value is chosen. The parameters that best fit the data were $\overline{P} = 60\%$ and $\sigma_P = 5$.

3. Results

Fig. 2A–C shows the number line judgments for the two attentional conditions, for the three different numerosities, averaged over all subjects. In all cases the numberlines for the single-task conditions (black symbols) are virtually linear, while those of the dual-task conditions (blue[1] symbols) show a clear compressive non-linearity. The lines passing through the symbols are best fits of the log-linear Eq. (1). Fig. 2D shows how the *logarithmic index* (λ) of these fits depends on attention for all three number ranges (bars the fits of the group data of Fig. 2A–C, linked symbols individual data). For every subject, the logarithmic component was far larger in the dual-task than the single-task condition. A two-way repeated measures ANOVA on the λ values revealed a significant main effect for attentional condition ($F = 47.8$, $p = 0.006$), but no effect for range ($F = 1.17$, $p = 0.387$) and no interaction ($F = 0.85$, $p = 0.47$). Post-hoc t-tests revealed that in none of the range conditions for the single-task λ was significantly different from 0 ($0.77 < t < 1.55$; $0.22 < p < 0.49$), showing that the linear model was adequate for the single-task conditions.

We also fit our data with an alternative model, the Bayesian model of central tendency (Eq. (2)), where the *posterior* probability of a particular localization response is given by the normalized product of the sensory data (the likelihood distribution resulting from a given number of dots) and the "central tendency" *prior*, which draws the sensory estimates towards the center. The blue curves at right show the Bayesian predictions for the three numberlines, clearly capturing the pattern of results ($0.95 < R^2 < 0.98$).

4. Discussion

This study shows that when adult humans are asked to position a cloud of dots on a number line, they normally do so accurately: that is, *linearly*. However, when attentional resources are diverted by a concurrent demanding conjunction task, the judgments become distinctly non-linear, well described by a logarithmic relationship.

One potential confound of our results is that they may reflect memory, as well as attentive, processes, as the aver-

[1] For interpretation of color in Fig. 2, the reader is referred to the web version of this article.

age time of response in the dual-task condition was nearly twice that of the single-task condition (5.8 compared with 3.4 s). It would be interesting to pursue which of these two factors is more important. Furthermore, our data do not allow us to distinguish whether the attentional (and memory) manipulation disrupts the *encoding* of non-symbolic quantity, or the *mapping process* itself – or both. However, our results make clear that when both encoding and mapping are required – as they are in this and previous studies – the non-linearity emerges very clearly. This is interesting, as it suggests that withdrawing attention may reveal a more *native* representation of numbers, one which prevails in young children (Berteletti et al., 2010; Siegler & Booth, 2004; Siegler & Opfer, 2003) and unschooled adults (Dehaene et al., 2008). Attention is necessary to map numbers linearly onto the number line. We believe that it is the attentional manipulation that affects the results, as previous work has demonstrated that linear responses occur even under conditions of speeded response (Pinhas & Fischer, 2008). However, other work has shown that humans adults, normally linear on a numberline task, show logarithmic-like non-linearities when asked to gauge the randomness of random a series that oversampled small numbers (Viarouge, Hubbard, Dehaene, & Sackur, 2010). This suggests that both linear and compressed maps can coexist, and the use of one or the other may be due to a variety of task-driven strategic factors.

What is the neural substrate underlying the logarithmic encoding and/or mapping? One possibility, assumed by many, is that it reflects the bandwidth of neurons selective to number. In both non-human and human primates, neural responses in the intraparietal sulcus show a logarithmic-like tuning, with bandwidth proportional to preferred number (Nieder, 2005; Nieder & Merten, 2007; Piazza et al., 2004), consistent with a pre-attentive logarithmic mapping onto the numberline. A native logarithmic representation is also consistent with the fact that numerosity discrimination in both monkeys and humans shows Weber-Law behavior, with the root-variance of the discrimination increasing directly with numerosity.

However, another possibility that we have explicitly modeled is that the non-linearity does not reflect logarithmic encoding of numerosities, but a more general perceptual principle, *central tendency*, observed in almost all sensory systems. That numerosity may be subject to the central tendency is further support for the notion of number being a visual sensory attribute (Burr & Ross, 2008; Dehaene, 1997). The intuitive explanation for how the model introduces the compressive non-linearity is that for constant Weber fraction, the sensory likelihood function at high numbers is broad, so the *prior* will be effective, while at low numerosities, the likelihood function is narrow, so the prior impacts less on it. This may also explain why numerosities over a larger scale tend to show more compression than over a smaller scale (Berteletti et al., 2010). The model has only two parameters, the mean of the distribution, which error minimization sets near the center of the numberline (excluding the small, easily recognizable numbers), and the width of the prior (relative to the Weber fraction). It has no more free parameters than the logarithmic model of Eq. (1), and the parameters remain fixed for all three numberlines: yet it accounts for 97% of the variance of all number lines.

What exactly is the prior, and how does it depend on attention? The prior could be computed from the sensory input, a "running average" calculated over trials, much in the same way as the system does in the "method of constant stimuli", where subjects compare quantities on individual trials against an estimate of the mean. Alternatively it may depend on the output numberline, which has finite length. For example, if a physical numerosity of 100 were perceived as 130 (through noise fluctuations), it could not be placed higher than 100 – and this effect could propagate down. We suspect that both these factors contribute to the prior.

