ORIGINAL PAPER

The benefit of symbols: monkeys show linear, human-like, accuracy when using symbols to represent scalar value

Margaret S. Livingstone · Krishna Srihasam · Istvan A. Morocz

Abstract When humans and animals estimate numbers of items, their error rate is proportional to the number. To date, however, only humans show the capacity to represent large numbers symbolically, which endows them with increased precision, especially for large numbers, and with tools for manipulating numbers. This ability depends critically on our capacity to acquire and represent explicit symbols. Here we show that when rhesus monkeys are trained to use an explicit symbol system, they too show more precise, and linear, scaling than they do using a one-to-one corresponding numerosity representation. We also found that when taught two different types of representations for reward amount, the monkeys systematically undervalued the less precise representation. The results indicate that monkeys, like humans, can learn alternative mechanisms for representing a single value scale and that performance variability and relative value depend on the distinguishability of each representation.

Keywords Monkeys · Numbers · Symbol learning

Introduction

Many studies indicate that humans are innately endowed with an approximate non-symbolic representation of number and that an analogous approximate numerosity system

M. S. Livingstone (✉) · K. Srihasam
Department of Neurobiology, Harvard Medical School,
220 Longwood Ave, Boston, MA 02115, USA
e-mail: mlivingstone@hms.harvard.edu

I. A. Morocz
Department of Radiology, Brigham & Women's Hospital,
Harvard Medical School, 75 Francis St, Boston, MA 02115, USA

is present in other animals, which is adaptive for comparing and evaluating food sources, threats, and social situations (Dehaene 1997). Untrained human infants and macaques behave as if they naturally assess and keep track of the number of fruits or conspecifics (though only for quite small numbers; Hauser et al. 2000). The psychophysics of approximate numerosity judgments are remarkably similar in humans and animals, in that both show magnitude-dependent variations in accuracy (Dehaene 1997); i.e. it is more difficult to distinguish 8 objects from 9 objects than it is to distinguish 1 object from 2 objects, even though both pairs differ by 1. Thus the standard deviation of approximate numerical estimates increases in proportion to the magnitude being estimated (Krueger 1982), a property that Gibbon called scalar variability (1977). In this way, judgment of approximate numerosity in both humans and animals is similar to more fundamental sensory processes in that the threshold for discrimination between two stimuli scales with their magnitude (Weber 1843).

The increasing discrimination threshold with numerosity magnitude has been interpreted in two ways: (1) that numerosity is internally represented by a linear 'accumulator' whose variability is proportional to the accumulated amount (Gibbon 1977; Whalen et al. 1999). (2) That the internal representation of numerosity is logarithmically compressed, with constant variability, so that larger numbers are subjectively closer together on a mental 'number line' (Dehaene and Mehler 1992; Dehaene 1997). Although it is debated which of these models best explains the way numerosities are internally represented (Cantlon et al. 2009; Dehaene et al. 2009), it is clear that the approximate number sense shows scalar variability, a feature common to many, seemingly more fundamental, sensory perceptions.

Most human cultures have, in addition, developed symbol systems for representing numerosity exactly—i.e. much

more precisely than is possible using our evolutionarily older approximate number sense. Symbol systems allow us to discriminate smaller proportionate differences than we could for the same numbers of items; e.g. 99 vs 100, and to manipulate quantities using memorized abstract rules. Although chimps, monkeys and parrots can learn to associate Arabic numerals (or number words) with numbers of items (Ferster 1964; Matsuzawa 1985; Washburn and Rumbaugh 1991; Pepperberg and Gordon 2005), few have so far exhibited any benefit to using such symbolic, rather than one-to-one correspondence, representations. One benefit to an animal of using symbolic representation was shown by Boysen and Hallberg (2000), who found that chimpanzees presented with two sets of objects, in a task where they were given the unchosen number of reward items, were unable to suppress choosing the larger number of items, but when the two reward amounts were represented by numerals, they were able to choose the smaller numeral and receive the larger reward. Furthermore, Matsuno et al. (2006) showed that chimpanzees who had learned color symbols could make more precise color discriminations than chimpanzees who had not. Here we compare the behavior of highly trained macaque monkeys in discriminating numerals versus dots arrays representing the same system of value; we find that monkeys using symbols, compared to using the corresponding numbers of items, show substantially increased choice accuracy.

