

Moving along the number line: Operational momentum in nonsymbolic arithmetic

KOLEEN MCCRINK

Harvard University, Cambridge, Massachusetts

STANISLAS DEHAENE

*INSERM-CEA Cognitive Neuroimaging Unit, Orsay, France
and Collège de France, Paris, France*

AND

GHISLAINE DEHAENE-LAMBERTZ

INSERM-CEA Cognitive Neuroimaging Unit, Orsay, France

Can human adults perform arithmetic operations with large approximate numbers, and what effect, if any, does an internal spatial–numerical representation of numerical magnitude have on their responses? We conducted a psychophysical study in which subjects viewed several hundred short videos of sets of objects being added or subtracted from one another and judged whether the final numerosity was correct or incorrect. Over a wide range of possible outcomes, the subjects' responses peaked at the approximate location of the true numerical outcome and gradually tapered off as a function of the ratio of the true and proposed outcomes (Weber's law). Furthermore, an operational momentum effect was observed, whereby addition problems were overestimated and subtraction problems were underestimated. The results show that approximate arithmetic operates according to precise quantitative rules, perhaps analogous to those characterizing movement on an internal continuum.

Human adults possess an ability to estimate and manipulate approximate numerical magnitudes, which has been termed *number sense* (Dehaene, 1997). This ability appears to be largely independent of language and other symbol systems, since it is present in both infants (Xu & Spelke, 2000) and other animal species (Brannon & Roitman, 2003; Gallistel, 1990). In the present article, we present a new method that enables a direct, psychophysical evaluation of the precision of human numerical cognition in tasks of addition and subtraction of estimated sets comparable to those used in animals and in young children. The resulting *numerical psychophysics* paves the way to a deeper understanding of the mechanisms of approximate arithmetic.

Much of the research in numerical cognition started with psychophysical work in nonhuman animals (Meck & Church, 1983; Platt & Johnson, 1971). Platt and Johnson systematically varied the number of barpresses that rats must make in order to obtain a reward. By plotting the distribution of barpresses before the rats went to the feeder, they found that the number of presses formed a curve around the target amount. This indicated that the rats had an approximate representation of how many presses were required. This curve was relatively narrow for small magnitudes but wider for large magnitudes, indicating that the variability of the representation increased as the number of required presses increased. This scalar variability con-

forms to Weber's law, which states that the just-noticeable difference increases linearly with increases in the reference magnitude or, more generally, that the discriminability of two magnitudes depends only on their ratio (Izard, 2005). Scalar variability is a classic signature of the approximate magnitude representation of numbers, observed in a broad variety of tasks in human adults, infants, and animals (Feigenson, Dehaene, & Spelke, 2004; Gallistel, 1990).

In addition to the representation of approximate numerosity, there is behavioral evidence that both animals and infants can perform elementary arithmetic computations over these numerosities. Flombaum, Junge, and Hauser (2005) found that rhesus macaques can spontaneously add two quantities of foodstuffs, and Brannon, Wusthoff, Gallistel, and Gibbon (2001) found that pigeons can subtract one set of light flashes from another and, subsequently, peck the estimated number of flashes that are left over. Wynn (1992) found that very young infants are able to add and subtract small sets of objects, and McCrink and Wynn (2004) found that preverbal infants can discriminate correct and incorrect outcomes of large-number problem sets (such as $10 - 5$, or $5 + 5$).

This approach to studying the core quantity system of estimated magnitudes has also been applied to human adults. Whalen, Gallistel, and Gelman (1999) and Cordes, Gelman, Gallistel, and Whalen (2001) had adults tap out

K. McCrink, mccrink@wjh.harvard.edu

a target number of keypresses under conditions in which counting was suppressed and found that adults perform similarly to the rats in Platt and Johnson (1971). That is, human adults approximated the correct number, with a variability that increased linearly with increases in magnitude. Weber’s law also accounts for the distance effect observed when subjects compared the approximate numerosities of two sets of dots (van Oeffelen & Vos, 1982). Barth, Kanwisher, and Spelke (2003) demonstrated that these approximate numerical representations are not bound to a particular modality; adults who had to compare magnitudes across modalities (number of dots vs. number of tones) performed similarly when comparing two magnitudes within a particular modality (e.g., Dot Array A vs. Dot Array B). Piazza, Giacomini, Le Bihan, and Dehaene (2003) showed that numerosity is represented cortically within the left and right intraparietal sulcus, a cortical region involved in mental representation and manipulation of number in both human and nonhuman primates (see Dehaene, Piazza, Pinel, & Cohen, 2003, for a review).

Although there is clear evidence for a representation of numerical magnitude, research on how adults actually manipulate these magnitudes during mental arithmetic is sparse. There are demonstrations of addition with Arabic numerals (e.g., Ashcraft & Stazyk, 1981; Dehaene & Cohen, 1991; Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999) and with nonsymbolic numerosities, such as estimated arrays of objects or tones (Barth et al., 2006; Barth, LaMont, Lipton, & Spelke, 2005; Lemer, Dehaene, Spelke, & Cohen, 2003; Pica, Lemer, Izard, & Dehaene, 2004). In the most detailed studies to date, Barth et al. (2006) presented short movies in which sets of objects and/or tones were subtracted or added and then asked subjects whether a proposed outcome was greater or smaller than the actual result. Both adults and 5-year-old children performed above chance and improved as the ratio between the true and the proposed outcomes increased. However, such a distance effect might be due merely to a comparison of the proposed outcome with one of the operands, without any actual calculation (for instance, one may note that 40 is larger than 30 – 20 merely by noting that it is larger than the first operand). Only indirect arguments were presented against the use of such shortcuts. Previous studies also provided little or no evidence on the precision and bias with which subjects can perform approximate arithmetic and did not verify whether Weber’s law holds.

Here, we designed a systematic experiment to better probe how subjects perform approximate addition and subtraction tasks. For each addition and subtraction problem, across several hundred trials, we systematically presented a range of proposed outcomes. Crucially, we asked subjects whether each outcome seemed correct or not. Thus, we could verify that there is a well-defined range of numbers, close to the true outcome of the operation, that subjects judge as correct and that this range varies systematically with the operation being presented. This procedure enabled us to evaluate Weber’s law, which predicts that the range of outcomes that subjects judge as correct should increase with increases in the size of the true outcome and should depend on the ratio of the true and the proposed outcomes.

The present design also enabled us to precisely measure which quantity subjects preferentially expect for each arithmetic operation. In this way, we made a new observation: Outcomes to addition problems are systematically overestimated, whereas outcomes to subtraction problems are systematically underestimated. In the Discussion section, we will examine to what extent this *operational momentum* constitutes a numerical equivalent on a mental number line of the well-established phenomenon of *representational momentum*, which has been reported for a variety of internal continua for space, action, or the pitch of sounds (Ashida, 2004; Freyd & Finke, 1984; Freyd, Kelly, & DeKay, 1990; Getzmann, Lewald, & Guski, 2004).

