Supporting Information

Butterworth *et al.* 10.1073/pnas.0806045105

SI Text

Language. Warlpiri is in the Pama-Nyungan language family. It is a classifier language with three generic types of number words: singular (*jinta*), dual plural (*-jarra*, *jirrama*), and greater than dual plural (*jirrama manu jinta*, *marnkurrpa*, *wirrkardu*, *panu*). Anindilyakwa, probably unrelated to any other Australian language, is the major indigenous language spoken on Groote Eylandt. (It is also spoken in some small communities on neighbouring islands and on the nearby East Arnhem Land coast.) Like Warlpiri, Anindilyakwa is a classifier language, with nine noun classes and four possible number categories: singular, dual, trial (which may in practice include four), and plural (more than three). Anindilyakwa has a base-5 number system, apparently appropriated from the Macassan traders who visited the northern coast of Australia, including Groote Eylandt, from about the 17th century onward. It appears to be the case that the base-5 system is reserved for special cultural enumeration events (eg, distributing turtle eggs to recipients). In Anindilyakwa, numerals are adjectival, and must agree with the nouns they qualify (1). Because there are nine noun classes, enumerating in Anindilyakwa is complex. However, the number names are 1 (*awilyaba*), 2 (*ambilyuma* or *ambambuwa*), 3 (*abiyakarbiya*), 4 (*abiyarbuwa*), 5 (*amangbala*), 10 (*ememberrkwa*), 15 (*amaburrkwakbala*), and 20 (*wurrakiriyabulangwa*). The word for 20 is invariable; that is, it does not change its form in different grammatical contexts (1). The Anindylakwan number system is not formally introduced to members of the community until they reach adolescence. Stokes observes that "In traditional Aboriginal society nothing used to be counted that was outside normal everyday experience. When asked for what purpose counting was used in the old days, *the old women who know the number names* [emphasis added] say that counting was used for turtle eggs" (1; p. 39). She also maintains that "numerals in Anindilyakwa are adjectival. They are complicated by the number of noun classes [nine], because all adjectives must agree with the nouns they qualify." (1; p. 41) Although these languages contain quantifiers such as "few," "many," "a lot," "several," and so forth, these are not relevant number words because they do refer to exact numbers, and the theoretical claim is about exact numbers. Ordinals, such as "first," "second," and "third" would be more problematic. However, these words do not exist in either Warlpiri (2) or Anindilyakwa (1)

Indigenous Communities. Formal school-based instruction is conducted in English, with the aid of indigenous assistants who help by making links between the children's indigenous language and English. Formal schooling, which should begin at age 5, is the first time that children encounter English count words and related number concepts. Although schooling is nominally compulsory in the Northern Territory, relatively few children attend regularly (3). It should also be noted that the traditional method of learning in both sites is by observation not by instruction, and that up to the time that the child attended school, this is how he or she would have learned (4). This is the typical learning method among hunter-gatherers according to Kearins (4).

Methodology. The *Guidelines for ethical research in indigenous studies* (5) were followed in setting-up and conducting the study, as well as working with the communities. Research assistants spent approximately 3 weeks in the communities before data collection, to (a) become familiar with communities' social practices, and for the communities to become familiar with the research assistants; (b) learn rudimentary aspects of the indigenous language; (c) instruct indigenous helpers in research practices; and (d) familiarize children with test materials. Western research practices were strange to the indigenous assistants, who accommodated them with humor even though they would not interact with children in such ways. To acquaint helpers with research practices and to familiarize children with test materials (eg, counters), familiarization sessions were conducted. Children played matching and sharing games using test materials (counters and mats, and play-dough disks respectively). For the matching games, the interviewer put several counters on her mat and children were asked to make their mat the same. Children had little difficulty copying the number and location of counters on the interviewer's mat. For the sharing game, children shared three and six play-dough disks among thee doll recipients (child, mother, and father doll), the purpose of which was to see whether children's activities would be affected by the status of recipients. They were not: all children shared the play-dough disks in a one-to-one fashion.

Analysis of Trends Across Numerosities. The analysis of the relationship between the responses and the targets examined whether there was a discontinuity between small ($\leq$4) and large numbers ($>$4). For memory for counters tasks there was a linear trend for each language group [English: $y = 1.51x - 0.55$ ($r^2 = 0.96$); Warlpiri: $y = 1.07x + 1.12$ ($r^2 = 0.96$): Anindilyakwa: $y = 1.16x + 0.98$ ($r^2 = 0.99$)[. Similar linear trends were found for cross-modal matching tasks [English: $y = 0.87x + 0.83$ ($r^2 = 0.89$); Warlpiri: $y = 1.09x - 0.06$ ($r^2 = 0.99$): Anindilyakwa: $y = 1.11x - 0.13$ ($r^2 = 0.99$)] and nonverbal addition tasks [English: $y = 0.61x + 2.78$ ($r^2 = 0.80$); Walrpiri: $y = 0.87x + 2.15$ ($r^2 = 0.99$); Anindilyakwa: $y = 0.89x + 2.03$ ($r^2 = 0.99$)]. Moreover, a Bayesian likelihood analysis (6) comparing the frequency distributions for target numerosities 2 and 8 in the memory for counters task and 2 and 7 for the cross-modal matching task suggests that distributions were equivalent (ie, the odds in favor of the distributions being equivalent were greater than 2.5:1 for all comparisons, independent of age, and language group). As expected, Bayesian likelihood analysis comparing small with small numerosities and large with large numerosities also revealed odds in favor of distributions being equivalent.

Scalar Variability as an Index of Nonverbal Enumeration. Discriminating between precise and approximate representations of a number is difficult because of representational noise. Cordes and colleagues (7) outline a solution for determining whether number judgments represent enumeration activity (either verbal or nonverbal). They proposed a coefficient of variation (the ratio of the SD/Mean) for adjusting increasing variability in the estimates of means and standard deviations associated with increases in the numerosity of the target. This approach overcomes limitations of conventional regression analyses that do not take into account simultaneous variations of means and standard deviations. We examined performance on the memory for counters, cross-modal matching, and nonverbal addition tasks as a function of test site. An analysis of the models of variability, taking all of the NT participants as one group in order to increase the statistical power of our test, revealed that only one group in one task departed significantly from scalar variability (8). These were the Melbourne (English-speaking) children in the cross-modal matching task ($\beta = 0.44$, $P = 0.02$, $r^2 = 0.67$). However, this did not show a binomial distribution, which would require a negative slope of 0.5. Thus, even children who could count did not show evidence of verbal counting in a task where it may have been a useful strategy. Of course, this does not mean that they were not counting verbally, only that the distribution did not reveal it according to Cordes and colleague's criterion (7).

Strategic Use of Spatial Location by NT Children. Most NT children used a strategy that reproduced arrays similar to the interviewer's target arrays, independent of set size when correctly recalling counters. In contrast, English-speaking urban children tended to recall the number of counters without reference to where the interviewer placed them on her mat [χ^2 (1, n = 284) = 71.39, $P < 0.001$]. Younger NT children tended to reproduce the interviewer's arrays in correctly recalling counters more often than their older peers [small sets F (1, 16) = 5.06, $P < 0.05$; large sets F (1, 25) = 14.88, $P < 0.001$) (χ^2 (1, n = 225) = 5.15, $P < 0.05$] (8). This may reflect superior visuo-spatial memory skills in Aborigine children in remote sites (4).

Methodological Differences Between Our Study and Previous Research on Number Concepts in Indigenous Communities. Although it is true that the NT children in our study were legally obliged to go to school from the age of 5, not all of them did so, and even fewer went regularly. We would note here that the Mundurukú children also attended school and attempts were made to teach numbers to the Pirahã (9). The Pirahã also traded with outsiders. Everett (9) writes that, "Riverboats come regularly to the Pirahã villages during the Brazil nut season. This contact has probably been going on for more than 200 years. Pirahã men collect . . . [but] in this 'trade relationship' there is no evidence whatsoever of quantification or counting or learning of the basis of trade values." From this, and from his attempts to teach them about numbers, Everett concludes that the Piraha do not count, cannot count, and put no value on the practice of counting. He also noted that "many societies in the Amazon and elsewhere have borrowed number words as they develop economic ties that require numerical abilities" (9; p. 634). Moreover, he also noted that many Pirahã have at least some knowledge of Portuguese, although we are not told how many.

There are many differences between our participants from those of Pica and colleagues (10) and Gordon (11): Pica's study used nine monolingual children, mean age 4 (10); Gordon used nine monolingual adults, age 55(11). The controls were French adults. Gordon's study only used adults, and only four of these completed all of the tasks. There were no controls.

There were also important differences in the exact number tasks that were used. Gordon used seven versions of a number matching task and a kind of subtraction task (11). Pica and colleagues (10) only used one exact number task, subtraction. This was presented on a computer screen. Subtraction is notoriously difficult for 4 year olds in our culture and further, if subtraction is not a procedure that is part of the culture, why should one expect that even the adults would perform similarly to the French adult controls? By contrast, we used three very different exact number tasks, plus one: sharing disks made from play-dough, that could be treated as exact number or as continuous (so that we could see how the child dealt with the remainder). Pilot work showed that these tasks were culturally appropriate.

In a subsequent paper, Everett (9), who has worked with the Pirahã for many years, and who hosted Gordon in his studies with them, says that "on the videotape he made of his experimental setting, the Pirahãs say repeatedly that they do not know what he wants them to do, and they have repeated these comments since Gordon's visits. Gordon did not realize that they were confused because he was unable to communicate with them directly, and he did not request help in interpreting the Pirahãs' comments on his experiments." (9; p. 644)

The study of Mundurukú could also be interpreted as showing that possession of a number-word vocabulary, in addition to the two core systems of knowledge, is not sufficient for the development of concepts of exact numbers, because adults and children who were bilingual in Mundurukú and Portuguese are no better than monolingual Mundurukú speakers on the exact number tasks (10) (see Fig. 3*C*).

It should be noted that in the nearest comparable study (10), the kind of distribution accepted as reflecting an exact representation of small numerosities are similar to our own distributions. However, the Pica and colleagues distributions described as approximate, numbers above 4, are not like ours. (See Figs. 4 *A* to *C* for frequency graphs, and Ref. 10, Fig. 2 for comparison.)

1. Stokes B (1982) A description of the mathematical concepts of Groote Eylandt aborigines. in *Language and Culture. Work Papers of SIL-AAB, Series B*, ed Hargrave S (Summer Inst. of Linguistics, Darwin, Australia), Vol 8, pp 1–152.
2. Bittner M, Hale K (1995) Remarks on definiteness in Warlpiri. in *Quantification in Natural Languages*, ed Bach E (Kluwer, Dordrecht), pp 81–105.
3. Northern Territory Department of Education (1999) *Learning lessons: An independent review of Indigenous education in the Northern Territory* (Darwin, NT: Government Printing Office).
4. Kearins J (1981) Visual spatial memory of Australian Aboriginal children of desert regions. *Cognitive Psychol* 13:434–460.
5. Australian Institute of Aboriginal and Torres Strait Islander Studies (2000) Guidelines for ethical research in indigenous studies. Available at www.aiatsis.gov.au/_data/assets/pdf_file/2290/ethics_guidelines.pdf. Accessed Feb, 2003.
6. Glover S, Dixon P (2004) Likelihood ratios: A simple and flexible statistic for empirical psychologists. *Psychon B Rev* 11:791–807.
7. Cordes S, Gelman R, Gallistel CR, Whalen J (2001) Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychon B Rev* 8(4):698–707.
8. Butterworth B, Reeve R (2008) Verbal counting and spatial strategies in numerical tasks: Evidence from indigenous Australia. *Philos Psychol* 21:443–457.
9. Everett D (2005) Cultural constraints on grammar and cognition in Pirahã. Another look at the design features of human language. *Curr Anthropol* 46:621–646.
10. Pica P, Lemer C, Izard V (2004) Exact and approximate calculation in an Amazonian indigene group with a reduced number lexicon. *Science* 306:499–503.
11. Gordon P (2004) Numerical cognition without words: Evidence from Amazonia. *Science* 306:496–499.