Materials and methods

Four juvenile male rhesus macaque monkeys participated in these experiments beginning training at 1 year of age. In a compartment of each home cage, a touchscreen monitor (Elo TouchSystems, Menlo Park, CA) was mounted at eye level with a stainless steel juice delivery tube just below. Software for stimulus presentation, reward delivery, data collection, and data analysis were written in MatLab (Mathworks, Natick, MA). Pairs of stimuli were presented on the left and right side of the monitor. Each stimulus numerosity was randomly chosen from the values 0–21, with the constraint that the two values could not be the same. The numerals were 5 cm in height, and the circles containing the dots were 9 cm in diameter. For the dot stimuli, visual features such as dot size, dot color and dot position (within a circular boundary) varied randomly; the dots were drawn so that when two dots overlapped they differed in color, and the smaller dot was drawn in front; thus all the dots were easily seen and discriminable by a human observer, even for the highest numbers of dots. For the shapes control task the dots were replaced by differently colored shapes; for the equal area task the total area covered by the dots was constrained to be constant. The elements were not allowed to

overlap in either the shapes or the equal area conditions. For the numerals, the letters X Y W C H U T F K L N R were used to represent numbers 10–21 because we chose not to use double digits; we will therefore refer to them as 'numerals'.

The animal chose one of the two stimuli by touching the screen within 2.25 cm of the center of the stimulus. He was always rewarded by the same number of drops as the numerosity of the chosen stimulus; the drops were delivered at 4 Hz, each drop accompanied by a beep sound. Each stimulus pair was presented for 10 s or until the animal responded by touching either stimulus. A new stimulus was presented 3 s after the end of the preceding trial. Each monkey was allowed to work, alone, for at least 3 h per day, and they usually stopped working after 2 h, or 300–400 trials. For the first 3 months of training, they were given ad lib water during non-test periods, and juice rewards. After they stopped taking the free water in their cages and were drinking predominantly the juice provided during test periods, we switched the reward to water and stopped providing fluids before the daily testing period. We offered water in the afternoon, after testing, but the monkeys usually did not take any; their daily fluid intake was always more than 30 ml/kg. They always had ad lib food and have been steadily gaining weight.

The monkeys were trained alternately on choosing between pairs of numerals and pairs of dot patterns. They started with 0 and 1 and progressed to the next higher number when their performance on the highest number reached >75% averaged over all possible combinations with other numbers, for both dots and numerals. The reward pulses were long (yielding several drops per pulse) when the maximum numbers were small, and decreased in length as the monkeys attained larger maximum numbers. They reached a stable level of performance for both dots and numerals from 0 to 21 after 4 months of training. Then they were tested on 0–21 for both dots and numerals, on alternate days, for 1 month to ensure that their performance was stable, and then tested for 2 months to obtain the data presented here (Fig. 1). Control experiments (Fig. 3) were run subsequently. Only data from the first 200 trials each day, excluding the first 10 trials, were used for analysis because after 200 trials the monkeys usually started playing around and working sporadically, and during the first few trials the experimenter was still in the room, which distracted them.