METHOD

Subjects

The subjects were 12 college-age adults with a mean age of 24 years. They gave informed consent to participate and were paid €15.

Design and Stimuli

Each subject viewed 598 randomly presented short videos of math operations, half of which displayed addition problems and half of which displayed subtraction problems (see Table 1 for specific operand and outcome types). The trials of interest made up 90% of the study. The other 10% were distractors, which were added to ensure that the subjects could not respond simply on the basis of accepting 8 or 16 (the most common true outcomes from which the other outcomes radiated) as always correct. The problems that appeared in the main experimental trials were repeated eight times in total, and the distractor problems were repeated twice.

All displays at start contained a gray occluder with a dark red fixation cross in the center of the screen. In the addition videos, a set of objects (n_1) moved from offscreen left and moved toward a centered occluder, taking 1.5 sec to arrive behind the occluder. A second set of objects (n_2) moved from offscreen right to join the other set behind the occluder (also taking 1.5 sec). In the subtraction videos, the first set of objects (n_1) moved onscreen from the left and went behind the occluder, and a subset of these objects (n_2) smoothly

Table 1
Experimental Design

Operation Presented	Possible Outcomes (Approximate Ratio)						
	1/2.0	1/1.5	1/1.25	1.0	1*1.25	1*1.5	1*2.0
Main Experimental Trials (90% of Trials)							
6 + 2	4	5	6	8	10	12	16
6 – 2	2		3	4	5	6	8
12 + 4	8	11	13	16	20	24	32
12 – 4	4	5	6	8	10	12	16
24 + 8	16	21	26	32	40	48	
24 – 8	8	11	13	16	20	24	32
8 + 8	8	11	13	16	20	24	32
32 – 16	8	11	13	16	20	24	32
4 + 4	4	5	6	8	10	12	16
16 – 8	4	5	6	8	10	12	16
Distractor Trials (10% of Trials)							
4 + 20	12	16	19	24	30	36	48
30 – 10	10	13	16	20	25	30	40
7 + 5	6	8	10	12	15	18	24
20 – 15		3	4	5	6	8	10

Note—Subjects received 14 distinct arithmetic problems (left column), each paired with either the true outcome (1.0 column) or with an incorrect outcome (other columns). The proposed outcomes were set to be in an approximately fixed ratio with the true outcome. This ratio varied from one half of the true outcome to twice the true outcome.

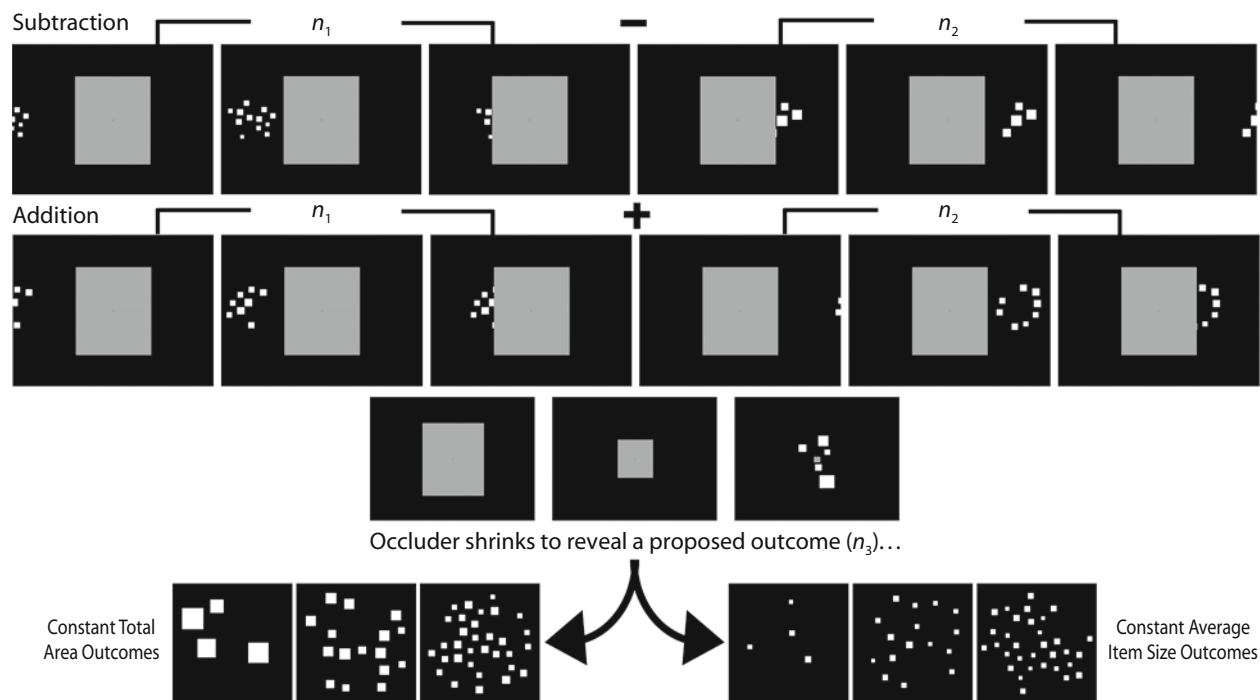


Figure 1. A schematic of addition and subtraction movies and the two test outcome types. A set of n_1 objects moves onto the screen, slows, and then accelerates behind the occluder. In the subtraction movies, a subset of n_2 objects smoothly continues to move from behind the occluder, slows, and then leaves offscreen. In the addition movies, a set of n_2 objects appears onscreen from the right-hand side and joins the first set of objects behind the occluder. In both cases, the occluder then shrinks and disappears to reveal a proposed outcome set (n_3), which can be either a constant-area outcome or a constant average item-size outcome (examples at bottom).

continued the trajectory and moved offscreen to the right. After the completion of the operation, the occluder then disappeared, revealing either an incorrect or a correct outcome (n_3). The fixation cross remained but lost visibility against the black background once the outcome was revealed (Figure 1). Each video took approximately 4 sec. A sample of these videos are available online at www.yale.edu/infantlab/mccrink_multimedia.html.

To prevent the use of nonnumerical cues, the sets of objects representing the numerical operands n_1 and n_2 were designed and generated using MATLAB, so that object size changed but the total summed area of the objects in a given set was always fixed. As a consequence, each addition problem yielded a total area behind the occluder of 11 cm², and each subtraction problem yielded a total area of 0 cm². For example, in the case of a 16 – 8 subtraction problem, an array of 16 squares with sum total area of 6 cm² would go behind the occluder, and an array of 8 squares that also had a total area of 6 cm² would exit out the other side. Thus, total area could not be used to respond differentially to problems with an outcome of, say, 4, 8, 16, or 32. As a result of this manipulation, average item size covaried inversely with numerosity during the presentation of the operands (i.e., sets with smaller numerosities had larger objects). However, it is not clear how this covariation within each set could be used as a cue to the operation result. To further ensure that the subjects could not rely on perceptual variables, two outcome types were created. Constant-area outcomes possessed an identical amount of total area (5.5 cm², the intermediate value between 0 and 11 cm²). If the subjects held expectations of occluded objects solely in terms of area, they would exhibit a flat performance curve on these trials, because total area never changed across the proposed outcomes. The second type of outcome, with constant average item size, consisted of items with an average size of 0.2 cm² per object. Again, if the subjects based their responses on the average item size of outcomes, they should fail to differentiate between these trials and should show

a flat performance curve. If these two types of outcomes led to identical performance curves, always centered at a similar location on the number line close to the true outcome, we could conclude that numerosity, which was the only common variable across both outcome types, was the basis for the subjects' judgments.

