STEPHEN LAURENCE AND ERIC MARGOLIS

Linguistic Determinism and the Innate Basis of Number

The ability to represent natural numbers is at the center of a lively controversy about the innate structure of the mind. As elsewhere in the study of cognition, there is a continuum of positions that implicate differing amounts of innate structure, but for our purposes it will be useful to distinguish three general approaches—what we will call empiricism, weak nativism, and strong nativism. *Empiricist* accounts maintain that there are no innate number-specific representations or number-specific cognitive systems of any kind and that the natural numbers are acquired on the basis of general cognitive resources that are responsible for the acquisition of a wide variety of concepts. *Weak nativist* accounts implicate considerably more innate structure, including number-related cognitive systems and representations of approximate quantity, but these accounts draw the line at concepts for specific natural numbers. They maintain that even though concepts for the natural numbers have a good deal of innate support, they have to be learned all the same. Finally, *strong nativist* accounts maintain that concepts for at least some specific natural numbers are innate and that these innate concepts are a crucial factor in the explanation of why the human mind is suited for mathematics.

In one respect, weak and strong nativists are natural allies. Both help themselves to domain-specific innate structure. Yet in other respects, weak nativists are closer to empiricists, since weak nativists and empiricists tend to agree that concepts for the natural numbers are a cultural achievement, like writing and agriculture. They view these concepts not as part of our innate endowment but as fundamentally owing to invention and discovery—a view we will refer to as the *Cultural Construct Thesis*. Of course, there is no disputing that culture influences mathematical cognition. Culture affects the richness of our numerical knowledge, the techniques we rely upon for exploiting numerical quantity, and the conventional means we use for recording and communicating mathematical information. But in adhering to the Cultural Construct Thesis, weak nativists and empiricists embrace the

This chapter was fully collaborative; the order of the authors' names is arbitrary. We would like to thank Peter Carruthers for helpful comments on an earlier version of this chapter.

more interesting claim that our very concepts for natural numbers are themselves cultural products and that any given concept for a natural number—whether it is TWO or TWO THOUSAND—owes as much to culture and learning as any other.

The emerging consensus in psychology is that weak nativists and empiricists are right to maintain the Cultural Construct Thesis and that strong nativism, because it stands in opposition to this thesis, is untenable. No doubt there are many reasons why the consensus has shifted in this direction, but one study that might be taken to provide particularly powerful support for this shift is Peter Gordon's (2004a, 2004b) high-profile cross-cultural study of number concepts among the Pirahã tribe in the Amazon. Gordon himself was primarily interested in testing the neo-Whorfian view that concepts for natural numbers are dependent on linguistic devices that not all languages share. He has described his study as constituting "a rare and perhaps unique case for strong linguistic determinism" (Gordon 2004a, p. 498). But proponents of the Cultural Construct Thesis would have a lot to gain if it could be shown that concepts for natural numbers are dependent on language in the way that Gordon supposes. Strong nativism would no longer appear to be a live option.

In this chapter, we take a careful look at Gordon's study and its implications for theories of numerical concepts. Other critical discussions of Gordon's work have noted general difficulties when cross-cultural data are used to draw inferences about the relationship between language and thought (R. Gelman and Gallistel 2004; R. Gelman and Butterworth 2005). Though we share these concerns, we believe there are even more fundamental objections to Gordon's experiments and that these objections are well worth exploring in the broader context that includes not just Gordon's linguistic determinism but also the Cultural Construct Thesis. We will argue that Gordon's experiments don't provide any support for either view and, consequently, that they don't diminish the prospects for strong nativism. Does this mean that we reject the use of cross-cultural research in the study of numerical concepts? Absolutely not. In fact, we hope that one of the benefits of our critical discussion will be a clearer picture of how cross-cultural research might be productively brought to bear on the psychology of number.

Here is how the chapter is organized. In section 1, we set the stage for our discussion by providing an overview of nativist approaches to numerical concepts. This includes a brief sketch of our own approach, which provides an illustration of a contemporary form of strong nativism. In section 2, we review Gordon's experiments and the prima facie case that his data support the type of linguistic determinism that he advocates as well as the Cultural Construct Thesis. In section 3 we turn to our objections. Finally, in section 4, we offer some thoughts on how future cross-cultural research might help to contribute to a better understanding of our most basic numerical capacities.

1 Nativism About Number

We begin, in this section, with a brief overview of the recent history of nativist theorizing about number and a sketch of our own general approach, which falls squarely in the strong nativist camp. Once this background is in place, we'll be in a position to turn to Gordon's study in section 2.

1.1 *From Strong to Weak Nativism*

Strong nativism can be traced back at least as far as Plato, but for contemporary theorists, the place to begin is with Rochel Gelman and C. R. Gallistel's landmark book *The Child's Understanding of Number.* Gelman and Gallistel put strong nativism back on the map by drawing much-needed attention to a wealth of data that proved troubling for the empiricist models that dominated developmental psychology in the 1970s. To account for this data, Gelman and Gallistel posited an innate system of representation with much the same structure as a conventional counting system, including its own stock of ordered discrete symbols. They referred to these symbols as *numerons*, but these were, in effect, innate natural number concepts. As Gelman and Gallistel saw it, the task of learning a conventional counting system isn't a matter of constructing the concepts from experience. It is primarily a matter of noting the correspondence between the public conventional system and the innate one and establishing an appropriate mapping between the two.

Despite Gelman and Gallistel's influence, contemporary theorizing about number is dominated by weak nativist accounts. One of the reasons Gelman and Gallistel's numerons fell out of favor is owing to an observation made by Karen Wynn (1990, 1992a). Wynn traced the developmental trajectory as children learn their public language counting system and noted that children generally take a long time to learn the meanings of individual count words even once they are familiar with the count sequence and with the procedures involved in counting (i.e., reciting count words in order while tagging one and only one item per word). Children can be counting for up to a year before they understand that counting is a way of enumerating a collection and before they understand the numerical significance of each of their count words. For example, before the age of three and a half, a child might be able to count as high as "six" and yet, when asked to give three items, the same child will often just grab a random number. Findings of this kind are problematic for Gelman and Gallistel's numeron hypothesis, since it's puzzling why it should take so long to interpret a conventional system in terms of a highly similar innate system. Wynn argued that the solution to the puzzle is that children don't have access to an innate stock of numerons; the innate system for representing numerical quantity must take a different form. Her suggestion was that it amounts to a system known as *the accumulator*, which uses mental magnitudes to represent approximate numerical quantities—the bigger the magnitude, the bigger the quantity represented (Meck and Church, 1983).[1] Two characteristic features of the accumulator are (i) that it has

1. Gelman rejects Wynn's critique on empirical grounds, citing data which she takes to show that children have more precocious counting skills than are revealed by Wynn's give-a-number task (R. Gelman 1993). At the same time, Gallistel and Gelman (1992, 2000; Gallistel, Gelman, and Cordes 2005) have followed Wynn in supposing that the preverbal system of numerical representation is the accumulator (though see Leslie, Gallistel, and Gelman, this volume). What allows Gelman to reject Wynn's critique while simultaneously embracing Wynn's suggestion about the accumulator is that Gelman views the accumulator as conforming to the counting principles. In other words, for Gelman, the accumulator's mental magnitudes are supposed to serve much the same function as numerons (e.g., Gallistel and Gelman 1992). See Laurence and Margolis (2005) for arguments that the accumulator should not be construed in this way.

more difficulty distinguishing between numbers that are closer to one another than it does numbers that are further apart (*the distance effect*) and (ii) that its discriminative capacity degrades as numbers become larger (*the magnitude effect*). So while the accumulator may represent numerical quantity, it lacks the precision that is integral to the natural numbers. As a result, though Wynn's commitment to the accumulator wasn't merely a throwback to prior empiricist models, it also wasn't the strong nativist position that Gelman and Gallistel had defended. In our terms, Wynn's proposal amounted to a form of weak nativism. It postulated a fair amount of innate structure without requiring specific natural number concepts to be innate.[2]

Much of the evidence in favor of the accumulator has come from experiments with animals (Gallistel 1990). Rats, pigeons, and many other species have been shown to be sensitive to approximate numerical quantity, and in a variety of tasks their behavior shows the telltale signs of the accumulator—increasingly variable discrimination both as the target numbers become larger and as they come closer together. It's important to bear in mind, however, that the animals are responding to *numerical* quantity, and that experimentalists have gone to great lengths to control for the many non-numerical properties that tend to correlate with number (e.g., duration for sequentially presented items and surface area for static spatial displays). One of the most elegant experiments along these lines is Elizabeth Brannon and Herbert Terrace's (1998) study of rhesus monkeys. Brannon and Terrace presented monkeys with four stimulus displays, each with one to four items, and trained the monkeys to touch the displays in ascending numerical order. Following training, the monkeys were tested on pairs of novel stimuli with as many as nine items, where the task was to indicate their numerical order. Both in the training period and in the test trials, Brannon and Terrance used stimuli that varied widely in terms of their shapes and sizes (see figure 8.1). This ensured that the monkeys couldn't solve the task by focusing on such non-numerical features as total surface area, total circumference, or surface luminance. Also, because all of the stimuli used for the test trials were novel, the monkeys couldn't fall back on memorized patterns from the training sessions. Despite these rigorous conditions, the monkeys did surprisingly well, responding far above chance levels (see figure 8.2). Much of the interest of Brannon and Terrace's study lies in the fact that it shows that monkeys can appreciate the ordinal relations among sets of different sizes. But success clearly depends on being able to discriminate the sets in terms of numerical quantity—in order to put them in ascending numerical order, the monkeys need to determine the numerical quantities of the different sets. What's more, the monkeys made more errors when

2. We should point out that Wynn may have thought of herself as a strong nativist, since she described the accumulator as delivering fairly precise representations for the first few numbers and only losing precision for numbers above 3. Wynn (1992c) also claimed that infants are able to appreciate the precise solutions to simple arithmetic problems using small numbers (see below). Nonetheless, it pays to construe Wynn's critique of the numeron hypothesis as opening the way for weak nativist theorizing, since the accumulator's representations *aren't* precise in the way that Wynn took them to be and since weak nativists now appreciate the ways in which the accumulator falls short of providing precise representations of the natural numbers (Carey 2001; Spelke 2003; Laurence and Margolis 2005).