Because the monkeys were allowed up to 10 s to make a choice before the next stimulus was presented, their reaction times could be as long as 10 s, but histograms of their reaction times for the first 200 trials each day peaked at less than 600 ms, with a long tail extending homogeneously from 1 to 10 s. This long tail means that average reaction times, as well as the harmonic means and medians, were biased toward longer durations than the peak of the reaction
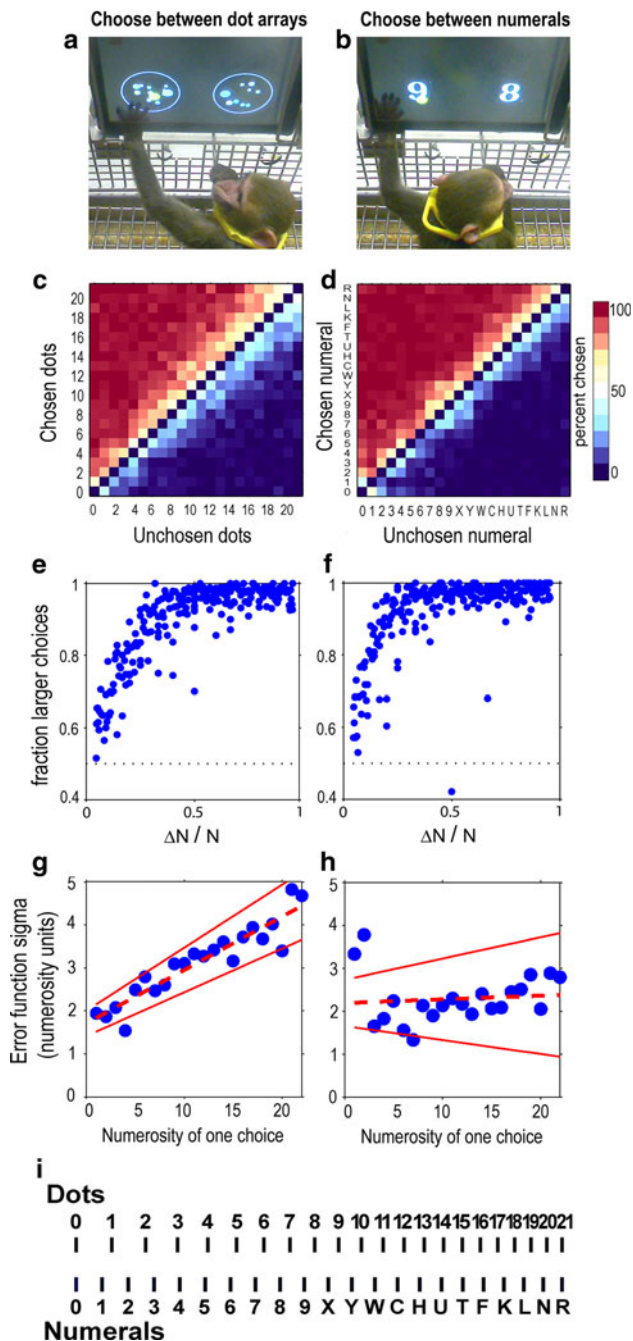
Fig. 1 Young monkeys distinguish surprisingly large numerosities. **a** Monkey using a touchscreen to choose 10 dots in preference to 9 dots (the *large yellow dot* appears where he touches the screen). The dots vary in color, size, and position randomly (the colors are much more distinct from each other than they appear on this video image). The monkey's mouth is on the juice tube. **b** Monkey choosing the numeral 9 in preference to 8. **c** Average performance choosing between two dot patterns for 4 monkeys on alternate days over a 2-month period (∼6,000 trials per monkey). The color of each square in the matrix indicates the percentage of trials in which the monkey chose the number corresponding to the square's *vertical position* over the number corresponding to its *horizontal position*. Thus all the *squares* in the *upper left* represent trials when the monkey chose the larger numerosity over the smaller, and the squares in the *lower right* show trials when the monkey chose the smaller numerosity. Note the increasing *number of 'errors'* for proportionally smaller differences between choices. **d** Average performance for choosing between two numerals for 4 monkeys on alternate days over a 2-month period (∼6,000 trials per monkey). The same data replotted, at a reviewer's request, graphically; the fraction of times the larger numerosity (**e**) or numeral (**f**) was chosen as a function of the difference between the two choices, divided by the larger choice. The standard deviation of the error function for choices between dot patterns (**g**) and numerals (**h**). The *dotted red lines* show the best linear fit to the data, with the 95% confidence intervals indicated by thin *red lines*. **i** Scales for dots (*top*) and numerals (*bottom*) compressed so that the standard deviations in **g**, **h** are constant on that scale; the scales were calculated from the best fitting lines in **g** and **h**