Procedure

The subjects were told to continuously fixate on the cross in the center of the screen and to observe the movies up until the presentation of the outcome, at which point the screen went blank and they signaled whether they thought the answer was correct or incorrect by pressing the “j” or “f” key, respectively. A blue circle appeared on the screen to indicate that their answer had been recorded and that they could press the space bar to advance to the next video. No other feedback was provided. The subjects were told to respond as quickly as possible but that it was more important to respond correctly than quickly. This served the purpose of ensuring that the subjects used a more *online* implicit process, rather than any sort of explicit heuristic or reasoning process.

RESULTS

Figure 2 displays the percentage of time the subjects perceived each revealed outcome as correct, for all addition and subtraction problems in the main design. For each operation, the results trace a well-defined unimodal curve with a clear peak at the approximate location of the true outcome. Thus, for each operation, the subjects appeared to accept as correct only a well-defined range of plausible answers. This finding clearly establishes a human competence for approximate nonsymbolic arithmetic.

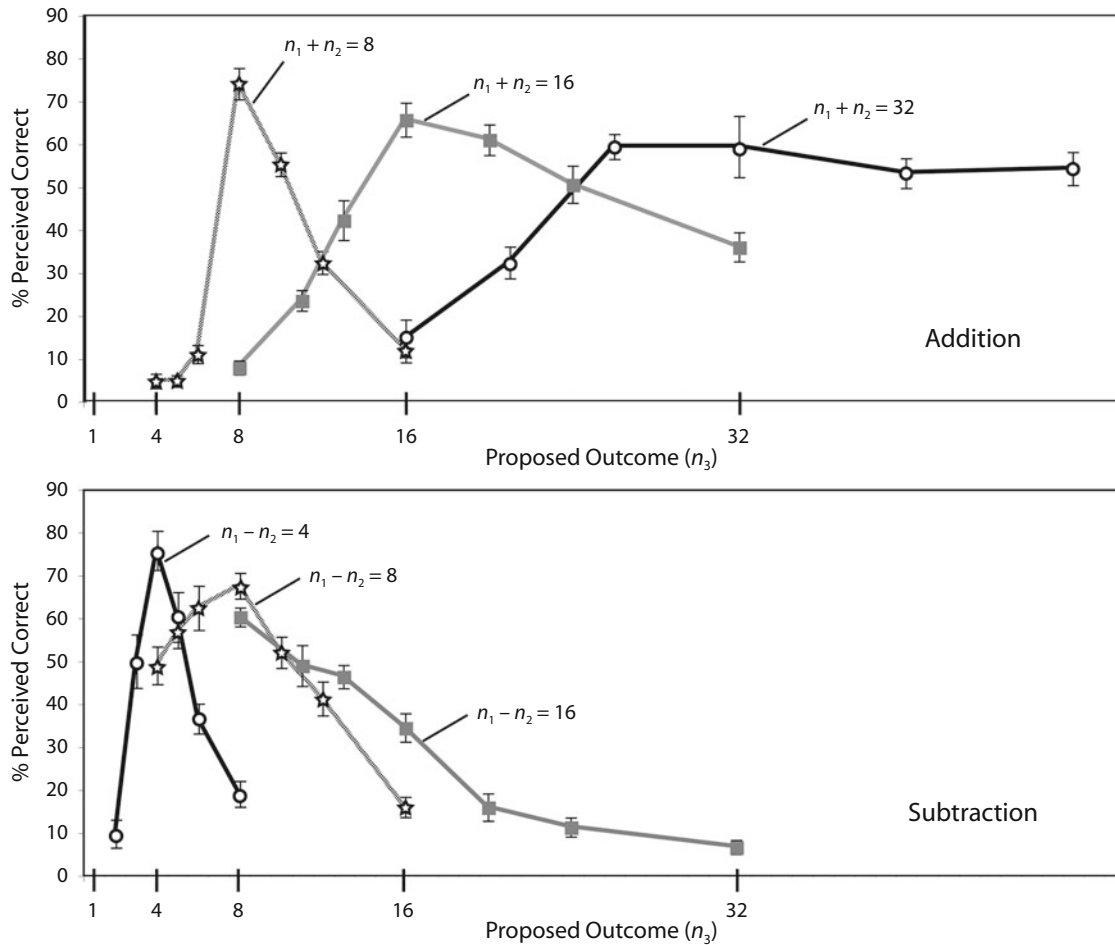


Figure 2. Performance curves for the main addition and subtraction problems. For each operation and each value of the true outcome, the figure shows the percentage of time that the subjects perceived the operation as correct as a function of the proposed outcome (n_3).

It is also clear from Figure 2 that the breadth of the curves increases as the magnitude of the true outcome increases. For example, in the addition case, the function for an outcome of 32 in the addition case is much wider than the function for an outcome of 8. However, as is illustrated in Figure 3, once plotted on a logarithmic scale, the curves for different outcomes become similar, even across addition and subtraction problems. Their similar width across the log scale indicates that performance was determined mainly by the ratio of the proposed and true outcomes, which is an indication of Weber’s law.

To evaluate the significance of these patterns, the data from nondistractor trials were analyzed with a $7 \times 5 \times 2$ ANOVA with subjects as a random factor and within-subjects factors of ratio of the true and proposed outcomes (approximately 1/2, 1/1.5, 1/1.25, 1, 1.25, 1.5, or 2.0), problem size (a five-level factor as defined by the sum of the two operands, which could be 8, 16, 24, 32, or 48), and operation type (addition or subtraction). Initially, a fourth factor of test outcome type (constant area or constant item size) was analyzed. However, there was no main effect of test outcome type [$F(1,11) = 1.53, p = .217$,

and examination of the response curves indicated essentially identical performance as a function of outcome type, confirming that the subjects based their responses on the common variable of number, rather than on any nonnumerical parameter such as density or size. This factor was thus dropped from further analyses.

The ANOVA revealed a highly significant main effect of ratio [$F(6,66) = 28.85, p < .001$], indicating that performance was driven mostly by the distance between the proposed outcome and the true outcome of the operation, this distance being measured by their ratio (or equivalently, by their difference on a logarithmic scale).

There was also a main effect of problem size [$F(4,44) = 13.09, p < .001$], suggesting that the subjects tended to more frequently respond *correct* to larger problems. An interaction between operation type and problem size [$F(2,22) = 9.41, p = .001$] showed that this bias was more pronounced for addition than for subtraction.

There were also interactions between ratio and problem size [$F(24,264) = 6.96, p < .001$] and between ratio and operation [$F(6,66) = 9.27, p < .001$], and a triple interaction between ratio, operation, and problem size [$F(10,110) =$

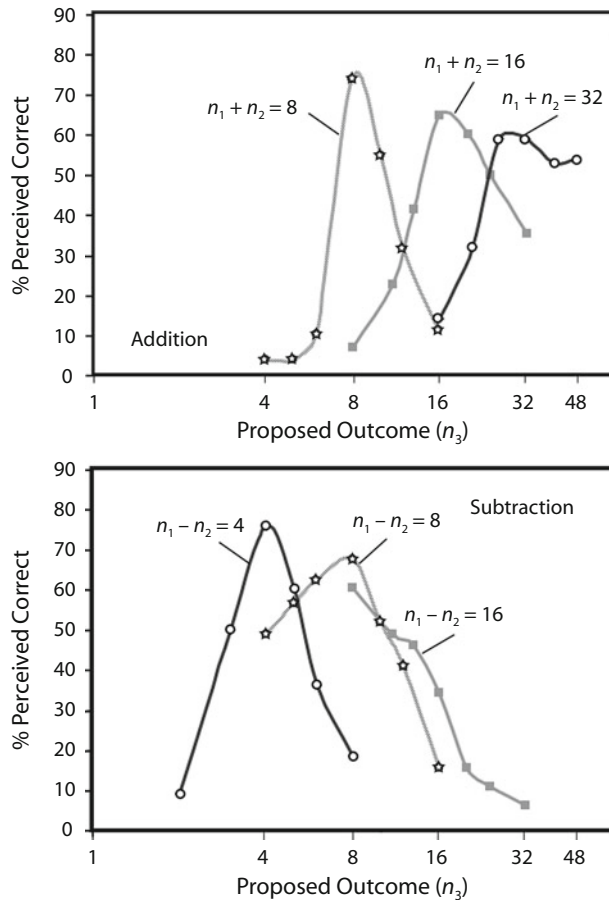


Figure 3. Performance curves for the main design problems plotted on a logarithmic scale. The figure shows the percentage of time that the subjects perceived the operation as correct as a function of the proposed outcome's (n_3) position on a logarithmic scale. A clear enhancement of symmetry of the response curves (a consequence of Weber's law) is evident in comparison with Figure 2.

6.58, $p < .001$]. These interactions reveal a novel effect that we have termed the *operational momentum* effect, which becomes particularly large as problem size increases. A close examination of Figure 2 shows that although the responses to small problems tend to be centered on the true outcome, as problem size increases they become increasingly "lopsided" and displaced in the direction of larger numbers for addition and of smaller numbers for subtraction. Figure 4, which summarizes performance as a function of ratio, depicts this momentum effect most clearly; the subjects tended to more readily judge as correct those outcomes that were biased in the direction of operation.

Quantitative Evaluation of the Operational Momentum Effect

For each problem, the location of the peak response indicates which outcome the subject considers correct for a given operation. Estimating this value would allow us to quantitatively evaluate the amount of bias in the subjects' responses and, therefore, allow determination of the magnitude of the momentum effect. We therefore designed a nonlinear curve-fitting analysis that allowed us to estimate

this parameter, which we termed the *estimated preferred outcome* (EPO). The procedure also yielded an estimate of the width of the curve, thus quantifying the uncertainty the subject has in making his judgment (hereafter, *uncertainty range*, or UR).

Our fitting procedure is based on the observation that, although they are asymmetrical on a linear scale, the performance curves become quasisymmetrical and Gaussian-shaped once plotted on a logarithmic scale (see Figure 3). Indeed, this is a classical finding in approximate numerical cognition (e.g., Nieder & Miller, 2003) and is a natural consequence of Weber's law (Izard, 2005). We therefore fitted the performance curves with a log-Gaussian curve with four free parameters: center, width, height, and offset from zero (corresponding to a small probability of responding at random on some trials):

$$p(x) = \alpha_1 + \alpha_2 G\left(\frac{\log(x) - \alpha_3}{\alpha_4}\right),$$

where the α_i are free parameters and G is the Gaussian probability density function

$$G(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}.$$

As is illustrated in Figure 5, the fits were generally very good (mean $r^2 = 93.3\%$, minimum = 86.2%). Very similar results were obtained when the more complicated function, introduced by Piazza et al. (2003) for *same-different* judgments, was used instead of the Gaussian curve. For easier interpretation, the parameters α_3 and α_4 , which are the mean and standard error of the Gaussian on the logarithmic scale, were then converted back to a linear scale according to the formulas

$$EPO = e^{\alpha_3}$$

and

$$UR = EPO \times e^{\alpha_4} - EPO \times e^{-\alpha_4}.$$

Table 2 shows the resultant EPOs and URs for each problem. Statistical analyses on these values provided several clues concerning the mechanisms of approximate arithmetic. First, the EPO correlated significantly with the true outcome, both overall ($r^2 = 75\%$, $p < .001$), within addition ($r^2 = 95\%$, $p < .001$), and within subtraction ($r^2 = 61\%$, $p < .05$). The overall regression slope of 1.09 ± 0.182 (1 SE), thus indistinguishable from a slope of 1, indicates that the subjects' estimates of each outcome were nearly equal to the true outcomes on average, thus confirming the subjects' overall competence for approximate arithmetic.

However, when the same regression was performed separately for addition and for subtraction, the slope was above 1 (1.17 ± 0.12) for addition and below 1 (0.385 ± 0.14) for subtraction. (A slope value above 1.0 indicates a shift of the preferred outcome toward values larger than the true outcome, and a slope value of less than 1.0 indicates a shift toward values smaller than the true outcome.) As is shown in Figure 6, all EPOs were larger than the true outcome for addition, and all but one were smaller than

Table 2
Estimated Preferred Outcomes, Percentages of Deviation Relative to the True Outcome, and Range of Correct Responses for Each Problem

n_1	Operation Sign	n_2	True Outcome	Estimated Preferred Outcome	Percent Deviation	Size of Acceptable Range
6	+	2	8	9.0	13	4.0
12	+	4	16	17.0	6	9.6
24	+	8	32	34.7	9	41.7
8	+	8	16	20.7	29	17.5
4	+	4	8	8.5	7	2.9
6	-	2	4	4.1	1	2.4
12	-	4	8	7.4	-7	7.9
24	-	8	16	9.6	-40	8.2
32	-	16	16	4.9	-69	9.6
16	-	8	8	5.8	-27	8.8

Note—See Figure 4 and the text for an explanation of how these parameters were derived. The momentum effect appears as a consistently negative deviation for subtraction problems and as a consistently positive deviation for addition problems.

the true outcome for subtraction. This finding provides another clear indication of operational momentum. To demonstrate this statistically, we computed the percentage of response deviation calculated as the signed difference between the EPO and the true outcome, expressed as a percentage of the true outcome (Table 2). Across problems, this value was significantly positive for addition [$M = 15.1\%$; $t(6) = 3.61, p = .01$] and significantly negative for subtraction [$M = -30.6\%$; $t(6) = -3.48, p = .01$], thus revealing a significant difference between addition and subtraction.

We also used regressions to evaluate which parameters of the proposed arithmetic operation, if any, modulated the size of the momentum effect. A multiple regression performed on the absolute percentage of deviation, with independent factors of true outcome, n_1 , and n_2 , revealed significant positive correlations with both n_1 ($r^2 = 46\%$, $p <$

.01; slope, 1.00 ± 0.33) and n_2 ($r^2 = 55\%$, $p < .01$; slope, 2.19 ± 0.60), but not with true outcome ($r^2 = 6\%$, $p = .41$). Note that the slope of increase with n_2 is about twice larger than the slope of increase with n_1 , although this difference does not reach significance. A subsequent stepwise regression examining the separate contributions of n_1 and n_2 to percentage of deviation revealed that a model incorporating n_2 accounted for a significant amount of variance above and beyond a model with n_1 only (an r^2 change of 27.1%, $p = .006$). A complementary analysis of the effect of n_1 beyond n_2 revealed an additional effect of n_1 as well, but to a lesser extent (an r^2 change of 18.1%, $p = .02$). Thus, although both operands have some influence on the amount of deviance from the actual outcome, the second operand may have a slightly greater effect than the first in determining the amount of operational momentum. This issue cannot be resolved with the present data, however, because the present stimulus set was not designed to systematically vary n_1 and n_2 orthogonally. Note in particular that, in all main experimental trials, the second operand was always smaller than or equal to the first, an undesirable feature for comparing the size of their impact on the momentum effect. Additional experiments will be needed to resolve this point.

Determinants of the Precision of Approximate Arithmetic

We now turn to the determinants of the UR. For which problems were the subjects least precise in their approximations? If one assumes that the operands n_1 and n_2 are encoded according to Weber’s law, the final uncertainty can be calculated from the variance associated with each of the operands. The basic theory underlying this process was developed by the second author (S.D.) and was published in Appendix B of Barth et al. (2006; for a similar approach, see Cordes, Gallistel, Gelman, & Latham, 2003). The simplest model predicts that, for both addition and subtraction, the variance associated with the result of an operation should

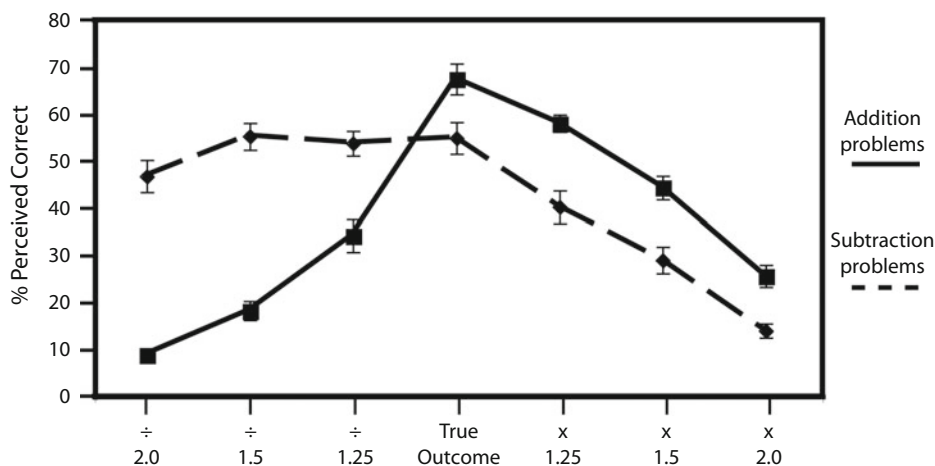


Figure 4. Performance curves for addition and subtraction collapsed across all experimental problems. The percentage of operations judged correct is plotted as a function of the approximate Weber ratio, which is a measure of the numerical distance from the proposed outcome to the true outcome. The momentum effect is evident; subjects are more likely to judge as correct the overestimated addition problems and the underestimated subtraction problems.

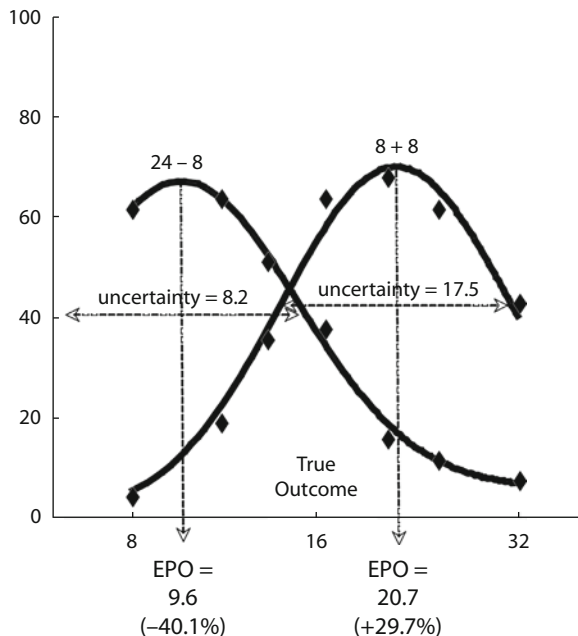


Figure 5. Illustration of the computation of the estimated preferred outcomes (EPOs) for two sample addition and subtraction problems with a true outcome of 16. The percentage of operations judged correct is plotted on a logarithmic scale (diamonds = actual data points) and is fitted by a Gaussian curve (thick line = best fit). This Gaussian fit defines both a central tendency (EPO) and an uncertainty range (the range of values with ± 1 standard deviation).

be equal to the sum of the variances associated with each of the operands. This would imply that the UR is proportional to the square root of the sum of squares of the operands n_1 and n_2 . We tested this model by linear regression of the UR on this variable, across all problems listed in Table 1.

For addition, this model provided an excellent fit ($r^2 = 92.9\%$, $p = .0005$). However, the slope of 1.82 was significantly different from 1, suggesting a higher variance than that due to the operands alone. This result is not unreasonable, since some additional variance might be expected to arise from the computation of the operation. Barth et al. (2006) proposed that this term might be proportional to the size of the operation outcome. From this hypothesis, they found its weight to be 1.3 times larger than the operand-induced variance. Once we used these predetermined values, the corresponding predictor provided a very tight fit ($r^2 = 95.4\%$, $p = .0002$), with a slope of 1.01.

For subtraction, however, both the simple model and the more complex model failed to capture significant variance in the UR. Only a nonsignificant trend was observed (respectively $r^2 = 43.7\%$, $p = .11$, slope = 0.15; and $r^2 = 38.4\%$, $p = .14$, slope = 0.12). Although the regression might become significant if a greater range of subtraction problems was tested, what is most striking is the very low slope, which indicates that the variance in the subtraction outcome is much lower than would be expected on the basis of uncertainty associated with the operands. For instance, for a problem such as $32 - 16$, the large size of the operands should make it very difficult for subjects to have any confidence in the outcome. On the contrary, an

inspection of Table 2 shows that the UR for $32 - 16$ is identical to that for $12 + 4$.

What could be the explanation for the high precision achieved by the subjects in the subtraction task? One key assumption of the variance estimates above is that the two operands (e.g., 32 and 16) are encoded by independent random variables, so that their variances add. This assumption might be violated in our paradigm; there might be a significant covariance between those two estimates, in which case the variance of $n_1 - n_2$ would be much lower than the variance of n_1 plus the variance of n_2 (see Cordes et al., 2003, for a similar assumption). The high covariance would mean that subjects do not separately estimate the two operands. This hypothesis seems particularly plausible given that our movies gave the impression that the first operand passed behind the screen and that some of its objects continued their trajectory as the second operand. Note that such a covariance term, in addition problems, would actually *increase* the uncertainty range and, thus, might provide an alternative explanation for the high addition variance described above.

A second explanation for the higher precision in subtraction is the apparent closeness of the *zero boundary*, which could have resulted in a floor effect, thus reducing variance. For the addition problems, the response space is essentially open ended (toward infinity). For the subtraction problems, however, there is an obvious lower bound of zero, which may have led to decreased variance. Again, new experiments with a more systematic variation of n_1 , n_2 , and outcomes will be needed to disentangle the sources of noise of nonverbal approximate calculation.

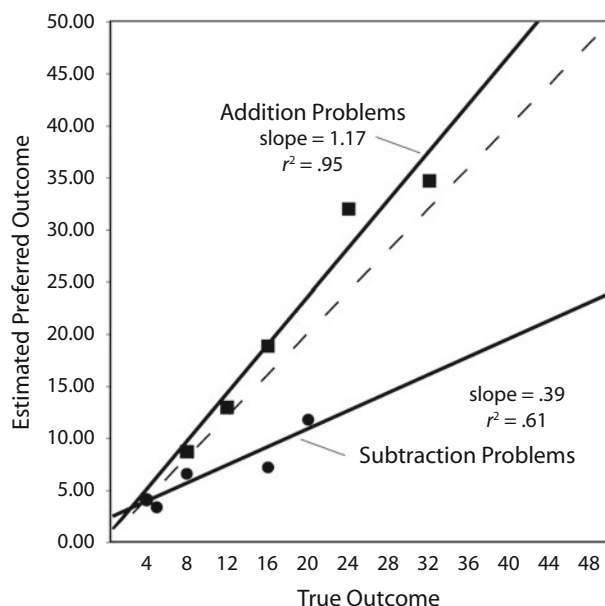


Figure 6. Linear relation between the true outcome and the subjects' estimated preferred outcome. The distinct linear regression lines for addition and subtraction problems suggest that subjects can estimate the outcome of nonsymbolic arithmetic operations, yet do so with a bias toward larger numbers for addition and toward smaller numbers for subtraction.

DISCUSSION

Our experiments afford two main conclusions about approximate nonsymbolic arithmetic by human adults: one concerning the mean performance (what average value do subjects consider to be a correct answer to an arithmetic problem?), and the other concerning the variability of responses (what is the range of the values considered correct?). Concerning mean performance, we discovered that the mechanism of human approximate arithmetic produces what we term *operational momentum*: a systematic bias toward larger values for addition problems and smaller values for subtraction problems. Concerning variability, we found that human adults can add and subtract numerical magnitudes via an estimation system that conforms to Weber's law: Variability increases in proportion to problem size.

Our results confirm and extend previous studies of nonverbal numerical operations (Barth et al., 2006; Cordes et al., 2003; Whalen et al., 1999). By giving the subjects in the present study two basic operations and a full range of operands and outcomes, we were able to precisely quantify approximation behavior and to observe interactions between operation and outcome. One result, which mirrors the findings in animals and infants, is that exact verbal representations of number are not necessary in order to perform arithmetic operations. As was predicted, the subjects who were forced to estimate the outcome of the addition or subtraction of two sets of objects were able to do so but represented this outcome in a "fuzzy" manner. They were much more likely to respond that an answer was correct as the ratio between the presented outcome and the correct outcome became closer to 1.0. The subjects also exhibited greater variability in their responses as the true outcome grew. This scalar variability is a reflection of Weber's law for numbers and is a hallmark of the magnitude estimation system frequently found in nonhuman animals (Brannon et al., 2001; Platt & Johnson, 1971) and preverbal human infants (Lipton & Spelke, 2003).

A novel finding of the present study is that approximate mental arithmetic is subject to an *operational momentum*; as numbers were mentally added or subtracted, the subjects seemed to be carried further along than necessary on their number line, toward large numbers for addition problems and toward small numbers for subtraction problems. For example, when given the problem of $24 + 8$, they were much more likely to judge 40 as correct than to judge 26 as correct, despite the fact that both of these incorrect outcomes differed from the correct outcome (32) by about 25%.

The Possible Role of Arithmetic Heuristics

Before considering high-level representational accounts of this effect, it is important to eliminate possible confounds. Could the effect, for instance, be due to a simple size heuristic according to which, in an addition situation, one should expect the outcome to be greater than the first operand and, in a subtraction situation, one should expect an outcome that is less than the first operand? This heuristic predicts that, for addition problems $n_1 + n_2 = n_3$, any n_3 that is smaller than either n_1 or n_2 can be quickly rejected, as with subtraction problems $n_1 - n_2 = n_3$ in

which n_3 is larger than n_1 . Together, these shortcuts would create an asymmetry in subjects' correctness judgments.

The results shown in Figure 4, however, suggest that this hypothesis does not suffice to explain the present results. For the central three ratios, 1/1.25, 1.0, and 1.25, given the problems that we used, the strategy above does not apply (see Table 1). Nevertheless, an ANOVA restricted to the average performance of subjects for these trial types revealed an interaction of operation and Weber ratio [$F(2,22) = 22.34, p < .001$], indicating the presence of the operational momentum effect even within this restricted subset of problems. For example, in response to $8 + 8$, the subjects preferred 20 to 13, although both outcomes are larger than 8 (and are approximately equidistant from the true outcome of 16). Likewise, the subjects presented with $16 - 8$ preferred 6 as a correct answer to 10, although both of these outcomes are smaller than 16 (and equidistant from the true outcome of 8). Thus, although the size heuristic may contribute to enhancing the difference between the response curves for addition and for subtraction, it cannot fully explain the momentum effect.

The Momentum Interpretation

What might be the mechanisms of approximate arithmetic that yield this momentum effect? We envisage at least two nonexclusive interpretations. One proposal is that the operational momentum in approximate arithmetic is a numerical version of the broader phenomenon of representational momentum (for a thoughtful review, see T. L. Hubbard, 2005). In a classic example of representational momentum (Freyd & Finke, 1984), subjects saw a shape at each of three successive orientations. Their memory of the final position of the shape was then biased in the direction of implied motion. The parallels to the present study are obvious; one can readily conceive of our experiment as a numerical test of representational momentum along an internal number line, with the extrapolated direction being movement along the number line, instead of rotations through particular spatial orientations. Much in the same way that moving objects have physical momentum, our mental calculations of addition and subtraction may themselves have momentum. It should be noted that representational momentum has been previously observed for the nonspatial continuum of pitch of tones (Freyd et al., 1990), thus confirming that it may extend to abstract concepts of represented motion.

The metaphor of approximate addition and subtraction as a displacement on the number line has been proposed by several authors, including Restle (1970), Dehaene and Cohen (1991), Gallistel and Gelman (1992), and, most recently, E. M. Hubbard, Piazza, Pinel, and Dehaene (2005). It meshes well with the metaphor of a mental number line, which has been found useful to account for many behavioral *number sense* tasks (E. M. Hubbard et al., 2005). A strong relation between representations of number and of space in normal subjects was first revealed by the SNARC effect (spatial-numerical association of response codes; Dehaene, Bossini, & Giroux, 1993), in which subjects are faster to respond to small numbers with their left hand and to larger numbers with their right hand. This number-space interaction has been found in other tasks, such as a

probe detection in which an irrelevant number was found to direct attention to the left side of the display if it was small and to the right side if it was large (Fischer, Castel, Dodd, & Pratt, 2003). Neuropsychological evidence for the number line comes from patients who suffer from spatial neglect. When asked to state the middle number between two values, these patients respond with numbers that deviate toward the nonneglected side (Zorzi, Priftis, & Umiltà, 2002). Brain-imaging evidence also suggests the involvement of parietal lobe spatial and attentional areas during mental calculation (Dehaene et al., 2003), particularly during approximation (Dehaene et al., 1999).

This body of evidence led E. M. Hubbard et al. (2005) to infer that the computational mechanisms that subvert movements of attention and updating of parietal maps in adults may be crucial for arithmetic operations as well. During these operations, the locus of activation would be shifted along the number line, in the direction of the operation (toward larger numbers for addition and smaller numbers for subtraction). Thus, E. M. Hubbard et al. actually predicted that a momentum should be found in mental arithmetic; this parietal remapping mechanism would lead to the side effect of “pushing” the participants’ estimates further along the number line. Note, however, that this interpretation still constitutes little more than a metaphor, since there is no physical movement or inertia in parietal attention.

Freyd (1987) speculated that representational momentum should occur when the underlying stimulus dimension is continuous and not when the underlying dimension is discrete, regardless of whether the specific stimulus happens to be presented in a discrete (implied motion) or continuous (smooth motion) format. Given that we used the (discrete) set of integers as stimuli, our results, if interpreted as a form of representational momentum, may seem to contradict Freyd’s hypothesis. However, a key aspect of current theories of numerical cognition, in both humans and animals, is that integers, with the possible exception of the very small numerosities 1, 2, and 3, are thought to be represented internally by approximate and continuous *scalar* quantities (Dehaene, 1992; Gallistel & Gelman, 1992). Indeed, the hypothesis that the number line forms a continuum is crucial to current mathematical models that treat approximate numerical decisions within the same framework as any other psychophysical decision (e.g., Dehaene, in press). In this context, the finding of a momentum effect with approximate numerosities is not incompatible with Freyd’s hypothesis. In fact, unpublished data from our laboratory suggest that the momentum effect can be found, although in much smaller size, when the operands are presented as symbolic numerals (e.g., Arabic digits; Knops, Viarouge, & Dehaene, 2007). Even this effect, however, would not contradict Freyd, since it is believed that Arabic numerals, during approximation tasks, are quickly converted to the same number line continuum used for representing approximate numerosity.

The Compression Interpretation

An entirely different explanation of the observed momentum effect can also be envisaged. Operational momentum might be a consequence of the compression and

expansion necessary to add or subtract on a compressed number line. Although this remains heavily debated, evidence suggests that on the internal number line, more representational space is dedicated to small numbers than to larger numbers (see, e.g., Dehaene, 2001, 2003; Izard, 2005; Nieder & Miller, 2003; Siegler & Opfer, 2003). In other words, Fechner’s law would hold for the numerical continuum: The internal scale $s(n)$, in the Thurstonian sense, would be a logarithmic or power function, rather than a linear scale [where $s(n) = n$]. Addition by juxtaposition of segments on such a compressed number line might then lead to a result shifted toward larger numbers, and subtraction to a result shifted toward smaller numbers. To see this most clearly, imagine that $s(n) = \log(n)$ and that subjects mistakenly add the internal representatives of n_1 and n_2 —that is, their logarithms. They would then reach a point on the number line at coordinate $\log(n_1) + \log(n_2)$, which corresponds to $\log(n_1 \times n_2)$ —the representation of the product of the operands, rather than their sum (i.e., a considerable overestimation of the addition result). Subtraction would similarly be replaced by division, this time leading to a considerable underestimation.

We are simply using a pedagogical example and are not suggesting that errors as extreme as this actually occur (Figure 2 clearly shows that they do not). It is likely that, in the course of the operation, the relevant neuronal network that implements the addition or subtraction process can first “undo” the internal compression of the operands, thus avoiding gross inaccuracies (see Dehaene, 2001). Nevertheless, if this internal *decompression* is inaccurate, a small compressive bias might persist, thus causing the observed momentum effect. One interesting possibility is that the first operand sets the start location $s(n_1)$ on the internal number line, whereas the second operand n_2 first has to be scaled according to the local amount of compression present at that location before being added or subtracted. According to this model, only the second operand would be transformed, and any bias in this transformation would yield a momentum effect determined mostly by the size of n_2 .

Clearly, such a compressive account of the momentum effect must remain highly speculative at present. A more extensive quantitative study of operational momentum will be needed to probe whether such a mechanism or its many possible variants can fit the data and what mathematical rules can be proposed for over- and underestimation effects in approximate calculation.

CONCLUSION

The present research supports theories that hypothesize that organisms have a magnitude estimation system that gives rise to approximate representations of numbers. These representations can be used in arithmetic operations (such as addition and subtraction), and the outcomes of these operations are approximate as well, exhibiting discrimination functions and variability in accordance with Weber’s law. The findings from the present study offer further support to the view that numerical operations are analogous to movement on an internal continuum, since

both seem to operate according to a momentum law. By dissecting the mechanisms of numerical estimation and mental arithmetic, we ultimately should gain greater insight into the intricate systems that underlie computation in both of these evolutionarily ancient cognitive domains.

AUTHOR NOTE

Correspondence concerning this article should be addressed to K. McCrink, Department of Psychology, Harvard University, 33 Kirkland St., Cambridge, MA 02138 (e-mail: mccrink@wjh.harvard.edu).

REFERENCES

- ASHCRAFT, M. H., & STAZYK, E. H. (1981). Mental addition: A test of three verification models. *Memory & Cognition*, *9*, 185-196.
- ASHIDA, H. (2004). Action-specific extrapolation of target motion in human visual system. *Neuropsychologia*, *42*, 1515-1524.
- BARTH, H., KANWISHER, N., & SPELKE, E. (2003). The construction of large number representations in adults. *Cognition*, *86*, 201-221.
- BARTH, H., LA MONT, K., LIPTON, J., DEHAENE, S., KANWISHER, N., & SPELKE, E. (2006). Non-symbolic arithmetic in adults and young children. *Cognition*, *98*, 199-222.
- BARTH, H., LA MONT, K., LIPTON, J., & SPELKE, E. (2005). Abstract number and arithmetic in preschool children. *Proceedings of the National Academy of Sciences*, *102*, 14116-14121.
- BRANNON, E., & ROITMAN, J. (2003). Nonverbal representations of time and number in animals and human infants. In W. Meck (Ed.), *Functional and neural mechanisms of interval timing* (pp. 143-182). Boca Raton, FL: CRC Press.
- BRANNON, E., WUSTHOFF, C., GALLISTEL, C. R., & GIBBON, J. (2001). Numerical subtraction in the pigeon: Evidence for a linear subjective number scale. *Psychological Science*, *12*, 238-243.
- CORDES, S., GALLISTEL, C. R., GELMAN, R., & LATHAM, P. (2003, November). *Nonverbal arithmetic in humans*. Poster presented at the annual meeting of OPAM (Object Perception, Attention, and Memory), Vancouver, BC.
- CORDES, S., GELMAN, R., GALLISTEL, C. R., & WHALEN, J. (2001). Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychonomic Bulletin & Review*, *8*, 698-707.
- DEHAENE, S. (1992). Varieties of numerical abilities. *Cognition*, *44*, 1-42.
- DEHAENE, S. (1997). *The number sense: How the mind creates mathematics*. New York: Oxford University Press.
- DEHAENE, S. (2001). Subtracting pigeons: Logarithmic or linear? *Psychological Science*, *12*, 244-246.
- DEHAENE, S. (2003). The neural basis of the Weber-Fechner law: A logarithmic mental number line. *Trends in Cognitive Sciences*, *7*, 145-147.
- DEHAENE, S. (in press). Symbols and quantities in parietal cortex: Elements of a mathematical theory of number representation and manipulation. In P. Haggard, Y. Rossetti, & M. Kawato (Eds.), *Attention and performance XXII: Sensorimotor foundations of higher cognition*. New York: Oxford University Press.
- DEHAENE, S., BOSSINI, S., & GIRAUX, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, *122*, 371-396.
- DEHAENE, S., & COHEN, L. (1991). Two mental calculation systems: A case study of severe acalculia with preserved approximation. *Neuropsychologia*, *29*, 1045-1074.
- DEHAENE, S., PIAZZA, M., PINEL, P., & COHEN, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, *20*, 487-506.
- DEHAENE, S., SPELKE, E., PINEL, P., STANESCU, R., & TSIVKIN, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science*, *284*, 970-974.
- FEIGENSON, L., DEHAENE, S., & SPELKE, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307-314.
- FISCHER, M. H., CASTEL, A. D., DODD, M. D., & PRATT, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nature Neuroscience*, *6*, 555-556.
- FLOMBAUM, J. I., JUNGE, J. A., & HAUSER, M. D. (2005). Rhesus monkeys (*Macaca mulatta*) spontaneously compute addition operations over large numbers. *Cognition*, *97*, 315-325.
- FREYD, J. J. (1987). Dynamic mental representations. *Psychological Review*, *94*, 421-438.
- FREYD, J. J., & FINKE, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *10*, 126-132.
- FREYD, J. J., KELLY, M. H., & DEKAY, M. L. (1990). Representational momentum in memory for pitch. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *16*, 1107-1117.
- GALLISTEL, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.
- GALLISTEL, C. R., & GELMAN, R. (1992). Preverbal and verbal counting and computation. *Cognition*, *44*, 43-74.
- GETZMANN, S., LEWALD, J., & GUSKI, R. (2004). Representational momentum in spatial hearing. *Perception*, *33*, 591-599.
- HUBBARD, E. M., PIAZZA, M., PINEL, P., & DEHAENE, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience*, *6*, 435-448.
- HUBBARD, T. L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin & Review*, *12*, 822-851.
- IZARD, V. (2005). *Interactions between verbal and non-verbal numerical representations: Theoretical and empirical approaches*. Unpublished doctoral thesis, University of Paris VI.
- KNOPS, A., VIAROUGE, A., & DEHAENE, S. (2007). *Dynamic representations underlying symbolic and nonsymbolic calculation: Evidence from the operational momentum effect*. Manuscript submitted for publication.
- LEMER, C., DEHAENE, S., SPELKE, E., & COHEN, L. (2003). Approximate quantities and exact number words: Dissociable systems. *Neuropsychologia*, *41*, 1952-1958.
- LIPTON, J. S., & SPELKE, E. S. (2003). Origins of number sense: Large-number discrimination in human infants. *Psychological Science*, *14*, 396-401.
- MCCRINK, K., & WYNN, K. (2004). Large-number addition and subtraction by 9-month-old infants. *Psychological Science*, *15*, 776-781.
- MECK, W. H., & CHURCH, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*, 320-334.
- NIEDER, A., & MILLER, E. K. (2003). Coding of cognitive magnitude: Compressed scaling of numerical information in the primate prefrontal cortex. *Neuron*, *37*, 149-157.
- PIAZZA, M., GIACOMINI, E., LE BIHAN, D., & DEHAENE, S. (2003). Single-trial classification of parallel pre-attentive and serial attentive processes using functional magnetic resonance imaging. *Proceedings of the Royal Society B*, *270*, 1237-1245.
- PICA, P., LEMER, C., IZARD, V., & DEHAENE, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*, 499-503.
- PLATT, J. R., & JOHNSON, D. M. (1971). Localization of position within a homogenous behavior chain: Effects of error contingencies. *Learning & Motivation*, *2*, 386-414.
- RESTLE, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology*, *83*, 274-278.
- STIEGLER, R. S., & OPFER, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, *14*, 237-243.
- VAN OEFFELEN, M. P., & VOS, P. G. (1982). A probabilistic model for the discrimination of visual number. *Perception & Psychophysics*, *32*, 163-170.
- WHALEN, J., GALLISTEL, C. R., & GELMAN, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science*, *10*, 130-137.
- WYNN, K. (1992). Addition and subtraction by 5-month-old human infants. *Nature*, *348*, 749-750.
- XU, F., & SPELKE, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*, B1-B11.
- ZORZI, M., PRIFTIS, K., & UMLTÀ, C. (2002). Neglect disrupts the mental number line. *Nature*, *417*, 138-139.