What purpose does the prior – and central tendency in general – serve? As others have argued, a prior based on the statistics of the sensory events can improve performance – measured as the sum of total error – at the expense of reducing veridicality (see Jazayeri & Shadlen, 2010, for detailed account). Effectively, under conditions of great uncertainty, performance can be improved by considering the past history of events. But why does the regression to the mean occur only under conditions of divided attention? The most straight-forward possibility is that the sensory Weber fraction is higher under conditions of divided attention (Burr et al., 2010), so the prior becomes more effective (as it is the relative widths of prior and sensory likelihood that determines the extent of central tendency). Another possibility is that attention induces a more qualitative change in the processing, acting on the *prior* itself by reducing its width, and hence the mode of encoding and/or mapping numbers. We are currently investing these possibilities quantitatively.

To conclude, the current results show that attention can change the pattern of mapping numerosity estimates onto a numberline. The study provides support for the idea that mapping numbers onto space is a universal intuition, but that the native mapping principle is non-linear. The non-linearity could arise either from an intrinsic logarithmic representation of numbers, as many have assumed, or from a more general principle of central tendency of perceptual judgments. Either way it would seem that the linear numberline is a cultural invention, depending on formal education. But even after it has been instilled by years of schooling, it remains strongly dependent on the availability of attentional resources.

Acknowledgments

Supported by ERC Grant 229445 "STANIB", Grant PRIN09 from the Italian Ministry of Universities and Science and the Australian Research Council. Supported by Italian Space Agency Project CRUSOE.

References

Agrillo, C., Dadda, M., Serena, G., & Bisazza, A. (2009). Use of number by fish. *PLoS One, 4*(3), e4786.

Ashkenazi, S., & Henik, A. (2010). A disassociation between physical and mental number bisection in developmental dyscalculia. *Neuropsychologia, 48*(10), 2861–2868.

Barth, H. C., & Paladino, A. M. (2011). The development of numerical estimation: Evidence against a representational shift. *Developmental Science, 14*(1), 125–135.

Berteletti, I., Lucangeli, D., Piazza, M., Dehaene, S., & Zorzi, M. (2010). Numerical estimation in preschoolers. *Developmental Psychology, 46*(2), 545–551.

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 42*(1), 189–201.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433–436.

Burr, D. C., Anobile, G., & Turi, M. (2011). Adaptation affects both high and low (subitized) numbers under conditions of high attentional load. *Seeing and Perceiving, 24*, 141–150.

Burr, D., & Ross, J. (2008). A visual sense of number. *Current Biology, 18*(6), 425–428.

Burr, D. C., Turi, M., & Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *Journal of Vision, 10*(6), 20.

Butterworth, B. (1999). *The mathematical brain* (pp. xv, 446 p.). London: Macmillan.

Cicchini, G., Arrighi, R., Cecchetti, L., Giusti, M., & Burr, D. (2011). Optimal coding of interval timing in expert drummers, string musicians and non-musical control subjects. *Journal of Vision, 11*(8), 1212.

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

Dehaene, S., Izard, V., Spelke, E., & Pica, P. (2008). Log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures. *Science, 320*(5880), 1217–1220.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition, 44*(1–2), 43–74.

Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., & Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development, 78*(4), 1343–1359.

Geary, D. C., Hoard, M. K., Nugent, L., & Byrd-Craven, J. (2008). Development of number line representations in children with mathematical learning disability. *Developmental Neuropsychology, 33*(3), 277–299.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*(6939), 534–537.

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*(5), 1457–1465.

Halberda, J., Mazzocco, M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature, 455*(7213), 665–668.

Hollingworth, H. L. (1910). The central tendency of judgements. *Journal of Philosophy, Psychology and Scientific Methods, 7*, 461–469.

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences of the United States of America, 106*(25), 10382–10385.

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience, 13*(8), 1020–1026.

Karolis, V., Iuculano, T., & Butterworth, B. (2011). Mapping numerical magnitudes along the right lines: Differentiating between scale and bias. *Journal of Experimental Psychology: General, 140*(4), 693–706.

Nieder, A. (2005). Counting on neurons: The neurobiology of numerical competence. *Nature Reviews Neuroscience, 6*(3), 177–190.

Nieder, A., & Merten, K. (2007). A labeled-line code for small and large numerosities in the monkey prefrontal cortex. *Journal of Neuroscience, 27*(22), 5986–5993.

Pepperberg, I. M. (2006). Grey parrot numerical competence: A review. *Animal Cognition, 9*(4), 377–391.

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., et al. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition, 116*(1), 33–41.

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron, 44*(3), 547–555.

Pinhas, M., & Fischer, M. H. (2008). Mental movements without magnitude? A study of spatial biases in symbolic arithmetic. *Cognition, 109*(3), 408–415.

Railo, H., Koivisto, M., Revonsuo, A., & Hannula, M. M. (2008). The role of attention in subitizing. *Cognition, 107*(1), 82–104.

Ross, J. (2003). Visual discrimination of number without counting. *Perception, 32*(7), 867–870.

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development, 75*(2), 428–444.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14*(3), 237–243.

Vetter, P., Butterworth, B., & Bahrami, B. (2008). Modulating attentional load affects numerosity estimation: Evidence against a pre-attentive subitizing mechanism. *PLoS One, 3*(9), e3269.

Viarouge, A., Hubbard, E. M., Dehaene, S., & Sackur, J. (2010). Number line compression and the illusory perception of random numbers. *Experimental Psychology, 57*(6), 446–454.

Xu, X., & Liu, C. (2008). Can subitizing survive the attentional blink? An ERP study. *Neuroscience Letters, 440*(2), 140–144.

Xu, F., Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Developmental Science, 8*(1), 88–101.