numerosity, we trained juvenile macaque monkeys to associate Arabic numerals and dot arrays with reward amount. We found that monkeys could easily learn dot-array numerosity and abstract symbolic representations of surprisingly high numerosities when we directly associated the symbols or the numerosities with reward amounts (Washburn and Rumbaugh 1991; Beran et al. 2008) using an automatic touch-screen apparatus. For several hours each day, in a section of their home cages, the monkeys were given choices between two simultaneously presented number stimuli; they chose one by touching it and were given the same discrete number of drops of juice as the chosen option. That is, there was no 'wrong' or 'right' choice, but the reward was proportional to the chosen number, so choosing the larger option was presumably better than choosing the smaller. The two choices presented were initially two Arabic numerals or two sets of dots randomly positioned inside a circle (Fig. 1a, b). The monkeys soon began to choose the larger of the two dot-array numerosities or the larger numeral (even though they were rewarded for choosing either side), so far for pairs from 0 to 21. Most interestingly, the monkeys, like humans, were more accurate when using numerals to represent the number of drops of juice reward than when using the seemingly more direct and intuitive one-to-one representation of dot-array numerosity. The monkeys' accuracy is impressive: for most choice pairs the monkeys chose the larger number *almost every time*, except for consecutive pairs of large numbers, like 21 vs. 20 (Fig. 1c, d).

time histogram. We therefore used only reaction times shorter than 800 ms to calculate the plots in Fig. 2b–d.

All experiments were done in accordance with procedures approved by the Harvard Medical School Standing Committee on Animals.

Results

To study the association between the approximate analog number system and the precise symbolic representation of
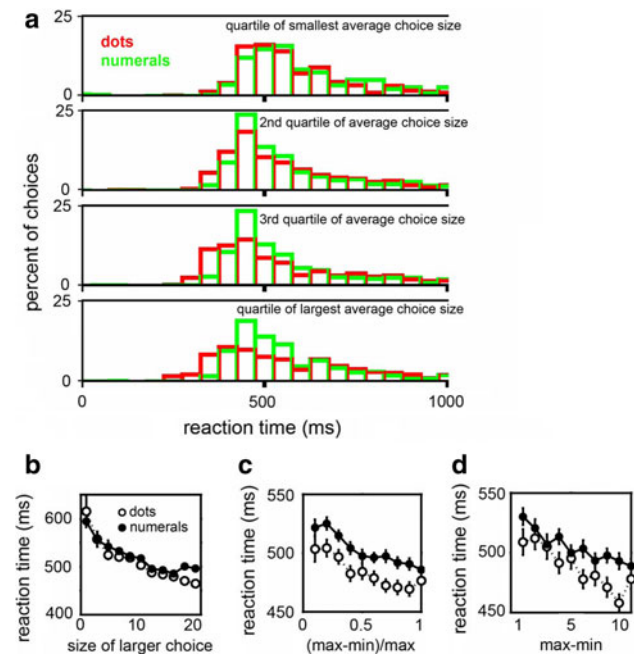
Fig. 2 **a** Histograms of reaction times less than 1 s for 4 monkeys choosing between pairs of dots (*red*) or pairs of numerals (*green*) for the quartile of choices in which the sum of the two choices was less than 10 (*top*); between 10 and 21 (*second panel*); between 21 and 32 (*next to* *last panel*); or smaller than 42 (*bottom*). **b** Average reaction times (±sem) as a function of the size of the larger of the two choices for dots (*open circles*) and numerals (*solid circles*). **c** Average reaction times (±sem) as a function of the numerical difference between the two choices divided by the larger choice, for dots (*open circles*) and numerals (*solid circles*). **d** Average reaction times as a function of the absolute difference between the two choices for choice pairs in which the maximum choice was at least 11 and the difference between choices was between 1 and 11. For panels **b–d**, averages were calculated using only reaction times less than 800 ms

When choosing between dot patterns, the monkeys showed more mistakes (choosing the smaller numerosity) when the difference between the two choices was small compared to their size (Fig. 1c); that is they showed scalar variability. When the monkeys chose between symbols (Fig. 1d), however, although they still tended to make more errors for small numerical differences, their inaccuracy did not scale with choice magnitude. To quantify this difference between choice behavior for numerals versus dots, we measured the fraction chosen for each numerosity, as a function of the difference between that numerosity and every other choice. We then fit a normal cumulative distribution function (error function) to each curve; the width of this distribution for each numerosity, the sigma of each function, is plotted in Fig. 1g for each dot numerosity and in Fig. 1h for numerals. These plots confirm that for dots the performance curves are distinctly broader for increasing numerosities, but this broadening with magnitude is almost absent for numerals.

Supposing that the relationship between response variability and numerical magnitude reflects a compression of

the underlying number scale (hypothesis 2 in the introduction), we distorted the two number scales so that the standard deviations would be constant on that scale. This distortion results in a scale that is relatively more compressed at higher numerosities for dots than is the comparable scale for numerals (Fig. 1i). However, these data do not directly address the question of whether the difference in behavior between numerals and dot arrays depends on differences in underlying scale.

The monkeys responded faster when at least one of the choices was large, compared to when both choices were small, probably reflecting higher motivation for choosing large numerosities (Fig. 2a, b). The fact that the monkeys' reaction times decreased, instead of increasing, with increasing number of dots indicates that they were not counting the dots. This is consistent with a study on a chimpanzee who learned to match numbers of dots to Arabic numerals, who also did not show increasing reaction times for increasing numbers of dots (Murofushi 1987). Reaction times also varied with choice difficulty, measured as the proportional difference between the two choices; i.e. with the difference between the choices divided by the larger choice, consistent with previous studies on monkeys using dot arrays (Brannon and Terrace 2000; Fig. 2c). This dependence on the difference between choices was true even when choice pairs were restricted to those with differences between 1 and 10 and maximum numerosities >11 (in order to rule out the possibility that faster choices for larger differences reflected maximum choice size rather than choice difficulty; Fig. 2d). The monkeys consistently chose slightly faster for dot pairs than for numeral pairs, though the difference was quite small (<20 ms).

We asked whether, for the dot arrays, the monkeys were basing their choices on the number of dots, or whether they relied on some low-level visual attribute. Two control experiments using patterns of random shapes (Fig. 3a), or patterns of dots in which the total dot surface area was always constant (Fig. 3c), indicate that the monkeys based their dot array choices on numerosity and not on low-level visual attributes. This is consistent with more extensive studies showing that monkeys preferentially base their judgments on numerosity in situations like this (Cantlon and Brannon 2007).

Since the monkeys learned independently to choose between dots-vs-dots or numerals-vs-numerals, we consider 3 increasingly abstract scenarios for what they learned: 1) two independent matrices of pairwise choices, 2) two independent rank orders, or 3) the number of drops of juice, or the hedonic value, symbolized by each stimulus. Monkeys can learn to rank order quite long lists of numerosity-based or arbitrary visual stimuli, and they can show transitive inference from one list to another (Treichler and Van Tilburg 1999; Terrace et al. 2003; Treichler et al. 2003). Nevertheless,
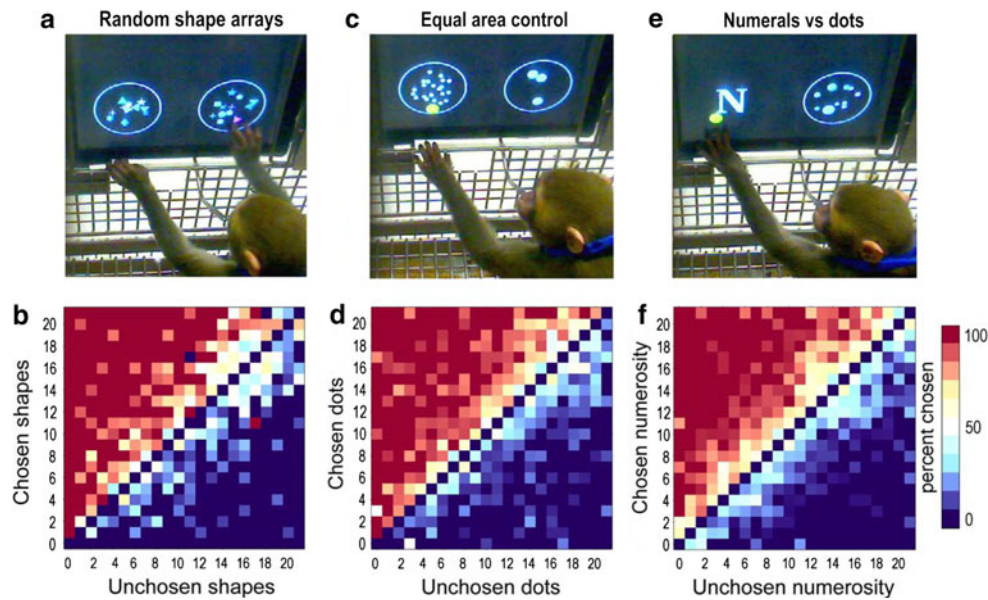
Fig. 3 Generalization and abstraction behavior of monkeys trained to associate dot arrays and numerals independently with the corresponding number of drops of juice. **a** Monkey choosing 12 shapes over 10. The stimuli are shapes of different colors (the colors were more distinctive than they appear on this video image). **b** Average performance of 4 monkeys for the first 2 days on this task. **c** Equal-area control, total dot area was always constant. The monkey chooses 21 small dots instead of 3 large dots. **d** Average performance of 4 monkeys for the first 2 days of the equal-area control task. **e** Transitivity of dot array learning and numeral learning. The monkey chooses the numeral *N* (= 20 drops) over 10 dots (= 10 drops). **f** Average performance of 4 monkeys over the first 4 days on the numeral vs. dots choice task. Their average performance did not change over the 4 days

we can rule out the possibility that in our task they learned two independent matrices of pairwise choices or two independent rank orders because when they were first presented with choices between one numeral and one dot array (Fig. 3e), they immediately started to choose, accurately, the larger one, irrespective of modality. This means they recognize dot numerosity and numerals as representing equivalent scalar values, for surprisingly large and precise amounts.

Intuitively it seems obvious that it would be more difficult to distinguish 20 dots from 21 than to distinguish the corresponding numerals, but is this because numerals and numerosities are scaled differently in our brains, as has been proposed for approximate versus precise numerosity representations in humans (Dehaene et al. 2008). We cannot tell whether the monkeys treat the numerals as representing numerosity, because we do not know how they quantify the reward: they could use number of drops, amount of juice, reward duration, or a more abstract hedonic scale. But, as discussed earlier, our controls (Fig. 3a–d) and the results of previous studies (Cantlon and Brannon 2007) indicate that monkeys do use the *numerosity* of dot arrays to make dot choices; therefore either they treat the numerals as equivalent to numerosities, or the numerals and the dot numerosities are each independently equivalent to some scalar measure of juice or hedonic value. In either case this hedonic scale, whatever it represents, is mutually equivalent to both numerosity and the learned numeral symbols,

so we now consider the mapping of the two different representations, dots vs numerals, to this scale.

We first calculated the 'perceptual' equivalence for dots and numerals, following the approach of Weber (1843). To do this, we assumed that Weber's perceptual just-noticeable-difference is proportional to the sigma of the error function. We then plotted the distance along each number scale corresponding to the cumulative number of standard deviations, for dot arrays and numerals separately (Fig. 4a). Because discriminability was dependent on magnitude for dots but not for numerals, the difference between two choices corresponding to a given perceptual increment depends on magnitude for dots, but not for numerals; i.e. the perceptual value of numerals increases almost linearly with increasing magnitude, but the perceptual value for dot arrays falls off with magnitude (Fig. 4a). This was Fechner's insight: if perceptual value is correlated with discriminability, and discriminability of a fixed stimulus difference decreases with stimulus magnitude (Weber 1843), then the perceptual response must fall off with stimulus magnitude (Fechner 1860). That the perceptual value of dot arrays, but not numerals, should fall off with stimulus magnitude is counterintuitive in our case since the hedonic value of any numerosity should be the same irrespective of whether it is represented by a numeral or dot array.

In Fig. 4b, we plot the relative value of each number of dots compared to each numeral, from 0 to 21, derived from the curves plotted in Fig. 4a. Counterintuitively, this
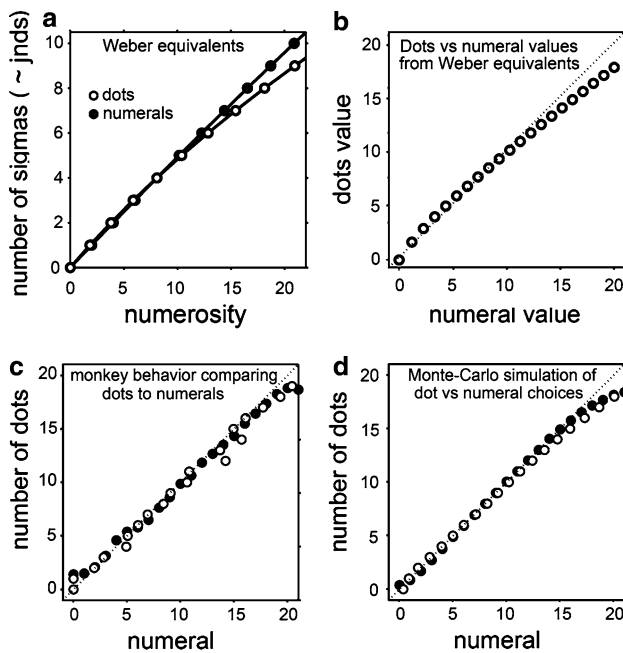
Fig. 4 Perceived value equivalence for dot numerosity and numerals. **a** Correspondence between the number of sigmas (of the error function) and stimulus numerosity for dots (*open circles*) and numerals (*filled circles*) based on the sigmas calculated for panels g and h in Fig. 1, using only dots-vs-dots and numerals-vs-numerals choice behavior. **b** Predicted equivalence between dot and numeral values based on dots-vs-dots and numerals-vs-numerals choice behavior. The ratio of dot value to numeral value was calculated from panel (**a**) for each numerosity 0–21. **c** Behavioral measure of dots vs numeral values based on numeral-vs-dots choice behavior. *Filled circles* show the numeral value that is behaviorally equivalent to each dot numerosity, and *open circles* show the number of dots equivalent to each numeral. The *dotted line* is dots = numerals. **d** Modeled value equivalents for numeral-to-dots matching using a Monte Carlo simulation (1,000,000 trials) with the same probability distribution as the behavioral distribution for dot-to-dot or numeral-to-numeral matching in Fig. 1 g and h; *filled circles* = numeral equivalent to each dot number; *open circles* = dot equivalent to each numeral. Note that both the monkeys' behavior and the predicted behavior show an underestimation of the value of dots compared to numerals at the highest numerosities

calculation predicts that twenty-one dots should be less 'valuable' to a monkey than the 'numeral' R, even though both represent 21 drops of juice. Therefore, given a choice between R and 21 dots, the monkeys should slightly prefer R. We tested this prediction using the task illustrated in Fig. 3e in which the monkeys chose between a dot array and a numeral: the same four monkeys were tested for 4 weeks on the dots-vs-numerals task. We calculated for each dot numerosity the proportion of times each possible numeral choice was made, and vice versa. Then for each dot numerosity and for each numeral, we fit an error function to the data and calculated the 50% choice probability, which gives us a numeral equivalent for each dot numerosity, and vice versa (Fig. 4c). The monkeys' equal choice behavior fell along the dots = digits identity line (dotted

line) for small numbers, but at higher numerosities the monkeys tended to undervalue the dots compared to the numerals, as predicted by our calculation (Fig. 4b) based on their dots-vs-dots and numeral-vs-numeral choices.

We then asked whether a simple model of numerosity discrimination based only on scalar variability could explain the difference between dots and numeral valuations. We did a Monte Carlo simulation of the numeral-vs-dots choice based on the monkeys' dots-vs-dots or numeral-vs-numeral choice accuracy. Each numeral or dot number was assigned a 'value' randomly drawn from a Gaussian choice probability with a sigma based on the data in Fig. 1g and h. The model then chose whichever stimulus had the higher 'value'. Not surprisingly the model accurately reproduced the variability in the dots-vs-dots and numeral-vs-numeral choice behavior in Fig. 1c and d (not shown), but the model also reproduced the undervaluation of dots at the higher end of the choice range (Fig. 4d). Therefore, the degree of scalar variability for the two visual representations alone can explain the fall off in estimated value for large numbers of dots and the compressed dot valuation scale.

Discussion

The monkeys in this study learned to use both arbitrary symbols and the number of dots in a display to represent reward amount. They showed compressed valuation when estimating numbers of dots but essentially linear valuation of numerals. The former behavior is common to humans and other animals (Dehaene 1997), but linear symbolic number comprehension is thought to be unique to educated humans (Dehaene et al. 2008).

Is numerosity comprehension required for any of the behaviors in this study? There are two reasons to believe that numerosity was indeed the cue on which the monkeys based their dot-array choices: (1) When initially tested with dot arrays in which the total dot area was constant or with arrays of arbitrary shapes, the monkeys immediately chose the larger numerosity, indicating that numerosity, rather than some low-level visual cue, was critical to their value judgments. (2) Cantlon and Brannon (2007) showed that even in the presence of other options, like color or shape, monkeys preferentially choose to categorize stimuli by their numerosity. Therefore, let us assume that the monkeys assigned a value to each dot pattern based on dot numerosity. Interestingly, the monkeys comprehended the numerosity of a dot array without depending on low-level visual cues, yet the reaction-time results indicate that they do so without counting the dots. The reaction-time results further suggest that the mental process of comprehending the numerosity of a dot array is slightly easier for the monkeys than is recognizing and evaluating learned symbols.

It is less clear that the monkeys used numerosity to evaluate the numerals. Just because we used numerals to represent reward amount does not mean that the monkeys treated the numerals, like the dots, as numerosity cues. Several species of non-human primates have been taught to rank order large numbers of arbitrary images, including numerals (Brannon and Terrace 1998; Terrace et al. 2003; Inoue and Matsuzawa 2007), and intuitively there is no reason to think that numerosity comprehension should play any role in this behavior. However, the monkeys in our study did, without explicit training, treat the dot arrays and the numerals as interchangeable representations for the same reward scale; i.e., the number of drops of juice.

It is not surprising that our monkeys should have regarded the two symbolic representations as interchangeable since we explicitly trained them to use both systems with the exact same reward scale. Previous studies have shown that monkeys can spontaneously interchange numerosity cues and ordinal value in rank order tasks, indicating an intuitive appreciation that analog quantities can correspond to ordinal position (Harris et al. 2007). Our study differs from the extensive studies on list learning in macaques (Treichler and Van Tilburg 1999), in which the goal is to avoid any detectable differential reward association for each item, in order to find out if monkeys can appreciate an ordinal relationship among items without any other cue than the knowledge of which item is rewarded in pairwise presentations. In our study, instead, we made it as clear as possible, using discrete juice drops and accompanying sounds, that each number was differentially, and uniquely, rewarded. That is, in list-learning studies the question is whether monkeys posses the cognitive capacity to learn rules and infer an underlying scale from conditional 2-choice discriminations, but here we explicitly provided a linear analog scale they could associate with the different visual stimuli.

Educated humans have a linear appreciation of a symbolic number scale, a non-linear appreciation of numerosity, and the ability to correlate the two representations. It is usually assumed that our linear symbolic number understanding is correlated with, and influences, our non-linear approximate number sense, as a consequence of repeated experience associating numbers of items or amounts with numerals. But our approximate number sense remains non-linear even though our understanding of numerals is linear. Our linear understanding of numerals is distinct from our approximate number sense in other ways: Most of us know instantly that 99 cents is more than 98 cents without imagining two piles of pennies and mentally comparing them–that kind of associative link would be less accurate and slower than people normally are. Furthermore, even with extensive experience using precise linear symbolic representations for numerosity, people tend to underestimate

large numbers of items, guessing that 100 dots are fewer than half that number (Krueger 1982; Izard and Dehaene 2008), even while presumably being able to differentiate precisely any pair of large numerals.

Dehaene has proposed that these differences indicate that we have two distinct number representations: a non-linear approximate number sense and a precise linear symbolic number sense that is achieved through education (Dehaene 1997; