Exact and approximate arithmetic in an Amazonian indigene group with a reduced number lexicon

Pierre Pica[1], Cathy Lemer[2], Véronique Izard[2], and Stanislas Dehaene[2,*]

[1] Unité Mixte de Recherche 7023 « Structures Formelles du Langage », CNRS and Paris VIII University, Paris, France

[2] Unité INSERM 562 "Cognitive Neuroimaging", Service Hospitalier Frédéric Joliot, CEA/DSV, 91401 Orsay cedex, France.

[*] Corresponding author

Abstract

Is calculation possible without language? Or is the human ability for arithmetic dependent on the language faculty? To clarify the relation between language and arithmetic, we studied numerical cognition in speakers of Mundurukú, an Amazonian language with a very small lexicon of number words. Although the Mundurukú lack words for numbers beyond five, they are able to compare and add large approximate numbers, way beyond their naming range. However, they fail in exact arithmetic with numbers larger than 4 or 5. Our results imply a distinction between a universal system of number approximation, and a language-based counting system for exact number and arithmetic.

One-sentence summary

Psychological experiments in speakers of Mundurukú, an Amazonian language where number words exist only up to five, indicate that they can perform approximate calculations, but are unable to calculate with exact numbers.

> All science requires mathematics. The knowledge of mathematical things is almost innate in us. . . . This is the easiest of sciences, a fact which is obvious in that no one's brain rejects it; for laymen and people who are utterly illiterate know how to count and reckon.
>
> Roger Bacon

Where does the human competence for arithmetic arise from? Two theories have been proposed. Some authors propose that humans, like many animals share a non-verbal "number sense", an evolutionarily ancient capacity to process approximate numbers without symbols or language (*1-3*). Mathematics would arise from this non-verbal foundation. Others, however, postulate that language plays a special role in arithmetic. Numerical competence would arise from a progressive linguistic construction, supported by verbal counting and the recursive character of the language faculty (*4-7*). In the absence of language, numerical competence would therefore be drastically limited.

To elucidate the relations between language and arithmetic, numerical competence must be studied in situations in which the language of numbers is either absent or reduced. In many animal species, as well as in young infants prior to the acquisition of number words, behavioral and neurophysiological experiments have revealed rudiments of arithmetic (*8-12*). Infants and animals only appear to represent the first three numbers exactly. Beyond this range, they can only approximate "numerosity", with a fuzziness that increases linearly with the size of the numbers involved. This and other neuroimaging and neuropsychological experiments have yielded a tentative reconciliation of the above two theories: exact arithmetic would require language, while approximation would not (*13-16*). This conclusion, however, has been challenged by a few case studies of adult brain-lesioned or autistic patients in whom the lack of language did not abolish exact arithmetic, suggesting that even complex calculation might also be independent of language (*17*).

In the final analysis, the debate cannot be settled by studying subjects who are raised in a culture teeming with spoken and written symbols for numbers. What is needed is a language deprivation experiment, in which neurologically normal adults would be raised without number words or symbols. While such an experiment is ethically impossible in our Western culture, some languages are intrinsically limited in their ability to express number, sometimes using a very narrow set of number words ("one, two, many"). These often endangered languages present a unique opportunity to establish the extent and limits of non-verbal arithmetic abilities.

Here, we study numerical cognition in native speakers of Mundurukú, a language which has number words only for the numbers one through five (*18, 19*). Mundurukú is a language of the Tupi family, spoken by approximately 7,000 people living in an autonomous territory in the Para state of Brasil (figure 1). Following two pilot trips in 2001 and 2002, one of us (P.P.) travelled through several villages during 2003 and was able to collect data from 55 speakers of Mundurukú in a computerized battery of numerical tests. Ten native speakers of French (mean age 50) served as controls. The Mundurukú entertain contacts with non-indigenous culture and individuals, mainly through government institutions and missionaries. Thus, several of them speak some Portuguese. A few, especially the children, now receive basic instruction. Thus, we distinguished six groups of subjects based on their age, bilingualism, and level of instruction (figure 1). Using a solar-powered laptop computer, we managed to collect a large amount of trials in classical arithmetical tasks, including a chronometric comparison test. This allowed us to test whether classical effects such as Weber's law and the distance effect, which characterize numerical cognition in Western subjects (*3, 20-23*), remain present in the absence of a well-developed language for number.

A first task explored the verbal expressions for numbers in Mundurukú (*24*). Participants were presented with displays of 1-15 dots, in randomized order, and were asked to say in their native language how many dots were present, thus permitting an objective analysis of the conditions of use of number words. No systematic variation across groups was identified, except for lack of use of the word for "five" in the younger children, and the results were therefore pooled across all groups (figure 2). The results confirm that Mundurukú has frozen expressions only for numbers 1-5. These expressions are long, often having as many syllables as the corresponding quantity. Given the universal correlation between word length and word frequency (*25*), this finding suggests that numerals are infrequently used in Mundurukú. The words for three and four are polymorphemic: ebapũg=2+1, ebadipdip=2+1+1, where "eba" means "your (two) arms". This possibly reflects an earlier base-2 system common in Tupi languages. Above 5, there was little consistency in language use, with no word or expression representing more than 30% of productions to a given target number. Participants relied on approximate quantifiers such as "few" (adesũ), "many" (ade), or "a small quantity" (bũrũmaku). They also used a broad variety of expressions varying in attempted precision, such as "more than one hand", "two hands", "some toes", all the way up to long phrases such as "all the fingers of the hands and then some more" (in response to 13 dots).