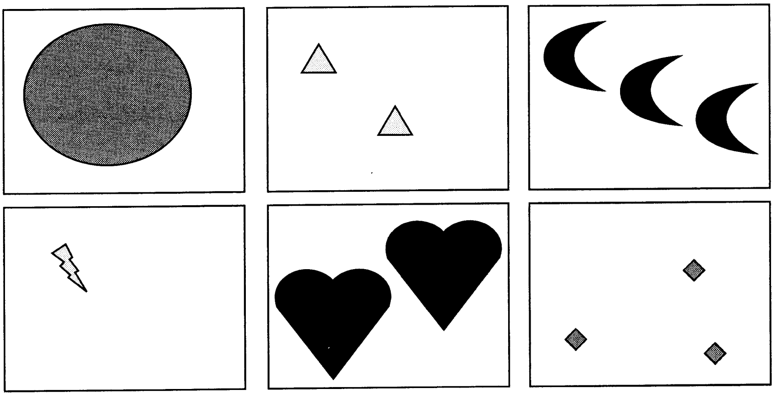

FIGURE 8.1 Examples of the types of stimuli used by Brannon and Terrace (1998). Non-numerical properties (e.g., surface area and total circumference) were carefully controlled for by varying the sizes and shapes of the elements with each trial.

the numerical comparisons involved finer distinctions. This strongly suggests not only that the underlying system of representation lacks the precision of the natural numbers but also that its representations are the mental magnitudes associated with the accumulator.

Work of this sort with animals has led to further experiments with humans and the discovery that humans of all ages—even infants—have access to the accumulator's approximate representations (Whalen et al. 1999; Xu and Spelke 2000; Lipton and Spelke 2003). Indeed, when Brannon and Terrace reran their experiment with human adults (instructing their participants to make their judgments as quickly as possible while being careful not to make errors), the results were nearly identical to the results for the monkeys (Brannon and Terrace 2002; see figure 8.2). The current consensus in psychology is that the accumulator is a ubiquitous cognitive system with an evolutionarily ancient history. But to embrace the accumulator as part of the innate structure of the mind is to take a good step away from an empiricist model of numerical cognition. This consideration, above all others, explains why so many theorists these days count as weak nativists. They adopt the view that we need at least this much domain-specific structure but assume we needn't go so far as to postulate representations of numerical quantity that are any more precise than the accumulator's mental magnitudes.

Still, weak nativists typically help themselves to more cognitive machinery than just the accumulator. Another system that is widely cited is the *object-indexing system* (also referred to as the *object-file system*). The object-indexing system is a mechanism of visual attention that is able to keep track of a small number of objects (up to four) by employing a comparably small number of representations that act like pointing devices. These representations, or indexes, function in parallel and track their respective objects by responding primarily to spatial-temporal properties. Object-based approaches to visual attention are well motivated apart from any concerns about numerical cognition (Scholl 2001). However, many psychologists have

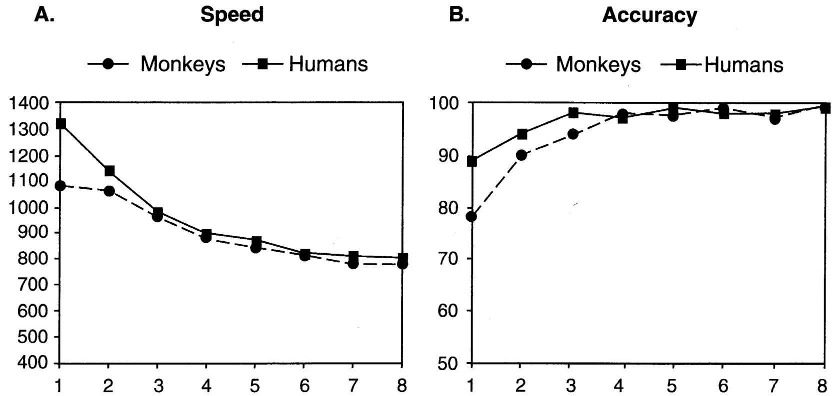

FIGURE 8.2 Brannon and Terrace's (1998) ordering task. The x-axis represents the numerical difference between stimuli; the y-axis represents time in milliseconds (2a) and accuracy (2b). Overall, monkeys and humans perform similarly. Both are quicker and more accurate in responding to larger numerical differences, though humans take slightly longer with stimuli differing by a value of just 1 or 2 and are slightly more accurate for judgments in this range. (Based on Brannon and Terrace, 2002, figure 26.5).

come to think that the object-indexing system explains a good amount of data that, at first glance, may have appeared to support the view that infants or animals can represent small precise quantities (see, e.g., T. Simon 1997; Leslie et al. 1998; Uller et al. 1999). An example, though not an uncontroversial one, is the proper analysis of Wynn's (1992c) classic addition/subtraction study with infants. Wynn showed five-month-old infants simple arithmetic events and measured their looking time for correct and incorrect results. For instance, in a 1+1 scenario, infants saw a single doll placed on a stage, followed by a curtain rising and blocking the view of the stage. A second doll was then visibly placed behind the curtain. Finally, when the curtain fell, the infants saw either the correct outcome (two dolls) or an incorrect outcome (one or three dolls). Because infants looked significantly longer at the incorrect outcomes, Wynn concluded that they can do simple arithmetic. Wynn's results have been replicated many times, and variations on the same basic procedure have been successful with monkeys and dogs (Hauser et al. 1996; Uller et al. 2001; West and Young 2002). But many theorists have felt that her reading of the data is too extravagant and that there is no need to suppose that infants are representing numerical quantity or any arithmetic facts. Perhaps instead a better explanation can be given directly, in terms of the object-indexing system. For example, in the unexpected outcome of 1+1=1, infants have an active index that is missing its object, and this may produce greater demands on attention, causing infants to look longer (Leslie et al. 1998). For theorists who are skeptical of strong nativism but who aren't necessarily empiricists, appeals to the object-indexing system have seemed quite attractive. Like the accumulator, the object-indexing system involves a limited amount of innate structure but no innate representations for specific natural numbers—in this case, no representations for the smallest natural numbers.

1.2 *Strong Nativism Reconsidered*

In our view, the retreat from strong nativism was too hasty. Weak nativism faces a rather serious difficulty that only serves to highlight the explanatory power of strong nativism. This is the challenge of explaining how precise numerical concepts are learned given the meager innate resources that weak nativists acknowledge.

In general, there is something puzzling about how one can acquire a system of representation that is richer than the one in which its learning takes place (Fodor 1975, 1981; Niyogi and Snedeker forthcoming). The very idea of learning fundamentally new concepts has an air of mystery about it. Unlike Fodor, we don't want to say that concept learning is simply impossible (Laurence and Margolis 2002 and forthcoming). However, it certainly is true that there is a substantial explanatory burden associated with proposals for learning new concepts, and number concepts are a case in point. As Stanislas Dehaene has put the point, "[I]t seems impossible for an organism that ignores all about numbers to learn to recognize them. It is as if one asked a black-and-white TV to learn about colors!" (Dehaene 1997, pp. 61–62). Weak nativists aren't unaware of the difficulty. They have debated the relative importance of the accumulator and the object-indexing system and have speculated about how these systems might support the acquisition of the natural numbers. They have also suggested that natural language may play an important role, perhaps even an essential role (Dehaene 1997; Gallistel and Gelman 2000; Carey 2004; Spelke 2003), in the acquisition process. While this work has in many ways been extremely fruitful, weak nativist models have, by and large, been short on details at just the point where they are supposed to explain how concepts for the natural numbers emerge from prior systems of representation (Laurence and Margolis 2005 and in prep.). In contrast, strong nativist models are far better equipped to provide a fully explicit and satisfying account, precisely because such models help themselves to more innate structure than is permitted within a weak nativist framework.

Of course, strong nativism comes in different varieties, just as weak nativism does, and some of these will be more plausible than others. The essential difference between strong and weak nativism is that strong nativism takes at least some natural number concepts to be innate. So one needn't adopt all of the commitments of Gelman and Gallistel's (1978) model to be a strong nativist; for example, one might suppose only a few natural number concepts are innate, or one might hold that the innate system of representation doesn't embody the counting principles. We will briefly sketch our own version of strong nativism as an alternative and (we believe) more plausible strong nativist position.

On our model, one of the core systems supporting natural number concepts is an innate *number module*.[3] The number module, as we construe it, contributes a small set of representations that correspond to the first few natural numbers—1, 2, 3, and perhaps 4. These representations have precise numerical content, but it's fairly minimal. They

3. See B. Butterworth (1999) for a related view, though Butterworth motivates and develops the idea of a number module in a different way than we do.

needn't carry with them an understanding of the quantitative relations among small collections or knowledge of mathematical facts and operations. In fact, as far as our model is concerned, the number module's representations needn't even be understood to be ordered. What makes them numerical representations is just that they serve to detect collections of specific sizes, for example, the representation corresponding to 2 is uniquely responsive to collections that have precisely two items, independent of whatever non-numerical properties the collections have. How might the number module be implemented? One option is that the module takes the form of a neural network that receives input from the object-indexing system and from comparable systems in other modalities. Such a network would have three or four output nodes, and its connections would be weighted so that each of these output nodes responds selectively to a particular numerical quantity. One way to accomplish this would be for the input nodes to provide enough activation so that any one of them would suffice to activate the 1 output node, any two the 2 node, and so on, while at the same time having inhibitory links so that each output node inhibits the activation of the output nodes corresponding to smaller numerical quantities. So the 2 output node inhibits the 1 output node, and the 3 output node inhibits both the 1 and 2 output node, and so on.[4] See figure 8.3.

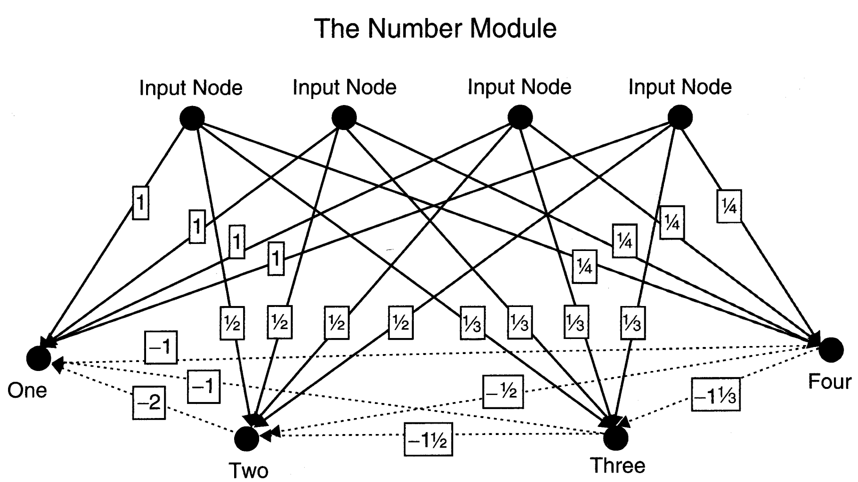

FIGURE 8.3 The number module. The network's input comes from the object-indexing system and from comparable nonvisual systems. The output nodes are selectively responsive to specific numerical quantities.