Crucially, the Mundurukú did not use their numerals in a counting sequence, nor to refer to precise quantities. They usually uttered a numeral without counting, although if asked to do so they could count very slowly and non-verbally by matching their fingers to the set of dots. Our measures confirm that they selected their verbal response based on an apprehension of approximate number rather than on an exact count. With the possible exception of "one" and "two", all numerals were used in relation to a range of approximate quantities rather than a precise number (figure 2). For instance, the word for five, which can be translated as "one hand" or "a handful", was used for 5, but also 6, 7, 8 or 9 dots. Conversely, when 5 dots were presented, the word for "five" was uttered only on 28% of trials, while the words "four" and "few" were each used on about 15% of trials. This response pattern is comparable to the use of round numbers in Western languages, for instance when we say "ten people" when there are actually 8 or 12. We also noted the occasional use of two-word constructions (e.g. "two-three seeds") which have been analyzed as permitting reference to approximate quantities in Western languages (*26*). Thus, the Mundurukú are only different from us in failing to count and in allowing approximate use of number words in the range 3-5, where Western numerals usually refer to precise quantities.

If the Mundurukú have a sense of approximate number, they should be able to process numbers non-verbally way beyond the range for which they have number words. If, however, concepts of numbers emerge only when number words are available, then the Mundurukú would be expected to be at chance level with large numbers. We tested this alternative using two estimation tasks. First, we examined number comparison, a task which, in Western subjects, has revealed an effect of numerical distance whether the targets are presented as sets of objects or symbolically as Arabic digits (*20, 21*). Participants were presented with two sets of up to 80 dots, and were asked to point to the more numerous set (figure 3a). Mundurukú participants responded way above chance level in all groups (minimum 70.5 % correct in the younger group; $p<0.0001$). There was no significant difference among the six groups of Mundurukú subjects, suggesting that the small level of bilingualism and instruction achieved by some of the participants did not modify performance. Although average Mundurukú performance was slightly worse than the French controls, thus creating a difference between groups ($F_{6,55}=2.58$, $p<0.028$), this appeared to be due to a small proportion of trials with random responses, distributed throughout the experiment, and probably due to distraction in some Mundurukú participants (this was the first test that they took).

Crucially, performance varied significantly as a function of the ratio of the two numbers, thus revealing the classical distance effect in Mundurukú subjects ($F_{3,138}=43.2$

$p<0.0001$). This effect was identical in all groups, including the French controls (group X distance interaction, $F<1$; see figure 3a). Response times were also faster for more distant numbers in Munduruku ($F_{3,90}=12.9$, $p<0.0001$, and $F_{3,26}=4.93$, $p<0.008$). Again, although the French controls were globally faster, thus creating a main effect of group ($F_{6,37}=4.59$, $p<0.002$), the distance effect was parallel in all groups (interaction $F<1$). Fitting the performance curve suggested a Weber fraction of 0.17 in Mundurukú, only marginally larger than the value of 0.12 observed in the controls. Thus, the Mundurukú clearly can represent very large numbers and understand the concept of relative magnitude (*27*).

We then investigated whether the Mundurukú can perform approximate operations with large numbers. We used a non-symbolic version of the approximate addition task, which is thought to be independent of language in Western subjects (*13-15*). Participants were presented with simple animations illustrating a physical addition of two large sets of dots into a can (figure 3b). They had to approximate the result and compare it to a third set. All groups of participants, including monolingual adults and children, performed considerably above chance (minimum 80.7% correct, $p<0.0001$). Performance was again solely affected by distance ($F_{3,152}=78.2$, $p<0.0001$), without any difference between groups nor a group by distance interaction (see figure 3) (*28*). If anything, performance was higher in this addition+comparison task than in the previous comparison task, perhaps because the operation was represented more concretely by object movement and occlusion. Mundurukú participants had no difficulty in adding and comparing approximate numbers, with a precision identical to that of the French controls.

Finally, we investigated whether the Mundurukú can manipulate exact numbers. The number sense view predicts that outside the language system, number can only be represented approximately, with an internal uncertainty that increases with number (Weber's law). Beyond the range of 3-4, this system cannot reliably distinguish an exact number *n* from its successor $n+1$. Thus, the Mundurukú should fail with tasks that require manipulation of exact numbers such as "exactly six". To assess this predicted limitation of Mundurukú arithmetic, we used an exact subtraction task. Participants were asked to predict the outcome of a subtraction of a set of dots from an initial set comprising from 2 to 8 items (figure 3c and d). The result was always small enough to be named, but the operands could be larger (e.g. 6-4). In the main experiment, for which we report statistics below, participants responded by pointing to the correct result amongst three alternatives (0, 1 or 2 objects left). The results were also replicated in a second version in which participants named the subtraction result aloud (figure 3d).

In both tasks and in all Mundurukú groups, we observed a collapse of performance when the initial number exceeded 5, which marks the end of the naming range in Mundurukú ($F_{7,336}=44.9$, $p<0.0001$). This collapse contrasted sharply with the good performance observed in the French controls, which was only slightly affected by number size ($F_{7,63}=2.36$, $p<0.033$). Thus, we observed a highly significant group effect ($F_{6,57}=9.10$, $p<0.0001$) and a group by size interaction ($F_{42,399}=2.44$, $p<0.0001$). The Mundurukú's failure was not due to misunderstanding of the instructions, because they performed better than chance and indeed close to ceiling when the initial number was below 4, and thus could be named relatively precisely (see figure 2). In fact, performance remained slightly above chance for all values of the initial number (e.g. 49.6% correct for 8-n problems, chance = 33.3%, $p<0.0001$). The entire performance curve over then range 1-8 could be fitted by a simple psychophysical equation which supposes an approximate Gaussian encoding of the initial and subtracted quantities, followed by subtraction of those internal magnitudes and classification of the fuzzy outcome into the required response categories (0, 1 or 2). Thus, the Mundurukú still deployed approximate representations, subject to Weber's law, in a task that the French controls easily resolved by exact calculation.