4. One set of connection strengths that would accomplish this is as follows. Each input node is connected with a strength of 1 to the 1 output node, with a strength of ½ to the 2 output node, with a strength of ⅓; to the 3 output node, and with a strength of ¼ to the 4 output node. In addition, the 2 output node is connected to the 1 output node with a strength of −2. The 3 output node is connected to the 1 output node with a strength of −1, and to the 2 input node with a strength of −1½. The 4 output node is connected to the 1 output node with a strength of −1, and to the 2 input node with a strength of −½, and to the 3 output node with a strength of −1⅓.

For our purposes, what makes the number module's representations numerical is that they are fully abstract (they aren't tied to a single modality, such as vision) and they function to correspond specifically to the number of things in a collection (as opposed to individual objects or non-numerical properties). In addition, they are precise, unlike the accumulator's mental magnitudes. This combination of features allows children to precisely represent the first few natural numbers, providing an effective starting point for acquiring the full system of natural numbers.

How do children get beyond this fairly minimal base to acquire concepts of natural numbers beyond 3 or 4? One possibility is that an external structured symbol system helps children to extend the innate system. The external symbol system might be a natural language counting system, though in principle it could just as well be a system based on body parts, written arithmetic symbols, marks, or other external symbols. To illustrate how the external system might help, imagine that children are able to detect the properties *one, two,* and *three* through the representations in the number module and that they map these directly to the words "one," "two," and "three." At this point, they needn't see these words as being part of an ordered system. They just hear the words used independently of one another and associate them with the properties that they correspond to, just as they would in learning any other individually presented words or symbols. Suppose as well that children learn the counting routine as a kind of game, only to discover that for the small count words the last word reached in a count happens to be the word that expresses the quantity of the collection. This allows children to determine that counting is a way of enumerating and to interpret the first few count words in terms of their innate numerical representations. Because of the newly acquired mapping between the innate numerical concepts and the first few words in the count sequence, children would then have a way of placing the concepts in order, even if they don't yet fully understand the quantitative significance of that order. What's more, because they can represent *one* (again, via the number module), they are in a position to detect the single most significant fact about that order. They can determine that the quantity associated with each subsequent term (for the first few terms) is exactly one more than the quantity associated with its predecessor. Finally, they can inductively infer that every term in the sequence, not just the first few terms, participates in the same pattern—each expresses a quantity that is exactly one more than the preceding term. This, in barest outline, is how children might come to acquire concepts for natural numbers according to our own strong nativist account (see Laurence and Margolis in prep. for more details). The cornerstone of the account is the innate number module, which allows children to represent small numbers with precision, especially *one*, giving them a foothold for acquiring further natural numbers.[5]

5. Leslie, Gallistel, and Gelman (this volume) present what we take to be another strong nativist alternative to weak nativism. This represents a radical reorientation from Gallistel and Gelman's recent work (e.g., Gallistel and Gelman 1992, 2000; Gallistel, Gelman, and Cordes 2005). Like us, Leslie, Gallistel, and Gelman (this volume) argue that an innate ability to represent a difference of 1 is essential for acquiring the integers; however, they employ a much higher standard for the conditions that must be met to possess numerical concepts (see their discussion of the computational compatibility constraint).

The dispute about the cognitive development of numerical concepts isn't about whether they have an innate basis but about how much innate structure is involved and whether, and to what extent, it is number-specific. The attraction of strong nativism, we've been suggesting, is that by helping itself to more innate structure than weak nativism, it is able to give a far more explicit account of the development of numerical concepts. But to maintain this advantage, strong nativists have to reject the Cultural Construct Thesis. In the next section we'll look at a body of recent cross-cultural data that would appear to support the Cultural Construct Thesis and hence provide a serious challenge to strong nativist models like our own.

2 The Whorfian Challenge of the Pirahã

In a highly influential recent study, Peter Gordon investigated the numerical abilities of the Pirahã, a tribe in a remote region of the Brazilian Amazon. Gordon's own interests in this group stem from his views about linguistic determinism. As Gordon puts it, the issue here is whether the absence of relevant linguistic structures, such as words and grammatical devices, "precludes the speakers of one language from entertaining concepts that are encoded by the words or grammar of [another] language" (Gordon 2004a, p. 496). Gordon sees the Pirahã as offering an ideal case study because they speak a language that differs from most familiar languages in that it has a paucity of words for expressing numerical quantity. Moreover, the few numerical words that the Pirahã language does have fall short of expressing precise numerical quantities. Gordon's claims regarding the Pirahã language largely derive from work by the linguists Daniel and Keren Everett, who are among the foremost authorities on the Pirahã language and culture. (The Everetts have lived and worked among the Pirahã for over twenty years, and it was their research team that facilitated Gordon's own studies with the Pirahã.) Gordon notes that the primary candidates for number words in the Pirahã language are "hói," "hoí," and "baagi" (or "aibai"), corresponding to "roughly one," "roughly two," and "many." Crucial to Gordon's analysis is that these terms lack the precision associated with natural number concepts (Gordon 2004a, p. 498):

> One particularly interesting finding is that "hói" appears to designate "roughly one"—or a small quantity whose prototype is one.... In Pirahã "hói" can also mean "small," which contrasts with "ogii" (=big), suggesting that the distinction between discrete and continuous quantification is quite fuzzy in the Pirahã language.

In addition, despite occasional trading relations with nearby Brazilians, the Pirahã don't use money and haven't adopted Portuguese counting words. In part, this is because the Pirahã maintain a strong isolationist cultural identity. According to Daniel Everett, "the Pirahã ultimately not only do not value Portuguese (or American) knowledge but oppose its coming into their lives" (Everett 2005, p. 626). The interesting question, then, is whether the Pirahã, despite their lack of counting terms, have the cognitive capacity to represent and manipulate exact numerical quantities. Gordon argues that they do not and that this fact provides direct support for a strong form of linguistic determinism.

Not too long ago linguistic determinism had few supporters in cognitive science. Linguistic determinism has always been associated with Benjamin Lee Whorf's rather naïve analysis of Native American languages. Once this analysis was discredited, linguistic determinism itself came into disrepute (see, e.g., Pinker 1994). Recently, however, linguistic determinism has been making a comeback, and there has been a resurging interest in the kind of sustained cross-cultural work that would be needed to test it (Gumperz and Levinson 1996; Gentner and Goldin-Meadow 2003; Levinson 2003). Gordon's study has certainly contributed to the revitalization of linguistic determinism and is a particularly important case study given its prominence in the literature. (Gordon reports his data in the prestigious journal *Science*, which is itself a good indication that linguistic determinism has regained a good deal of scientific respectability.)

We share Gordon's interests concerning both the status of linguistic determinism and the more specific claim that precise numerical concepts depend on language. But for the purposes of this chapter, we also want to place his study in a larger context that asks about the innate basis of precise numerical concepts. Undoubtedly, part of the reason why Gordon's study has received so much attention is that it strongly suggests that precise numerical concepts are a cultural construct. If precise numerical concepts are so dependent on contingent features of language, then there is no reason to suppose that they have a specific innate basis. Rather, the supposition must be that they are learned by exposure to the cultural practices that only certain languages embody and, consequently, that the strongest viable form of nativism is weak nativism. In the next section, we will take up both of these issues—the status of Gordon's claims about linguistic determinism and the implications for the nativism dispute. But before turning to our own assessment of what can be concluded from Gordon's data, it will be helpful to review his experimental procedures and to briefly describe the results as Gordon himself sees them.

Gordon reports data from eight experimental tasks, conducted on seven adult Pirahã subjects (six male, one female) ranging in age from 18 to 55.[6] The first six tasks all have a similar structure in that they ask subjects to produce an array of items that match the number of a target display. The two remaining tasks involve keeping track of the number of items placed into an opaque container or using the number of symbols on the outside of a container to distinguish it from another container. The tasks were designed to place varying demands on different cognitive skills that interact with numerical abilities. For Gordon, the question is whether any patterns emerge across these variations. He reports that there is indeed a crucial pattern. While his subjects were relatively successful with small numbers of items (up to two or three), their performance significantly decreased with larger numbers. Moreover, the variability in their estimates tended to increase as the quantities increased—a pattern that is suggestive of the use of mental magnitudes or accumulator-based representations. From all of this, Gordon concludes that the Pirahã's linguistic system confines them to analog representations of numerical quantity. "[T]hese studies show that the Pirahã's

6. Most of Gordon's experimental data are drawn from four of the seven participants. Gordon notes that the six adult males comprised all of the adult males in the two most accessible Pirahã villages and that children and women were generally inhibited from participating in experiments.

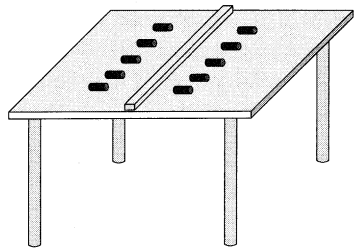

FIGURE 8.4 The one-to-one line match task. The experimenter arranges a linear array of items on one side of the table, and the subject's task is to place the same number of items on the other side.

impoverished counting system limits their ability to enumerate exact quantities when set sizes exceed two or three items" (Gordon 2004a, p. 498).[7]

Here, then, is a brief summary of Gordon's experiments.

One-to-One Line Match. Let's start with the group of tasks that require matching a target display's number. The first of these gives much of the flavor of the whole group. It's called the *one-to-one line match* task. Here's how it works. The experimenter and the subject sit on opposite sides of a table, and the experimenter lays down an array of batteries in a horizontal evenly spaced configuration. The subject is then asked to place batteries on his side of the table to "make it the same" (see figure 8.4). In effect, the task is to line up the batteries in one-to-one correspondence with the experimenter's array. Once the subject is finished, he is asked whether it is the same before being tested on another number. Regardless of the outcome, the experimenter always gives the cheerful response "aiyo!" (which is comparable to saying "OK!") and then proceeds with the next trial. Gordon made sure to always start with small quantities but then tested larger numbers (up to 8 or 9) in random order, with each tested quantity appearing two or three times.

Cluster Line Match. For this task the target group of items to be matched is a nonlinear configuration of nuts. As with the one-to-one line match task, the goal is to construct a linear horizontal array of batteries with the same number ("make it the same"). But since the nuts are not the same size as the batteries, this task can't be

7. Gordon sometimes talks as if the pertinent issue of linguistic determinism is whether the Pirahã can represent precise quantities greater than 2, which suggests that he may think that they can, at least under some conditions, represent exact quantities of 2 or less. However, when he talks about the ability to represent small exact quantities, he tends to align himself with work that identifies this ability with mechanisms of object-based attention, which do not employ specifically numerical representations (Gordon 2004a, p. 498). And since he maintains that the Pirahã don't have any words for precise numerical quantities (not even for 1), Gordon's linguistic determinism implies that the Pirahã shouldn't have *any* concepts for exact numerical quantities. For these reasons, we read Gordon as holding that the Pirahã aren't capable of representing any precise numerical quantities, not even quantities as small as 1 and 2. But in the end it doesn't matter whether Gordon himself goes this far, since it is clear that others opposed to strong nativism suppose that Gordon's data make a powerful case against there being any innate integer concepts.

solved merely by attending to the amount of stuff to be matched. Further, because of the nonlinear arrangement of the nuts, the task can't be solved simply by placing one battery directly in front of each nut.

Orthogonal Line Match. This time the array to be matched is a linear array, but it is positioned perpendicularly to the array that the subject is expected to create; the experimenter's array is vertical, the subject's horizontal. As in the cluster line match, this configuration prohibits the simple strategy of placing a battery directly in front of each target item. Moreover, were subjects to try to solve the task by using an estimate of overall length—another non-numerical strategy—there would be a telltale sign. Since vertical lines appear longer than same-sized horizontal lines, the reliance on mere length would cause subjects to overestimate the number of items needed to match the array, and they would end up placing too many batteries in their horizontal arrays.

Uneven-Line Match. This time the array to be matched is linear and horizontal but with different-sized gaps between the batteries that compose the array. In other words, the task is just like the original one-to-one line match task except that in the original task the batteries are evenly spaced.

Line Copying. This task differs from all of the previous ones in that a notepad is used. On one side of the pad's binding there is a horizontal array of lines. Subjects are expected to match the number by drawing lines on the other side. Visually, this looks as if you are extending the horizontal array. Part of the reason for this variation is the novelty of drawing for the Pirahã; drawing isn't a familiar activity for them. Also, the arrangement of the pad offers another variation where the task can't be solved by using the simplest non-numerical strategy—the new lines can't be placed one-for-one directly in front of the lines being matched.

Brief Presentation. The final experiment of the group is just like the cluster line match except that the array to be matched is visible for only a brief period—approximately one second. As a result, the matching procedure has to be done from memory. Gordon doesn't explain why he included this variation, but presumably the memory limitation further discourages the sorts of non-numerical strategies that motivated the cluster match task in the first place. Gordon reports that several subjects probably approached the cluster match task by positioning each battery to point to an individual nut. "Such a targeting strategy would be very familiar from their everyday use of bows and arrows for hunting and fishing" (Gordon 2004b, p. 4). This strategy would be far less effective, however, once the nut cluster is out of sight. One would have to remember the position of each nut, and that is not easy for larger arrays. Gordon also sees some value in manipulating requirements on memory as a way of testing numerical competence under varying task demands. "Any estimation of a person's numerical competence will always be confounded with performance factors of the task. Because this is unavoidable, it makes sense to explore how performance is affected by a range of increasingly demanding tasks" (Gordon 2004a, p. 497).

Of all the matching tasks, the results for the one-to-one line match were the best. Pirahã subjects created perfect matches for numbers in the 1–3 range and were suc-

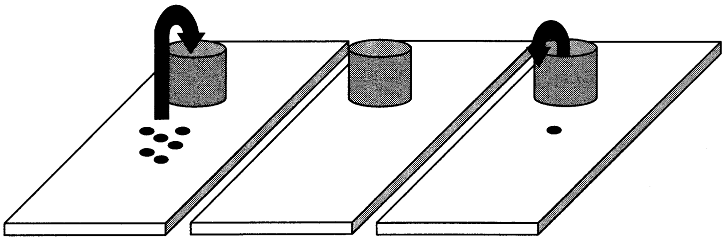

FIGURE 8.5 The nuts-in-can task. A group of nuts is shown for eight seconds and then placed into a can. As each nut is removed, one at a time, the task is to say whether any nuts are left in the can.

cessful about 75% of the time for numbers from 4–8. The overall trend for the other matching tasks was similar, in that their performance deteriorated as the numbers grew larger, but with some of the other matching tasks, their performance deteriorated even before getting to 3 and dropped well below the 75% success rate. For example, in the orthogonal line match they were at 100% for 1 and 2, but dropped to about 60% for 3–8, and didn't succeed at all for 9. In the brief presentation task, they were also at 100% for 1 and 2, but dropped to 75% for 3, and then below 50% for 5–9. The one exception to the general trend was the uneven-line match task (the one where they were supposed to match a horizontal array of batteries that were unevenly spaced). In this case, though performance dropped for 5 and 6, it rebounded for 7 and 8. Gordon notes that the reason for this exception is probably that the subjects were able to chunk the items in the larger arrays and then exploit their superior ability for matching small arrays.[8]

All of the tasks we've reviewed so far are variations on a theme. The goal is to create an array that matches the number of items in the experimenter's array. Gordon's two other tasks employ significantly different strategies for gauging numerical competence.

Nuts-in-Can. The first of these begins with a cluster of nuts laid in front of the subject for eight seconds. Then the nuts are placed in an opaque container (an oatmeal can). One by one the nuts are removed from the container and each time the subject is asked whether there are any nuts left. The inside of the container is revealed when the subject declares that the container is empty or once all of the nuts have been removed. (See figure 8.5.) The results for this task were similar to the various matching tasks in that set size was a major determinant of success. Pirahã subjects did poorly with larger numbers, but in this case they also had considerable difficulty with small numbers, achieving less than 100% success even for two nuts (the smallest number tested) and less than 75% success for three. For 5–9, their performance dropped even further, to below 50%.

Candy-in-Box. The last of Gordon's experiments also uses containers, this time cassette cases, each of which is covered by a picture depicting from one to six fish. The

8. The other results were as follows: For the cluster line match, 100% for 2 and 3, 75% for 4–8, and 0% for 9 and 10; for the line copying, 100% for 1 and a precipitous drop from 75% to 0% for 2–7.

subject sees a candy being placed in a case that is subsequently hidden behind the experimenter's back. Then the experimenter brings his hands forward holding two cassette cases—the original and a second case whose picture has precisely one more, or precisely one less, fish on it. The goal is to pick the case with the candy, and the candy is given as a reward for choosing correctly. Pirahã subjects did poorly on this task as well, falling below chance for some comparisons and never achieving much above 75% accuracy even for the smallest number comparisons—1 versus 2 and 2 versus 3.

For Gordon, the results of these eight tasks all point in same direction. "The present experiments allow us to ask whether humans who are not exposed to a [linguistic] number system can represent exact quantities for medium sized sets.... The answer appears to be negative" (Gordon 2004a, p. 498). And while the Pirahã may appear to have some limited ability to discriminate between sets with two or fewer items, Gordon takes this to be a reflection of the object-indexing system, not a truly numerical ability. As a result, Gordon's position amounts to the claim that the Pirahã's only numerical abilities are the approximate ones that are grounded in the accumulator and that humans share with many other animals. If Gordon is right, then concepts for precise numerical quantities would appear to be a cultural achievement, just as the Cultural Construct Thesis says. What's more, it's a cultural achievement that is only possible given the right sort of language. The reason why the Pirahã have no access to the precise numerical concepts that most of us take for granted is that their language doesn't allow them to entertain these concepts.

3 Critique of Gordon's Study

It's easy to see why Gordon's study of the Pirahã has attracted so much interest. Our facility with numbers appears to be a distinctively human characteristic and one that underlies many significant features of human life. If it turned out that a group of otherwise normal, intelligent human beings were incapable of entertaining precise numerical thoughts, and if it turned out this was because of contingent features of the language they speak, it would be big news. Both the Cultural Construct Thesis and the thesis of linguistic determinism, if true, would have deep implications for our understanding of the mind. Unfortunately, the experiments that Gordon has carried out aren't helpful for settling any of these important issues. The experiments suffer from a number of flaws that make it impossible to draw any meaningful conclusions about the Pirahã's numerical abilities, much less the relation between language and thought. In this section, we explain why. We begin with some general objections to Gordon's argument for linguistic determinism. We then take a close look at the details of his experiments, registering a series of objections that bear on both the status of linguistic determinism and the case against strong nativism.

3.1 *Correlation vs. Explanation*

For the moment, let's put aside the issue of strong nativism and focus just on the question of whether Gordon's study supports linguistic determinism. For the sake of argument, we will assume that Gordon's analysis of the Pirahã language is accurate and that the Pirahã have a significant linguistic deficit—in particular, that they have no linguistic expressions for precise numerical quantities. We will also assume that Gordon's study shows that the Pirahã have a significant cognitive deficit—namely, that they are unable to *conceptualize* precise numerical quantities. (Later, starting in section 3.3, we'll challenge the second of these assumptions, but for the time being we don't want to get embroiled in the details of Gordon's experiments.)

Given these assumptions, does Gordon's thesis of linguistic determinism follow? Unfortunately, no. The problem is that the most that can be concluded is that the linguistic and cognitive deficits are *correlated* in the Pirahã. However, as a number of critics of linguistic determinism have noted in other contexts, it's one thing to establish that linguistic and cognitive deficits are correlated, but quite another to show that the linguistic deficit *is responsible for* the cognitive one (Bloom and Keil 2001; Gleitman and Papafragou 2005; Pinker 1994). To establish the responsibility claim, a lot more would have to be done. Gordon would have to rule out the possibility that the determination relation goes in exactly the opposite direction. After all, it could be that the reason the Pirahã lack words for precise numerical quantities is because they lack concepts for precise numerical quantities, not the other way around. Similarly, Gordon would have to rule out the possibility that the conceptual deficit traces back to some other factor that has nothing to do with language. Reasonable alternatives of these sorts clearly need to be considered, yet Gordon's study fails to do so. Indeed, it's hard to see how the measures he employs could even begin to locate the source of the Pirahã's difficulties with numerical quantity, since Gordon's tests only presume to examine the numerical abilities themselves. The most they could tell us is whether a subject is capable of precisely enumerating a collection, not what prevents him from enumerating it if he can't.

To see the burden that Gordon faces, it might be helpful to say a little bit more about some of the competing explanations of why the Pirahã lack concepts for precise numerical quantities. One type of explanation appeals to cultural factors (i.e., cultural factors apart from language). We have already noted that the Pirahã have a strong identity as a people and are highly resistant to outside cultural influences. Everett (2005) characterizes their culture as one that places special significance on personal experience of the here-and-now and that has a corresponding indifference to abstractions. Everett speculates that this culturally based belief system ends up constraining how the Pirahã think and communicate and that this in turn is reflected in their language. If Everett is right, then the Pirahã should be expected to have difficulties with all sorts of abstractions, including numerical quantity, but the difficulty would trace back to their cultural outlook, not to an inherent limitation of their language. Though we ourselves are somewhat skeptical about the claim that the Pirahã are fixated on the concrete present, the general strategy of locating a cultural source of their difficulties with number isn't implausible. For example, one can easily imagine that a latent ability to represent precise quantities might be lost owing to lack of use. If the Pirahã simply fail to nurture and exercise this ability, then perhaps that is why they do so poorly on Gordon's numerical tasks. It's also not that hard to imagine other cultural factors that might be responsible. For example, the problem could be that the Pirahã aren't trained in a counting procedure and that learning concepts for natural numbers, especially larger ones, is inordinately difficulty without such a procedure. Note, however, that this explanation needn't invoke natural language, since counting itself needn't involve words; as we noted earlier, it can be based instead on body parts, tallies, or other types of external symbols.[9]

Because these alternative explanations invoke cultural practices, they might be thought to challenge Gordon's linguistic determinism at the cost of leaving the Cultural Construct Thesis perfectly intact. Of course, at this point we are simply accepting for the sake of argument that the Pirahã really do have the cognitive deficits Gordon claims they have, something we will be challenging shortly. But even if we grant that the deficits are real, there are further possibilities that have nothing to do with cultural practices. One of these is that the Pirahã, or the few subjects Gordon tested, suffer from a genetic anomaly. Gordon reports that there is no reason to suppose that his subjects were psychologically impaired, and Everett, who has lived among the Pirahã, flatly rejects the suggestion that the Pirahã have genetic defects, noting that they intermarry with outsiders (Everett 2005, p. 634). But the claim regarding intermarriage has to be taken with a grain of salt. First, despite having some contact with outsiders, the Pirahã remain a very small community, largely isolated from neighboring groups. Second, as Everett himself points out, the Pirahã's marriage system is "relatively unconstrained" in that it isn't unusual for Pirahã couples to share at least one biological parent (Everett 2005, p. 632). So while we wouldn't want to just jump to a genetic explanation, this possibility should not be ruled out a priori either. Clearly, a genetic explanation would have to be considered if the population in question were located in Boston or Chicago. We see no reason to think that things ought to be different for the Pirahã just because they are located in a more remote part of the world.[10]

The upshot of these considerations is that even if we take all of his results at face value, Gordon's experiments provide little or no support for linguistic determinism. On the face of it, there are any number of equally plausible hypotheses for why the linguistic and cognitive deficits might be correlated in the Pirahã. And since the most that Gordon's study could establish is that these deficits are correlated, it cannot even begin to rule out any of these alternatives.

3.2 *A Very Weak Correlation*

So far we have been supposing that Gordon's study shows that a conceptual deficit (the inability to think in terms of the natural numbers) is at least correlated with a linguistic deficit (a paucity of number words). We have only claimed that Gordon's study can't elucidate why the correlation obtains. In this section we want to go one step further by challenging the claim that there really is a meaningful correlation.

9. As it turns out, Everett did try to teach the Pirahã how to count using a linguistic counting system. His efforts were unsuccessful. See section 3.7 (below) for discussion of the significance of this outcome.

10. None of this implies that the Pirahã (or any other traditional people) are generally intellectually inferior. Genetic deficits can be quite focused and needn't involve general cognitive impairments. Exploring the possibility of a genetic anomaly would be no more presumptuous in this case than in other cases where a circumscribed cognitive deficit has been discovered (e.g., familial Specific Language Impairment).

Gordon's study, as it turns out, offers little evidence that there is. All he gives us is a *single* case study involving just *one* population—one data point, as it were—and this single case study is based almost entirely on just four subjects (see note 6)! Of course, sometimes a conclusion can be warranted on the basis of a single case study, even of a small population, but not when the issue is a sweeping claim like Gordon's thesis of linguistic determinism. It's one thing to say that the Pirahã, who happen to lack number words, are unable to solve certain tasks that require the use of precise numerical concepts. It's quite another to say that *in general* the representation of precise numerical quantities requires the linguistic means to express them and that numerical concepts are essentially dependent on number words. In order to justify these broader claims, additional case studies are absolutely essential. We need to look at other populations that also have a paucity of number words. The pressure to look in this direction increases all the more so when we recognize that Gordon's linguistic determinism is built around the finding that the Pirahã don't succeed on his numerical tasks—a negative result. What if other similar populations, or even just a single one, were to demonstrate precise numerical abilities despite a lack of number words? This in itself would overturn Gordon's negative finding, showing that precise numerical cognition isn't dependent on language after all. For this reason, it's extremely important to look beyond the Pirahã for further data points before drawing any conclusions about linguistic determinism.

Although Gordon doesn't discuss any other cross-cultural work, there is a body of earlier research that bears on the topic. As R. Gelman and Butterworth (2005) point out, much of this earlier research suggests that the link between language and number is nowhere near as tight as Gordon claims. For example, Australian Aborigine speakers of Warlpiri (a language similar to Pirahã in its paucity of number words) show no evidence of lacking numerical concepts. The linguist Robert Dixon offers this summary of what has been observed of Warlpiri speakers and other Aborigines (Dixon 1980, p. 108):

> [N]o special significance attaches to the absence of numeral systems in Australian languages; it is simply a reflection of the absence of any need for them in traditional culture. Aboriginal Australians have no difficulties in learning to use English numerals; Kenneth Hale has commented that "the English counting system is almost instantaneously mastered by Warlpiris who enter into situations where the use of money is important. . . ."

Likewise, Everett has observed that other Amazonian tribes freely borrow number words from their Brazilian neighbors when the need arises (Everett 2005, p. 634). Susan Schaller discusses a similar sort of case, involving a deaf adult who managed to function without having acquired a natural language by miming to communicate. According to Schaller, he readily learned the Arabic numerals for 1 to 20, matching these to corresponding sets of crayons. "He found the symbols for numbers easy compared to signs or words. Apparently, arithmetic already resided in his brain" (Schaller 1991, p. 61).

Admittedly, all of these claims are based on unsystematic observations, not carefully designed experiments, so it's fair to wonder exactly how precise and accurate the stated numerical competences are. It's also possible for Gordon to respond that in most cases the numerical ability comes only with the linguistic ability and that this is consistent with the numerical ability still being dependent on the linguistic one. However, the plausibility of such a response would depend on just how readily the numerical ability takes hold. If it is truly acquired "almost instantaneously," then it seems far more plausible to suppose that the numerical ability doesn't depend on language and that any newly acquired linguistic counting system simply provides evidence for the prior possession of numerical concepts.

There are also other relevant data to take into account (R. Gelman and Butterworth 2005). For example, Dixon reports an Aboriginal practice of using different parts of the palm to indicate the number of days until a planned event occurs, a system that apparently doesn't require possession of number words. And clinical studies have shown that precise numerical abilities can be preserved despite severe linguistic deficits and, in some cases, may develop without them as well. Hermelin and O'Connor (1990) describe a particularly impressive case of a speechless autistic man who can identify five-figure prime numbers and factorize numbers of the same magnitude, all based on exposure to a few examples expressed in standard Arabic notation (i.e., as opposed to natural language). (For work on aphasiac patients see, e.g., Rossor et al. 1995; and Varley et al. 2005.)

In sum, it shouldn't be granted that Gordon's study of the Pirahã establishes a genuine correlation between linguistic and numerical abilities. While he may have identified one instance where a population lacking number words also happens to lack precise numerical concepts, it is only one instance.[11] A broader examination of the evidence suggests that the pattern may not hold up elsewhere and certainly raises questions about whether we should expect it to. Taken together these points considerably weaken Gordon's case for his linguistic determinism. What they show is that even if we take him to have demonstrated that the Pirahã lack precise numerical concepts, Gordon provides little or no evidence that such concepts are dependent on language.

3.3 *A Null Effect*

We've seen that Gordon's results with the Pirahã don't suffice to establish his thesis of linguistic determinism. Even if we take him to have shown that the Pirahã can't represent precise numerical quantities, this may not be the result of the language they speak, and it doesn't tell us anything about the general class of cases where

11. A related study by Pica et al. (2004) was published in the same issue of *Science* as Gordon's article. This study examined the numerical abilities of the Mundurukú, another Amazonian tribe whose language has a highly impoverished vocabulary for numerical quantities (the Mundurukú language only has fixed terms for quantities up to 5, and none of these expresses a single precise numerical quantity). Pica et al. don't make the same strong claims concerning linguistic determinism as Gordon does. Nonetheless, they argue that the Mundurukú are incapable of precise numerical thought. Pica et al.'s test for precise numerical thought involved showing subjects a video of a certain number of dots (1 to 8) going into an opaque container that was previously shown to be empty. After a pause, some of the dots would exit the container. Subjects were then asked to say how many dots remained or to choose which of three images of a container (with zero, one, or two dots inside) depicted the correct result. There is much to admire about Pica et al.'s research, including their attention to non-numerical confounds and their use of French-speaking controls who were given the very same tasks as the Mundurukú. But in spite of these virtues, their study is subject to a number of the same criticisms that we will raise for Gordon's study—see notes 13, 15, and 19 below.

people speak a language that lacks words for precise numerical quantities. Still, Gordon's study might be thought to be somewhat suggestive. After all, he does seem to locate a population where a paucity of number words is associated with a conceptual system that doesn't register precise numerical quantities. And while it remains to be seen whether the same association holds up elsewhere—and, if so, why it does—Gordon's experiments might be thought to provide the first steps in a more encompassing research program as well as an experimental framework for investigating these questions. In much of the rest of this section (8.3.3–8.3.6), we will argue that this would be a mistaken view of the situation. Quite surprisingly, Gordon's experiments do not license any substantive conclusions about whether the Pirahã are capable of precise numerical thought; his experiments turn out to be a poor tool for gauging whether they have precise numerical concepts. If we are right about this, then Gordon's results with the Pirahã don't provide *any* support for either linguistic determinism or the Cultural Construct Thesis.

We'll begin with one of the most significant problems that we see with Gordon's experiments, a feature of his experimental procedures that affects nearly all of his tasks. This is that they are designed to elicit spontaneous responses and only spontaneous responses. In general there is nothing wrong with looking at spontaneous responses. If a group is given a numerical task and happens to respond correctly without any training or guidance, this would be an excellent sign that they have the mathematical concepts in question. The problem only occurs when their spontaneous responses are incorrect, when the result amounts to a *null effect*. What can be concluded then? Very little. The reason is that the negative outcome would be expected not only if they lack the relevant mathematical concepts but also if they have such concepts but don't habitually think in terms of them, or if they simply fail to understand the task. To make matters worse, Gordon's procedures compound the problem by reinforcing incorrect spontaneous responses rather than helping his subjects to appreciate what a correct response would entail.

One way to get a feel for this objection is to consider how Gordon's experiments look on the assumption that the Pirahã *are* capable of precise numerical cognition but that precise numerical quantities aren't salient in their culture. In that case, how would we expect Pirahã subjects to perform on Gordon's tests? Consider the various matching tasks. In these tasks, subjects are shown an array of batteries or nuts and told to *make it the same*. But what would "making it the same" mean to them? Presumably, if precise numerical quantities aren't salient for them, then their initial interpretation of the instructions wouldn't be to match the precise numerical value of the target array. Perhaps, instead, they'd suppose that they are meant to match the approximate numerical value, or match the total amount of stuff, or create a similar-looking visual pattern. Regardless of which it is, there would be nothing during the course of the experiment to cause them to revise their initial view of how to proceed. After all, whatever they do, Gordon replies with a cheerful "aiyo!" ("OK!"). The result is that no matter what interpretation they start out with, and no matter how misguided it may be, they are met with encouragement that tells them to continue in the same way. Under these conditions, we should certainly expect to see just the sort of poor results Gordon obtained. But if the Pirahã should fail Gordon's tests *even when it's assumed that they are capable of precise numerical thought*, the failure that Gordon documents cannot establish that such thought is beyond them.[12]

An analogy may help to clarify the situation. Suppose we were dealing with entirely nonlinguistic subjects—for example, chimpanzees—and we wanted to see if they are capable of enumerating precise numerical quantities. The challenge, of course, is to convey to an animal how to approach a task that measures this ability without having the luxury of being able to verbally state the instructions. Now imagine a scientist, like Gordon, starting with a small collection and then rewarding a chimpanzee for its spontaneous response no matter what it does. To keep things simple, we can suppose that the chimpanzee behaves appropriately—the experimenter places one item down, and the chimpanzee places one item down as well, or the experimenter places two down, and the chimpanzee responds with two. From here, the experimenter goes on to larger sets in random order, and the chimpanzee fails in one way or another to match the target number. It should go without saying that it would be irresponsible, on the basis of this outcome, to conclude that chimpanzees are *incapable* of enumerating precise quantities. In fact, a null effect under these conditions wouldn't be considered a publishable result. The experimenter simply hasn't done the necessary legwork to draw such a strong conclusion from the negative finding. With animals it's plainly obvious that we need to train them on a task to see what they are capable of, and that such training can sometimes take a substantial amount of time and effort. The true test of their abilities is not their spontaneous response on a task that they initially may not understand, but their behavior toward *novel stimuli* that are relevantly similar to the ones they have been trained on. Although the Pirahã aren't themselves nonlinguistic subjects, their situation is similar to the chimpanzee's in that the nature of Gordon's tasks can't be easily and directly conveyed to them verbally (by hypothesis, the Pirahã lack the needed vocabulary). But then, just as with the chimpanzee, there is no point in testing the Pirahã subjects until a serious effort has been made to fully convey what they are supposed to do. One obvious way to address this objection would be to adapt the standard procedures that are employed in comparative psychology, including the use of a battery of pretest trials and a system of

12. The situation is slightly more complicated when we turn to Gordon's nuts-in-can and candy-in-box tasks—the two that aren't simply variations on the one-to-one matching task. While there are a number of alternative interpretations of these tasks as well (e.g., the Pirahã might simply have seen them as guessing games), subjects did receive a certain amount of feedback on these tests. In the nuts-in-can task, subjects were shown whether or not they were right when they said that all the nuts had been removed from the opaque container. And in the candy-in-box task, a reward was built into the task in that subjects were given the candy when they correctly selected the box that contained it. None of this really helps, however, because there are serious questions about whether Gordon's Pirahã subjects understood what was expected of them despite the additional information that they received (see sections 3.5 and 3.6, below). This concern could have been obviated if Gordon had trained his subjects using a series of pretest trials and more explicit feedback about whether they were answering correctly.

rewards and penalties that are enforced until a criterion of success is reached. Gordon, however, did none of this. He simply recorded his subjects' spontaneous responses and left it at that.[13]

We should emphasize that our objection here is not that substantive conclusions can't be drawn on the basis of spontaneous judgments. Positive results on tasks involving spontaneous judgments can provide excellent evidence for the possession of cognitive capacities. Nor are we making the indiscriminate claim that null effects in psychology are always uninformative. We don't think that. It can often be useful to discover that a population fails a test for a given ability. But the test has to be implemented judiciously, otherwise the failure reflects more upon the method of investigation than upon the participants in the experiments. With Gordon's experiments, we see no reason to suppose that the results do reflect upon the participants—the Pirahã. What Gordon needs, and what he doesn't have, are credible procedures for conveying to his subjects what counts as success on his tasks. For this reason alone, Gordon's study does not support any conclusions about the numerical abilities of the Pirahã.[14]

3.4 *Non-numerical Performance Variables*

We've argued that a null effect on Gordon's tasks tells us little or nothing about numerical abilities of the Pirahã. This problem is exacerbated by the fact that several of Gordon's tasks incorporate irrelevant performance variables that have nothing to do with numerical cognition per se. So even if Gordon did manage, in some cases, to convey to his Pirahã subjects what counts as success, they might still fail for reasons that have nothing to do with a lack of numerical concepts. They might fail simply because the task designs make things unnecessarily difficult.

Take, for example, the matching task where the target array is presented for only one second before being covered up (the brief presentation task). Clearly, the time constraint requires that one memorize the array and then access this memory while constructing the match. But imposing these greater demands on memory doesn't help to clarify the numerical abilities under investigation; quite the opposite, it makes it harder to credit failure on this task to deficient numerical abilities. If the goal is to learn more about whether someone is capable of precise enumeration, then non-numerical factors, like memory load, should be *reduced*, not increased. The whole point of studying the Pirahã is to see if they are capable of going beyond the estimation of approximate quantity and whether they can

13. Pica et al.'s (2004) test of precise numerical ability among the Mundurukú (see note 11 above) is subject to much the same worry. As in Gordon's study, there seems to have been no systematic attempt to convey what counts as success on the task, no training, and no feedback.

14. Given that it would be wrong to conclude that the Pirahã are unable to form concepts for precise numerical quantities, it would also be wrong to conclude that they are unable to form concepts for precise numerical quantities because of their language. Moreover, the logic of the objection suggests that the proper test of Gordon's linguistic determinism is one that allows for, and may depend upon, a considerable amount of training. The only constraint is that the training shouldn't be on a *linguistic system*, particularly one that expresses precise numerical quantities.

make exact numerical comparisons. But by introducing a task that forces them to process the numerical information so quickly, Gordon is clearly encouraging estimation. Even people who do have a conventional counting system would be hard pressed to count out the number of items so quickly, particularly for larger numbers; failure to do so would hardly indicate that they can't represent these numbers.[15]

Another example of a task made unnecessarily difficult is the one where a successful numerical match requires drawing (the line-copying task). Most of us take pencils and paper for granted and are perfectly comfortable drawing conventional representations of such things as people, houses, trees, snowmen, etc. But none of this is true of the Pirahã. As Gordon himself remarks, drawing is completely alien to them. When asked to draw familiar items—animals, trees, etc.—the best they can do is to produce "simple lines without form" (Gordon 2004b, p. 5). Gordon also notes the difficult nature of drawing tasks for the Pirahã. "Producing simple straight lines was accomplished only with great effort and concentration, accompanied by heavy sighs and groans" (Gordon 2004a, p. 306). But if drawing itself is that difficult for the Pirahã, why suppose that their poor results on this task tell us anything about their numerical abilities? If a six-year-old child, with considerably more familiarity with drawing, misrepresents the number of fingers on a normal human hand in a drawing, we don't suppose that this shows that she lacks the concept FIVE. Good tests of numerical ability should minimize such irrelevant task demands.

3.5 *Non-numerical Confounds*

We've seen that there are a number of reasons to be skeptical about Gordon's conclusions regarding the Pirahã's numerical abilities. Just because his Pirahã subjects failed his tasks doesn't mean that they are unable to represent precise numbers. Gordon's tests simply aren't sensitive enough to allow us to draw that conclusion. Interestingly, though, many of Gordon's tests wouldn't allow us to draw any conclusion about the Pirahã's numerical abilities even if he had got the opposite result: that is, even if the Pirahã had passed his tests with flying colors, we wouldn't be able to conclude that they have precise numerical concepts. This is because the tests don't sufficiently

15. A similar problem affects the Pica et al. (2004) study mentioned in note 11 above. In this case, the behavior of the French controls is illuminating, as they were at ceiling for only two of the eight subtractions. For the other six, they performed below 100%, typically between 80 and 90%. While it is unclear exactly why the French controls had difficulty with these mathematical tasks—simple problems involving quantities no larger than 8—some non-numerical features must have made the task challenging. After all, we know that the French controls *are* capable of precise numerical thought! One possible explanation is the rapid movement of the dots in the displays (see http://www.sciencemag.org/content/vol306/issue5695/images/data/499/DC1/1102085s1.mov for a video demonstration of the task). However, if the presentation speed made the task more difficult, this would obscure whether the Mundurukú are capable of using precise numerical concepts. Assuming they don't habitually think in terms of precise numerical quantities, increasing the performance demands in this way would only encourage them to fall back on approximate solutions.

control for the non-numerical properties that reliably correlate with number (e.g., the total volume, surface area, and circumference of the stimuli) and because many of the tests can be passed using relatively simple non-numerical strategies.

As we noted earlier, the concern about non-numerical confounds is a familiar one whenever there is a question about whether a given population has numerical concepts. Psychologists who study infants and animals take great care to isolate the many non-numerical variables to which their subjects might respond. For this reason, it's somewhat surprising that Gordon didn't at least take the precaution of varying his stimuli. Within any given task, he standardly used items of the same basic shape and size, thereby ensuring that number correlated with the total volume of material (among other things). The worry is that subjects might achieve a considerable amount of success on tasks intended to test numerical abilities simply by tracking a non-numerical property like volume. Moreover, it's not hard to see how, in some cases, one could even achieve 100% accuracy on Gordon's tests without representing numerical quantity at all, much less precise numerical quantity. For example, with the basic one-to-one line match task, all you need to succeed is the strategy of placing a new battery in front of each item in the target array. Following this strategy on this task allows subjects to perform exactly as if they were precisely representing the number of batteries in the array. And employing this strategy requires little more than the ability to identify and track the individual objects that are used in the task, something that the object-indexing system can support without the need for any numerical concepts at all. A similar strategy would allow subjects to perform perfectly on all of the other matching tasks, with the exception perhaps of the brief presentation task.[16] In the candy-in-box task, which is ostensibly a more difficult numerical task, the number of fish depicted on two boxes is supposed to allow subjects to determine which box holds the candy. But solving the task only requires keeping track of the configurations of the fish symbols, since each number is perfectly correlated with a simple pattern. For smaller numbers, it's just a matter of noting the difference between a point and a line, or a line and a triangle. For larger numbers the patterns are more complex but the same general strategy would work since the pattern for a given number was never varied. All you need to do is recall, after a brief occlusion, the pattern that was associated with the candy.[17]

So even complete success on the majority of Gordon's tasks wouldn't in itself indicate a facility with precise numbers. But just as noteworthy is the fact that the Pirahã *didn't* succeed on the tasks despite the possibility of using fairly simple non-

16. The brief presentation task could in principle be solved in an analogous manner, but this would require subjects to form a highly accurate mental image very quickly and to be capable of accurate and detailed inspection of the image.

17. The nuts-in-can task is perhaps the most difficult to fully succeed on using non-numerical strategies. While there certainly are non-numerical confounds in this case (e.g., volume of nuts), and the task could be solved perfectly for small numbers without numerical representations (again, using the object-indexing system), it is unlikely that there is a non-numerical strategy that would guarantee complete success on the task.

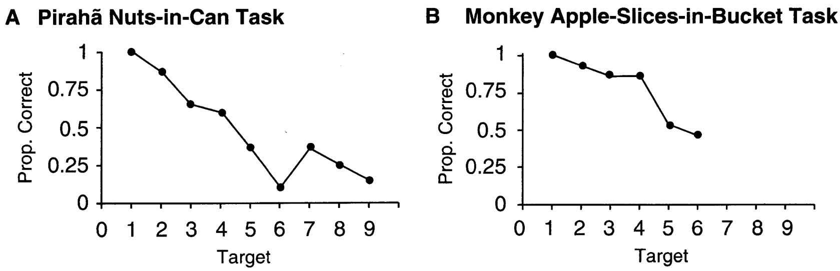

FIGURE 8.6 (a) shows how the Piraha performed on the nuts-in-can task (based on Gordon, 2004a, figure 1G). They had difficulties even with smaller numbers. (b) shows the results for rhesus monkeys on the related task of choosing between two buckets with differing numbers of apple slices (based on data in Hauser, Carey, and Hauser, 2000, figure 1, for choices between quantities that differed by 1).

numerical strategies. On five of Gordon's eight tasks (with no published data for a sixth), the Pirahã did so badly that they weren't able to succeed even for the numbers 1, 2, and 3. For example, in the candy-in-box task, the subjects were unable to reliably distinguish containers with one versus two fish on them (the case that began and illustrated the task). Likewise, the Pirahã subjects had less than perfect accuracy for two nuts in the nuts-in-can task (again, the case that began and illustrated the task), and they were only about 70% accurate for three nuts.

It's important to recognize just how substantial such failures are. Infants and animals, including dogs and monkeys, routinely succeed on quasi-numerical tasks that can be solved using non-numerical strategies (e.g., Wynn's addition and subtraction task; see section 8.1), and yet the Pirahã are failing on essentially similar tasks. Although a fully parallel study is not available, a study with rhesus monkeys by Hauser, Carey, and Hauser (2000) offers an instructive comparison. The monkeys watched as differing numbers of apple slices were placed into two opaque buckets. The question was which bucket they would approach first. Figure 8.6b shows the results. The monkeys chose the bucket with the larger number of slices more than 75% of the time for up to four slices, doing better than the Pirahã subjects did on the related nuts-in-can task. The Pirahã subjects fell significantly below 75% for as few as three nuts!

It's unclear what to make of the Pirahã's difficulties in cases where animals do better. Since in many of these cases the Pirahã ought to be able to succeed regardless of whether they can represent precise numerical quantity, we have to ask whether Gordon's tasks have features that may have inadvertently prohibited the Pirahã from revealing their true abilities.[18] We see a number of reasons why this

18. Another possibility, though one we think is unlikely, is that the Pirahã's deficits are far more profound than Gordon claims. Perhaps they not only lack the ability to represent precise numerical quantities but also lack the ability to form simple one-to-one correspondences, and even lack some of the basic capacities that rhesus monkeys exercise when evaluating the relative sizes of different sets.

might be so, but the most serious is that the tasks weren't accompanied by measures to ensure that the Pirahã understood what was expected of them.[19] In other words, it is possible that they simply didn't understand what they were supposed to do. We turn to this objection next.

3.6 *Gordon's Subjects Didn't Understand the Tasks*

We noted earlier that the matching tasks began with Gordon placing down a number of items and telling his subject to "make it the same." Then, regardless of how they responded, they were given the same encouraging feedback ("Aiyo!"). This combination of vague instructions and a guaranteed positive response is a dangerous mix. It has the troubling consequence that however the Pirahã initially interpreted the task, that interpretation was reinforced. As a result, the experimental conditions wouldn't have conveyed the intended goal except to people who guessed correctly from the start. And given that precise numerical quantities aren't supposed to be salient for the Pirahã, they would be extremely unlikely to start off with the right hypothesis. Further, as we just saw, the Pirahã failed abysmally on tasks that they could have completed with perfect accuracy using relatively simple non-numerical strategies. This would certainly make sense on the assumption that Gordon's subjects just didn't understand what they were supposed to do.

As it happens, we don't have to speculate about what Gordon's subjects were thinking. Daniel Everett has confirmed that they were unclear about what Gordon wanted from them and that they were self-conscious about their predicament (Everett 2005, p. 644):

> ...on the videotape he [Gordon] 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.

Everett's observation is troubling. Gordon's apparent indifference to whether the Pirahã were even trying to do what was expected of them in itself raises concerns about his experiments. Perhaps he assumed that the goal of each task was sufficiently obvious once an example or two was given, or that the most interesting response to measure is the one that involves the least coaxing—a spontaneous response. But we've seen that neither assumption is warranted. In any case, the simple fact that the Pirahã didn't understand Gordon's tasks shows that we shouldn't take Gordon's data

19. Interestingly, in spite of Pica et al.'s care to avoid non-numerical confounds, their precise numerical task (see note 11 above) could also be solved with 100% accuracy for small numbers, using a non-numerical strategy employing the object-indexing system, and yet the subjects performed at similar levels to the Pirahã. It is unclear why the Mundurukú ignored the more effective strategy for this task, but we suspect that it was a combination of lack of training and feedback and irrelevant performance variables, such as the speed of presentation of the stimuli, which may have encouraged numerical approximation.

at face value. The fact is that if the Pirahã did not understand the tasks, then they would be likely to fail them whether or not they are capable of precise numerical thought. So Gordon's data—his negative results—can't tell us anything substantive about the Pirahã's numerical abilities. And, of course, if they can't tell us whether the Pirahã are capable of representing precise numerical quantities, they can't tell us whether the Pirahã are capable of representing precise numerical quantities despite the limitations of their language. The result is that Gordon's study offers no support for either the Cultural Construct Thesis or linguistic determinism.

3.7 *Can the Pirahã Be Taught to Count?*

While Gordon's study focused on spontaneous judgments, it's worth noting that, at one point, the Pirahã were given explicit instruction on the Portuguese counting system. This program of education was administered by the Everetts and is briefly summarized in Everett (2005). Though the details are scarce, the information that is available is interesting for the further light that it sheds on Gordon's work.

Everett (2005, p. 626) reports that the attempts at instruction ended in failure:

> After eight months of daily efforts, without ever needing to call them to come for class (all meetings were started by them with much enthusiasm), the people concluded that they could not learn this material, and classes were abandoned. Not one learned to count to ten, and not one learned to add 3 + 1 or even 1 + 1 (if regularly responding "2" to the latter is evidence of learning).

This surprising result might initially seem to favor Gordon by offering additional evidence of the Pirahã's difficulties with precise number. But on the contrary, the Pirahã's difficulty doesn't sit at all comfortably with Gordon's linguistic determinism. If the problem for the Pirahã is that they have a *linguistic deficit*, as the thesis of linguistic determinism asserts, then teaching them number words in conjunction with the cultural practice of counting ought to give them just what they need to acquire concepts of natural numbers. An advocate of linguistic determinism should predict the Pirahã would overcome their alleged difficulties with precise numbers as they are exposed to the Portuguese counting system. (Advocates of the Cultural Construct Thesis should predict much the same thing, though they would be less focused on the linguistic character of the counting system.) Thus the Pirahã's reported failure to learn to count hardly supports Gordon's position. If anything, it argues against Gordon.[20]

20. Everett claims that Pirahã *children* easily learn to count in Portuguese as long as adjustments are made to how the words are pronounced and the instruction occurs in the context of an everyday task, such as stringing beads (personal communication reported in R. Gelman and Butterworth 2005, p. 9). This fact is consistent with linguistic determinism and the Cultural Construct Thesis, but it is also consistent with strong nativism. For example, Pirahã children might be capitalizing on an innate ability to represent precise quantity, coupled with the fact that they haven't yet absorbed their parents' strong aversion to foreign knowledge.

How, then, should we understand the Pirahã's failure to learn to count? Given the few published details about the instruction they received, we can only speculate. One possibility is that the Pirahã weren't motivated students. This conflicts with Everett's claim that they themselves had requested the classes in order to avoid being cheated in their trading relations (Everett 2005, p. 625). All the same, it is not unreasonable to suppose they were unreceptive to learning elements of the Portuguese language and culture. As we noted earlier, the Pirahã actively resist the knowledge and practices of outsiders. According to Everett:

> [T]he Pirahã ultimately not only do not value Portuguese (or American) knowledge but oppose its coming into their lives. They ask questions about outside cultures largely for the entertainment value of the answers. If one tries to suggest (as we originally did, in a math class, for example) that there is a preferred response to a specific question, they will likely change the subject and/or show irritation. (2005, p. 626)

The Everetts also put on a series of evening literacy classes for the Pirahã, again at the Pirahã's request. The results are telling (Everett 2005, p. 626):

> After many classes, the Pirahã (most of the village we were living in, about 30 people) read together, out loud, the word *bigí* "ground/sky". They immediately all laughed. I asked what was so funny. They answered that what they had just said sounded like their word for "sky". I said that indeed it did, because it was their word. They reacted by saying that if that is what we were trying to teach them, they wanted us to stop: "We don't write our language." The decision was based on a rejection of foreign knowledge; their motivation for attending the literacy classes turned out to be, according to them, that it was fun to be together and I made popcorn.

Given the Pirahã's contempt for foreign knowledge, one can imagine that the "math classes" were similarly valued simply as an excuse for getting together and that they weren't actually interested in engaging with the instruction.

3.8 *Summary*

On the face of it, Gordon's work in the Amazon seems to offer an ideal case study. What better way could there be of testing the dependence of number on language than looking at a population whose language has no terms for precise numerical quantities? Moreover, if Gordon is right that precise numerical concepts *are* essentially dependent on language, then his results would provide powerful support for the Cultural Construct Thesis, thereby undermining strong nativism. Given the suspicion with which strong nativism has come to be viewed in recent years, this outcome would be welcomed by a wide variety of theorists.

We've argued, however, that Gordon's study doesn't support either the thesis of linguistic determinism or the Cultural Construct Thesis. Since this has been a rather long section, it might help to offer a brief recap. We began with two general criticisms that were directed primarily to Gordon's linguistic determinism. Gordon claims to have established the dependence of precise number on language just by showing that the Pirahã do poorly on his eight tasks. Putting aside the issue of whether his tasks amount to good tests for numerical abilities, we noted that Gordon's argument doesn't establish a direction of dependence and that the correlation that his argument turns on is extremely weak—it amounts to a single case. Next we raised the question of whether Gordon's data do in fact show that the Pirahã are unable to represent precise numerical quantities. We argued that they do not. Part of the problem is that Gordon's experimental procedures focus on spontaneous judgments, yet incorrect spontaneous judgments can't tell us what subjects are or aren't capable of. This is all the more true when there exist plausible alternative explanations for their failure on such tasks, as there are in this case. Further, Gordon's vague task instructions and automatic positive feedback only serve to obscure the nature of the tasks for people, like the Pirahã, who don't habitually think in terms of precise numerical quantities. A rather different problem is that even if the Pirahã had succeeded on Gordon's tasks, this in itself wouldn't tell us about their numerical abilities either, since Gordon's materials didn't control for a variety of different non-numerical confounds. Given that we know that animals can take advantage of such confounds to solve related tasks, it bears explaining why the Pirahã didn't do better. It would appear that the reason for their poor performance is that Gordon's Pirahã subjects simply didn't understand what Gordon expected of them. In sum, despite the hope that an investigation of the Pirahã might settle the fundamental issues about the role of language and culture in mathematical cognition, Gordon's results leave things pretty much where they were. For strong nativists like ourselves, this means that a commitment to innate numerical concepts—something on the order of the number module outlined earlier—continues to be a genuine possibility. The debate between weak nativists and strong nativists remains unsettled.

4 Future Research

Why have these serious problems with Gordon's study been overlooked? We believe that there is something like an intellectual blind spot when it comes to evaluating exotic anthropological data. Paul Bloom reports a similar phenomenon with brain imaging data (Bloom 2006):

> In a recent study, Deena Skolnick, a graduate student at Yale, asked her subjects to judge different explanations of a psychological phenomenon. Some of these explanations were crafted to be awful. And people were good at noticing that they were awful—unless Skolnick inserted a few sentences of neuroscience. These were entirely irrelevant, basically stating that the phenomenon occurred in a certain part of the brain. But they did the trick: For both the novices and the experts (cognitive neuroscientists in the Yale psychology department), the presence of a bit of apparently-hard science turned bad explanations into satisfactory ones.

The scientific community isn't as objective as we'd all like to think. This means that we need to be more cautious when evaluating claims that play into our own weaknesses. We'd suggest that extra caution is often needed when considering claims about cognitive differences in exotic communities, just as it is needed when considering the implications of neuroimaging data. This isn't to say that we should abandon cross-cultural research on numerical cognition. On the contrary, we believe that cross-cultural data can provide an important source of evidence for understanding the nature of human mathematical abilities. And we'd very much like to see more systematic research along these lines. However, it is important that we guard against dropping our standards of evidence when we see phrases like "Amazonian tribe." For this reason, we'll end the chapter by assembling a set of minimal guidelines for future research, guidelines that build on the critical discussion in section 8.3.

First, future experiments need to make precise number more salient for the subjects. Given that the populations of particular interest are ones in which there is no communal practice of counting and that appear to have little regard for precise number, it is not enough to merely present tasks intended to test precise numerical abilities, or even to convey to the subjects that the tasks are broadly numerical. We know from various recent work that approximate number can be represented non-linguistically. So when a reliance on approximate number is the default response in a community, we need to find some way of conveying the goal of being more precise. Perhaps one way of pushing things in this direction would be to use stimuli that evoke situations where careful numerical comparisons would be more natural to make, for example, by asking which of two mothers has more children. This might be done using real families known to the local people or using photographs of people not personally known to the subjects (photos would have the advantage of allowing experimenters to control for various possible non-numerical confounds).

Second, measures need to be put in place for determining whether subjects understand the goal of a task. To foster better understanding, meaningful feedback could be given on a series of trials that precede testing. So long as the test trials use novel stimuli, we can exclude the possibility that good performance is achieved simply by memorizing the answers given in the pretest trials.

Third, measures need to be put in place to ensure that subjects are well motivated to succeed on numerical tasks. One possibility, again taking the lead from the literature on animals, would be to introduce a competitive paradigm. Tasks where subjects competed for a reward might substantially increase motivation. A related possibility is to employ a noncompetitive paradigm that offers rewards of differing value, where the greater reward is contingent upon a precise numerical discrimination. In effect, this is what Hauser et al. (2000) did in the study with rhesus monkeys that we described in section 83.5 (the one where different numbers of apple slices were placed into two buckets). This general approach would be easy to adapt for use with human adults.

Fourth, future experiments ought to be constructed so as to avoid excessive performance demands. The tasks shouldn't incorporate time constraints that encourage estimation and shouldn't be taxing for reasons that have nothing to do with the numerical judgments being elicited. It's perfectly fine for the tasks to be simple ones. In fact, the simpler the better. Any extra complications only make it difficult to say whether poor task performance is owing to poor numerical abilities or to inessential features of a task.

Fifth, future experiments should control for non-numerical confounds. There are many ways of doing this, but anthropologists might consider adapting experiments that have already been done with infants or animals. These tend to employ rigorous controls for non-numerical confounds but also have the advantage that they can be implemented without verbal instructions. See, for example, the study by Brannon and Terrace (1998) described in section 8.1.1 above. Although Brannon and Terrace's study was intended to test ordinal knowledge, ceiling performance on the task requires detecting precise numerical differences (e.g., the difference between 5 and 6). Ideally, to tease apart different hypotheses with respect to both linguistic determinism and the Cultural Construct Thesis, such a task might be run under three different conditions—one with prior training on number words, one with prior training on a nonlinguistic counting technique, and one with no prior numerical training. In principle, this would allow experimenters to determine the relative contributions of language and counting to the representation of precise numbers.

Sixth, to address the issue of linguistic determinism, test subjects should be given ample opportunity to *learn* precise numerical concepts. The issue of linguistic determinism is what people are capable of representing in the absence of the relevant linguistic features, not what they happen to represent in their ordinary experience. So while instruction shouldn't employ a linguistic counting procedure or the inculcation of number words, this still leaves room for various forms of instruction that do not turn on such linguistic devices, including the use of body parts or other external symbols.

Seventh, if at all possible, it would be beneficial to train and test children, not just adults, since this would help to clarify the source of difficulty in those cases where adults are resistant to instruction. There is anecdotal evidence that Pirahã children are able to learn precise quantities even if their parents can't (see note 20). But only a systematic evaluation can tell us if such claims hold up. One hypothesis as to why Pirahã children might do better than Pirahã adults is that the children haven't fully embraced their parents' strong aversion to knowledge based on other cultures. This could be tested, at least provisionally, by questioning children about their views of their Portuguese-speaking Brazilian neighbors.

Implementing these guidelines and suggestions would enable future cross-cultural research to make significant contributions to our understanding of numerical cognition and could help to resolve some of the large-scale issues regarding the nature and development of numerical concepts. We realize that fieldwork involves all sorts of practical limitations and that desirable experiments may not always be feasible. Nonetheless, if cross-cultural data are going to be more meaningful than the suggestive anecdotes that are already in the literature, we have to try to maintain the highest possible experimental standards.