Altogether, our results bring some light on the issue of the relation between language and arithmetic. They suggest that a basic distinction must be introduced between approximate and exact mental representations of number, as also suggested by earlier behavioral and brain-imaging evidence (*13, 15*). The Mundurukú have no difficulty with approximate numbers. They can mentally represent very large quantities of up to 80 dots, way beyond their naming range. They also spontaneously apply concepts of addition, subtraction and comparison to these approximate representations. This is true even for monolingual adults and young children that never learned any formal arithmetic. Those data add to previous evidence that number approximation is a basic competence, independent of language, and available even to preverbal infants and many animal species (*8-12*).

What the Mundurukú appear to lack, however, is a fast apprehension of exact numbers. Our results thus support the hypothesis that language plays a special role in the emergence of exact arithmetic during child development (*29-31*). What is the mechanism for this developmental change? It is noteworthy that the Mundurukú have number names up to 5, and yet use them approximately in naming. Thus, the availability of number names, in itself, may not suffice to create a mental representation of exact number. More crucial perhaps is that the Mundurukú do not have a counting sequence of numerals. Although some possess a rudimentary ability to count on their fingers, it is rarely used. By requiring an exact one-to-one pairing of objects with the sequence of numerals, verbal counting may promote a conceptual integration of the approximate number representation with the discrete object representation system (*29, 30*). Around the age of 3, Western children exhibit an abrupt change in number processing as they suddenly realize that each count word refers to a precise quantity (*31*). This "crystallization" of discrete numbers out of an initially approximate continuum of numerical magnitudes does not seem to occur in the Mundurukú.

Acknowledgements

This work was developed as part of a larger project on the nature of quantification and functional categories developed jointly with the linguistic section of the Departement of anthropology of the National museum of Rio and the Unité Mixte de Recherche 7023 of the CNRS, with the agreement of Funai and CNPQ. It was supported by INSERM, CNRS, the French Ministry of Foreign Affairs (P.P.), and a McDonnell Foundation centennial fellowship (S.D.). We are grateful to Manuela Piazza for her help with stimulus design, and to Celso Tawe and Venancio Poxõ for help in testing.

Figure legends

Figure 1. Location of the main Mundurukú territory where our research was conducted. Colored dots indicate the villages where participants were tested. The legend at bottom gives the size of the six groups of subjects and their average age.

Figure 2. Number naming in Mundurukú. Participants were shown sets of 1-15 dots in random order, and were asked to name the quantity. The graph shows the frequency with which a given word or locution was used in response to a given stimulus number. We only present the data for words or locutions produced on more than 2.5 % of all trials. Above 5, frequencies do not add up to 100%, because many participants produced rare or idiosyncrasic locutions or phrases such as “all of my toes” (a complete list is available from the authors).

Figure 3. Performance in four tasks of elementary arithmetic. In each case, the left column illustrates a sample trial. The graphs at right show the percentage of correct trials, in each group separately (M=monolinguals, B=bilinguals, NI=no instruction, I=instruction) as well as averaged across all the Mundurukú and French participants (right graphs). The lowest level on the scale always corresponds to chance performance. For number comparison (top two graphs), the relevant variable that determines performance is the distance between the numbers, as measured by the ratio of the larger to the smaller number. For exact subtraction (bottom two graphs), the relevant variable is the size of the initial number n1. The fits are based on mathematical equations described in Supplementary Online Material.

References and notes

1. C. R. Gallistel, R. Gelman, *Cognition* **44**, 43 (1992).
2. S. Dehaene, *The number sense* (Oxford University Press, New York, 1997)
3. S. Dehaene, G. Dehaene-Lambertz, L. Cohen, *Trends in Neuroscience* **21**, 355 (1998).
4. J. R. Hurford, *Language and Number.* (Basil Blackwell, Oxford, 1987)
5. N. Chomsky, *Language and the problems of knowledge* (MIT Press, Cambridge, 1988)
6. P. Bloom, *How children learn the meanings of words* (MIT Press, Cambridge, 2000)
7. H. Wiese, *Numbers, language, and the human mind* (Cambridge University Press, 2003)
8. M. D. Hauser, F. Tsao, P. Garcia, E. Spelke, *Proc R Soc Lond B Biol Sci* **270**, 1441 (2003).
9. A. Nieder, E. K. Miller, *Proc Natl Acad Sci U S A* **101**, 7457 (May 11, 2004).
10. E. M. Brannon, H. S. Terrace, *Journal of Experimental Psychology: Animal Behavior Processes* **26**, 31 (2000).
11. L. Feigenson, S. Carey, M. Hauser, *Psychol Sci* **13**, 150 (2002).
12. J. Lipton, E. Spelke, *Psychological Science* **14**, 396 (2003).
13. S. Dehaene, E. Spelke, P. Pinel, R. Stanescu, S. Tsivkin, *Science* **284**, 970 (1999).
14. E. S. Spelke, S. Tsivkin, *Cognition* **78**, 45 (2001).
15. C. Lemer, S. Dehaene, E. Spelke, L. Cohen, *Neuropsychologia* **41**, 1942 (2003).
16. S. Dehaene, L. Cohen, *Neuropsychologia* **29**, 1045 (1991).
17. B. Butterworth, *The Mathematical Brain* (Macmillan, London, 1999)
18. C. Strömer, *Die sprache der Mundurukú* (Verlag der Internationalen Zeitschrift "Anthropos", Vienna, 1932)

19. M. Crofts, *Aspectos da língua Mundurukú* (Summer Institute of Linguistics, Brasilia, 1985)

20. R. S. Moyer, T. K. Landauer, *Nature* **215**, 1519 (1967).

21. P. B. Buckley, C. B. Gillman, *Journal of Experimental Psychology* **103**, 1131 (1974).

22. J. Whalen, C. R. Gallistel, R. Gelman, *Psychological Science* **10**, 130 (1999).

23. H. Barth, N. Kanwisher, E. Spelke, *Cognition* **86**, 201 (Jan, 2003).

24. Materials and methods are available as supporting material on Science Online.

25. G. K. Zipf, *The psycho-biology of language.* (George Routledge, London, 1936), pp.

26. T. Pollmann, C. Jansen, *Cognition* **59**, 219 (May, 1996).

27. Comparison performance remained way above chance in two independent sets of trials where the two sets were equalized either on intensive parameters such as dot size, or on extensive parameters such as total luminance (see Methods). Thus, subjects did not base their responses on a single non-numerical parameter. Performance was however worse for extensive-matched pairs (88.3% versus 76.3% correct, $p<0.0001$). We do not know the origins of this effect, but it is likely that, like Western subjects, the Munduruku estimate number via some simple relation such as the total occupied screen area divided by the average space around the items, which can be subject to various biases (see J. Allik, T. Tuulmets, *Perception & Psychophysics* **49**, 303 (1991)).

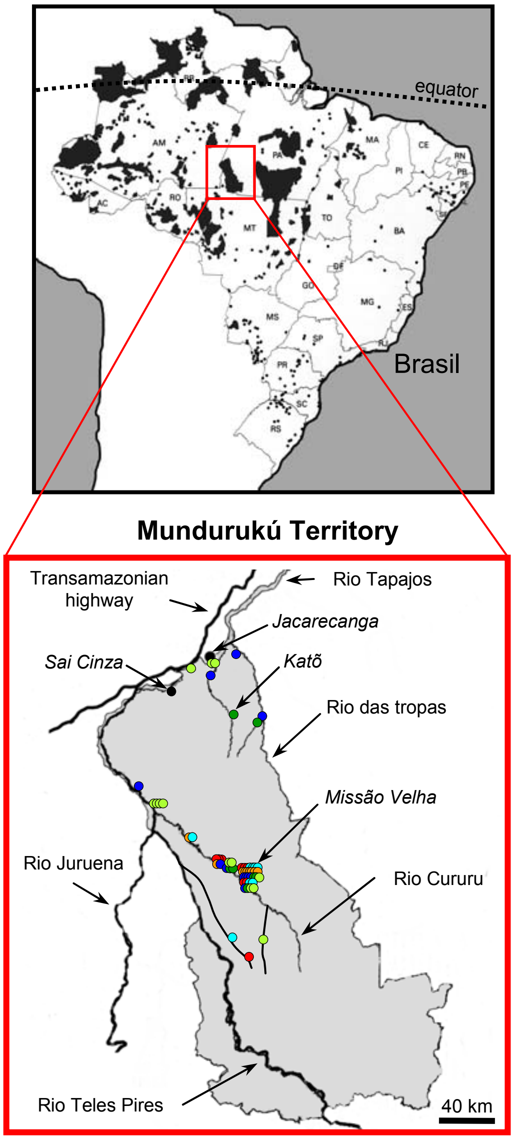
28. Performance remained above chance for both intensive-matched and extensive-matched sets (respectively 89.5 and 81.8% correct, both $p<0.0001$). Although this difference in stimulus sets was again significant ($p<0.0001$), it was identical in Munduruku and French subjects. Furthermore, performance was significantly above chance for a vast majority of items (44/51), and was never significantly below chance, making it unlikely that participants were using a simple short-cut other than mental

addition. For instance, they did not merely compare n1 with n3 or n2 with n3, because when n1 and n2 were both smaller than n3, they still discerned accurately if their sum was larger or smaller than the proposed number n3, even when both differed by only 30% (respectively 76.3 and 67.4% correct, both $p<0.005$).

29. E. Spelke, S. Tsivkin, in *Language acquisition and conceptual development* M. Bowerman, S. C. Levinson, Eds. (Cambridge University Press, Cambridge, 2001) pp. 70-100.

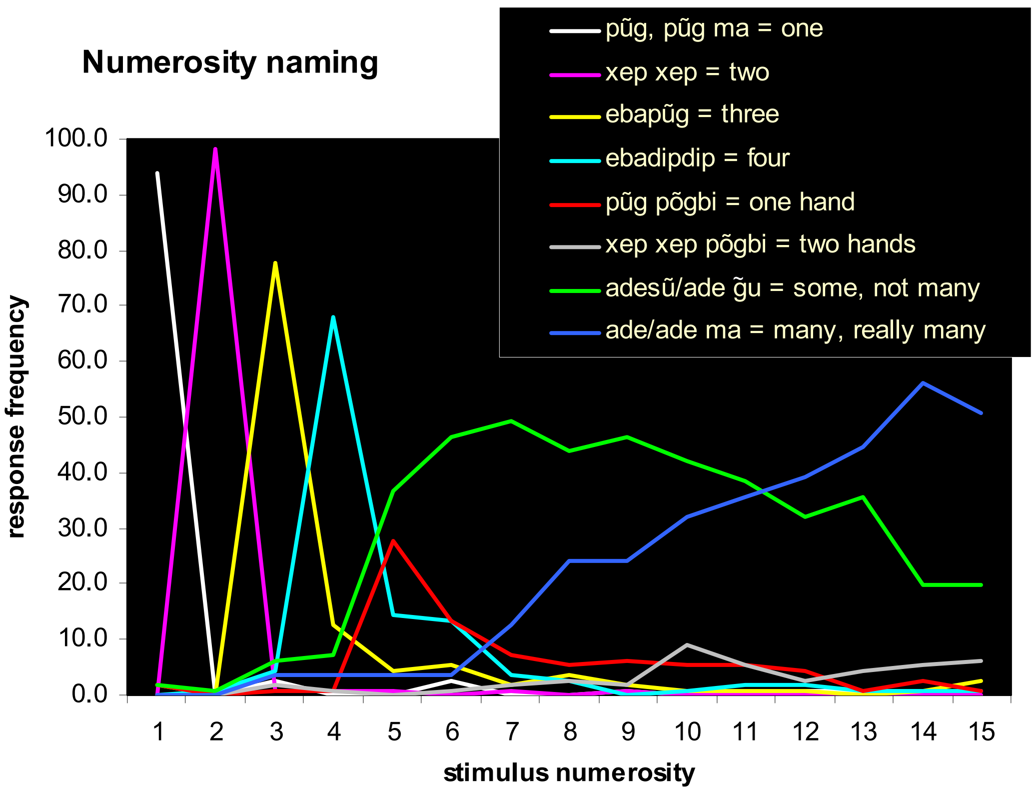
30. S. Carey, *Science* **282**, 641 (1998).

31. K. Wynn, *Cognition* **36**, 155 (1990).

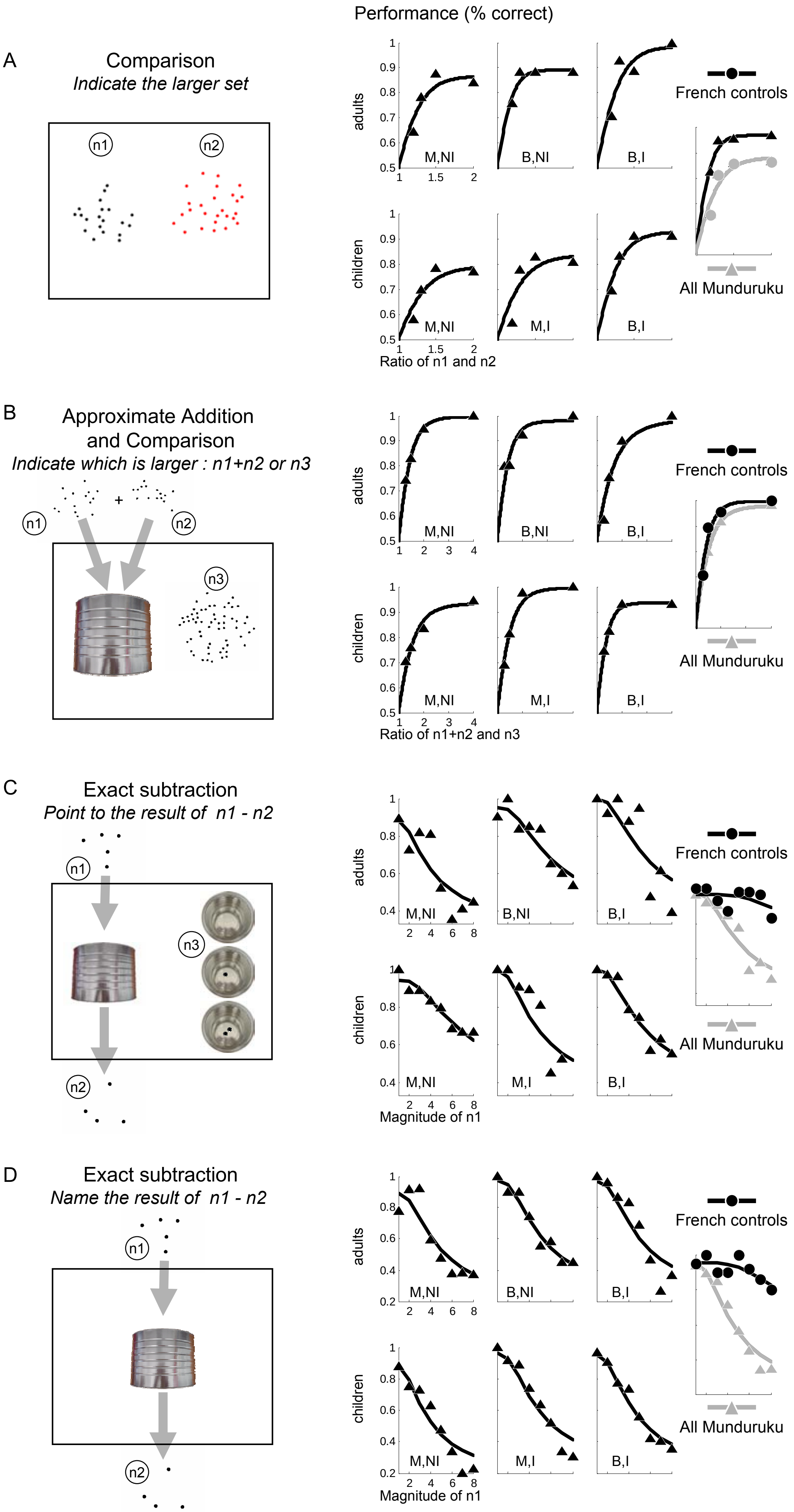


Adults	Children	
n=9 (55.5 y)	n=9 (4.7 y)	monolingual, no instruction
n=10 (59.3 y)		bilingual, no instruction
	n=7 (8.6 y)	monolingual, with instruction
n=7 (38.7 y)	n=13 (9.6 y)	bilingual, with instruction

Pica, Lemer, Izard & Dehaene, Figure 1



Pica, Lemer, Izard & Dehaene, Figure 2



Pica, Lemer, Izard & Dehaene, Figure 3

Supplementary on-line materials

Experimental methods

Participants were recruited locally at various villages through the active participation of the Mundurukú themselves. Instructions were always delivered in participants' native language. Stimuli were displayed on a solar-powered portable computer running PsyScope and PowerPoint software. For the numerosity naming and comparison tests, stimuli were occasionally presented on paper hand-outs (one stimulus per page).

Numerosity naming. Thirty sets of randomly arranged dots were presented sequentially. Each numerosity 1 through 15 was presented twice. Participants were asked to describe verbally the number of items. Two series of 15 sets were generated : in one, the extensive variables of total luminance and occupied area were equated across numerosities; in the other, the intensive variables of average dot size and spacing were equated.

Comparison. On each of 48 trials, two sets comprising 20-80 dots were presented side by side, the left set in black and the right set in red. Participants were asked to indicate the more numerous display. Half of the stimuli had the larger set on the right side. The Weber ratio w of the larger to the smaller numerosity, which determines the difficulty of the comparison operation in Western subjects (*1-3*), was systematically varied (w=1.2, 1.3, 1.5 or 2.0; 12 trials each). For each value of w, three pairs of numerosities were used, with different sizes of numerosity (small : 20-30 dots, middle : 30-60 dots, large : 40-80 dots). In half of the pairs total luminance and occupied area were equated across all stimuli, while in the other half, average dot size and spacing were equated.

In most cases, response times were measured to the nearest millisecond by asking participants to respond by depressing one of two large response keys coloured black and red like the stimuli. Some participants who refused to use the computer keys or failed to depress them properly were asked to merely point to the larger set. In the end, analyzable response choices and response times were obtained in 52 and 38 Mundurukú participants respectively.

Approximate addition and comparison. On each of 51 trials, a short sequence was presented to exemplify the addition of two large sets of dots (see figure 3). A can was first shown to be empty. It then rotated to the upright position, and two random sets of dots successively descended from the top of the screen into the can (duration of each set's motion: 5 seconds; inter-set delay, 0 seconds). Immediately after, a third set appeared on the right side of screen. Participants were asked to indicate the larger set (either the total hidden in the can, or the visible set). The same randomization and controls as in the comparison task were used. The ratio of n1+n2 and n3 was varied to manipulate task difficulty. The first three trials had a Weber ratio of 4, after which trials with w=1.3, 1.5 or 2.0 were randomly intermixed (16 trials each). The total numerosities ranged from 30 to 80, and one of them was split into two smaller numerosities according to a 2:1, 1:1, or 1:2 ratio. Analyzable data were obtained in 52 Mundurukú participants.

Exact subtraction. On each trial, the screen first showed an empty can. Then some dots fell into the can from the top, and finally some dots disappeared through the bottom (duration of each set's motion: 2 seconds, inter-set delay: 6 seconds; see figure 3). In the non-verbal multiple-choice version, participants had to select which of three displays matched the final content of the can: 2 dots, 1 dot, or no dot. All problems with an initial numerosity in the range 1-8 and a final numerosity in the range 0-2 were presented once or twice, for a total of 30 trials. In the verbal response version, participants described verbally the final contents of the can. All problems with an initial numerosity in the range 1-8 and a final numerosity in the range 0-4 were presented once or twice, for a total of 43 trials. There is no word for "zero" in Mundurukú but participants spontaneously offered short expressions such as "there is nothing left".

On half of the trials, dot spacing and total occupied area increased with numerosity, and they decreased with numerosity in the other half of trial. Dot size was assigned pseudo-randomly to one of two predetermined values. For the forced-choice task, the choice sets could be either of the same dot size as the problem sets, or of a different size.

Analyzable data were obtained in 54 Mundurukú participants in the multiple-choice task, and in 51 participants in the verbal response task.

Mathematical Theory

The number line model: basic formalism

Analog models of number processing assume that each numerical quantity is represented internally by a distribution of activation on an internal 'number line' (*3*-*5*). This internal representation is inherently noisy and varies from trial to trial. Assuming a specific form for this representation, tools from psychophysical theory can then be used to evaluate the optimal strategy and expected success rate (*6*).

Mathematically, the numerosity of a set of *n* dots is represented internally by a Gaussian random variable *X* with mean *q(n)* and with standard deviation *w(n)*.

$$p[X \in [x, x + dx)] = \frac{1}{\sqrt{2\pi}\, w(n)} \exp\left(-\frac{(x-q(n))^2}{2\, w(n)^2}\right) dx$$

The function *q(n)* defines the *internal scale* for number. In the literature, two main hypotheses about this scale are debated.

- According to the 'scalar variability' model, the internal scale is linear ($q(n) = n$) and the standard deviation w_n also increases linearly with *n* ($w(n) = w\, n$).

- According to the 'logarithmic number line' model, the internal scale is logarithmic ($q(n) = \mathrm{Log}(n)$) and the standard deviation $w(n)$ is a constant for all numbers ($w(n) = w$).

In those equations, w is the *internal Weber fraction* that specifies the degree of precision of the internal quantity representation. Both models lead to extremely similar predictions concerning behavior, with only subtle differences relating to asymmetries in observed distributions (*7, 8*). In particularly, both models predict a similar distance effect and Weber's law. In the present work, therefore, we do not attempt to distinguish between them, but rather use the model that leads to the simplest mathematical derivation in the present context (the 'scalar variability' model). We have verified that similar fits are obtained with the logarithmic number line model.

Comparison task

In the comparison task, participants are asked to decide which of two sets with numerosities *n1* and *n2* is the largest. In general, the optimal criterion for responding based on the internal observation of samples *X1* and *X2* may depend on the distribution of number pairs presented in the experiment. However, in the frequent situation where the same numbers have an equal probability of appearing in either order, the optimal criterion based on maximum likelihood consists simply in responding that the set with the greater internal representation is the larger (respond *n2>n1 iff X2-X2>0*). The value *X2-X1*, being the sum of two Gaussian random variables, is also a Gaussian random variable with mean *n1-n2* and standard deviation $w\sqrt{\mathrm{n1}^2+\mathrm{n2}^2}$. The error rate is the area under the tail of this Gaussian curve, or

$$p_{\text{comparison}} = \int_0^{+\infty} \frac{e^{-\frac{1}{2}\left(\frac{x+\mathrm{Abs}(\mathrm{n1}-\mathrm{n2})}{w\sqrt{\mathrm{n1}^2+\mathrm{n2}^2}}\right)^2}}{\sqrt{2\pi}\, w\sqrt{\mathrm{n1}^2+\mathrm{n2}^2}}\, dx = \frac{1}{2}\,\mathrm{erfc}\left(\frac{\mathrm{Abs}(\mathrm{n1}-\mathrm{n2})}{\sqrt{2}\, w\sqrt{\mathrm{n1}^2+\mathrm{n2}^2}}\right)$$

where erfc(x) is the complementary error function, given by $\mathrm{erfc}(z) = 1 - \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2}\, dt$

By letting $r = \frac{\mathrm{n2}}{\mathrm{n1}}$, one gets the predicted error rate

$$p_{\text{comparison}} = \frac{1}{2}\,\mathrm{erfc}\left(\frac{\mathrm{Abs}(r-1)}{\sqrt{2}\, w\sqrt{r^2+1}}\right)$$

Approximate addition task

In this task, participants are presented with three sets of numerosities *n1, n2* and *n3*, and are asked to compute the sum of $n1 + n2$ and to compare it with *n3*. In the general case, the optimal criterion based on maximum likelihood depends in a complicated way on the distribution of the numbers n1, n2 and n3 in the list of trials. However, a simplifying hypothesis assumes that participants respond by computing the sign of *N1+N2-N3*, where the N_i result from the conversion of the internal representations X_i back onto the number domain ($N_i = q^{-1}(X_i)$). This is where the mathematical derivation is much simpler for the 'scalar variability' model, because this relation reduces to $N_i = X_i$. The sum *N1+N2-N3*, being the sum of Gaussian random variables, is also a Gaussian random variable with mean n1 + n2 - n3, and standard deviation $w\sqrt{\mathrm{n1}^2 + \mathrm{n2}^2 + \mathrm{n3}^2}$. The error rate is the area under the tail of this curve, or

$$p_{\mathrm{ApproximateAddition}} = \frac{1}{2}\,\mathrm{erfc}\left(\frac{\mathrm{abs}(\mathrm{n1}+\mathrm{n2}-\mathrm{n3})}{\sqrt{2}\,w\sqrt{\mathrm{n1}^2+\mathrm{n2}^2+\mathrm{n3}^2}}\right) = \frac{1}{2}\,\mathrm{erfc}\left(\frac{\mathrm{abs}(r-1)}{\sqrt{2}\,w\sqrt{r^2+1+2\,\alpha(\alpha-1)}}\right),$$

where $r = \frac{\mathrm{n3}}{\mathrm{n1}+\mathrm{n2}}$, and $\alpha = \frac{\mathrm{n1}}{\mathrm{n1}+\mathrm{n2}}$ reflects the decomposition of the sum into two parts.

In general the error rate depends both on the ratio *r* and on the value of α. However the decomposition ratio α has a modest impact, as the term $2\alpha(\alpha\text{-}1)$ is always between -0.5 and 1 and is involved in a sum with a constant term 1. For this reason, we only considered the effect of the Weber fraction r in the fits, and replaced the term $2\alpha(\alpha\text{-}1)$ by its mean value in the experiment.

Exact subtraction tasks

In this task, the participants are presented with two sets of numerosities *n1* and *n2* and are asked to compute the substraction *n1-n2*. The result must be given either verbally in the production task, or by choosing amongst three possible values (0, 1 and 2) in the forced choice task. Again, the maximum likelihood view suggests that participants first compute mentally a mental representation of *N1-N2*, then examine where this value falls relative to fixed response criteria that divide the internal number line into multiple response domains. For the forced-choice task, those criteria separate the number line into three domains leading to responses 0, 1 or 2. In the production task, each possible verbal response *R* has two criteria $c_-(R)$ and $c_+(R)$ which defined a response interval on the number line. The requirement of having to respond to all targets further implies that $c_+(R) = c_-(R+1)$. Thus, the criteria were placed at the crossing points of the distributions associated with the different numerosities. Thus, $c_+(R)$ was defined by

$$\frac{1}{\sqrt{2\pi}\, w(R)} \exp\left(-\frac{(c_+(R)-q(R))^2}{2\,w(R)^2}\right) = \frac{1}{\sqrt{2\pi}\, w(R+1)} \exp\left(-\frac{(c_+(R)-q(R+1))^2}{2\,w(R+1)^2}\right)$$

This definition of the criteria is optimal for numerosity naming. The distribution for numerosity zero is not defined in the model, thus we modeled it by the distribution generated by the problem "1-1". Analysis shows that the particular choice of settings of the response criteria has only a weak influence on the final performance curve.

In the 'scalar variability' model, *N1-N2* is a gaussian random variable, with mean *n1-n2* and standard deviation $w\sqrt{\mathrm{n1}^2+\mathrm{n2}^2}$. Once the response criteria are defined, the probability of responding *R* to the problem *n1-n2* is given by

$$p(R/\mathrm{n1},\,\mathrm{n2}) = \int_{c_-(R)}^{c_+(R)} \frac{\exp\left(\frac{-(x-n1+n2)^2}{2\,(n1^2+n2^2)\,w^2}\right)}{\sqrt{2\pi}\; w\sqrt{n1^2+n2^2}}\, dx$$

The probability of responding correctly is given by substituting $R = n1\text{-}n2$ in the above expression. The probability of making an error is thus given by:

$$p_{\mathrm{ExactSubtraction}} = 1 - \int_{c_-(n1-n2)}^{c_+(n1-n2)} \frac{\exp\left(\frac{-(x-n1+n2)^2}{2\,(n1^2+n2^2)\,w^2}\right)}{\sqrt{2\pi}\; w\sqrt{n1^2+n2^2}}\, dx$$

Fitting the data

The above equations were used for fitting the data in figure 3. However, a small modification had to be introduced to obtained a good fit. All of the above equations characterize optimal performance and, in particular, imply that as the distance between the numbers increases, error rate drops to zero. However, perhaps because of frequent distraction in the environment in which the Munduruku were tested, we observed non-zero error rates in all conditions of our experiment (particularly in the uneducated adults and children). Thus, besides the parameter w, a second free parameter p was introduced, corresponding to a fixed probability of responding randomly on any trial. The final equations used to fit the observed probabilities of a correct response were therefore given by

$$p_{\text{correct}} = (1-p)(1-p_{\text{error}}) + p\, p_{\text{chance}}$$

with p_{error} specified by the above equations appropriate for each task, and p_{chance} determined by one over the number of possible responses in the task.

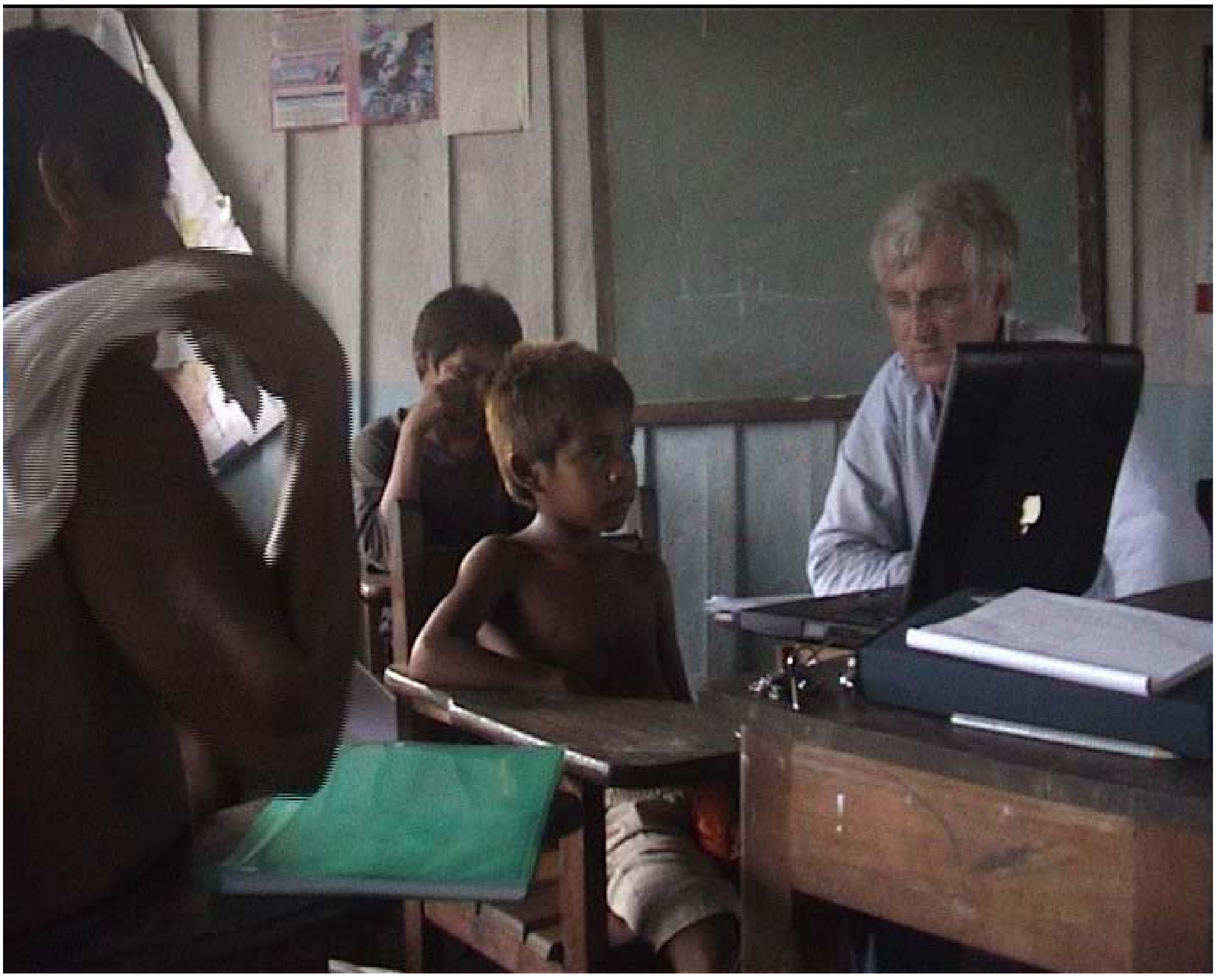
References for supplementary on-line material

1. R. S. Moyer, T. K. Landauer, *Nature* **215**, 1519 (1967).
2. P. B. Buckley, C. B. Gillman, *Journal of Experimental Psychology* **103**, 1131 (1974).
3. M. P. van Oeffelen, P. G. Vos, *Perception & Psychophysics* **32**, 163 (1982).
4. S. Dehaene, *Cognition* **44**, 1 (1992).
5. R. N. Shepard, D. W. Kilpatrick, J. P. Cunningham, *Cognitive Psychology* **7**, 82 (1975).
6. D. Green, J. A. Swets, *Signal detection theory and psychophysics* (Krieger Publishing Company, New York, 1966)
7. A. Nieder, E. K. Miller, *Neuron* **37**, 149 (Jan 9, 2003).
8. S. Dehaene, J. F. Marques, *Quarterly Journal of Experimental Psychology* **55**, 705 (2002).

Photos documenting the experiments

One of the villages where experiments were run.

A Mundurukú woman taking the number comparison test (paper version)



A child participating in the computerized tests under supervision of the first author (Pierre Pica)

A young participant taking the number comparison test.

One of the few educated adults who counted on his fingers and toes during the number naming test.

Photos © Pierre Pica CNRS UPRESA 7023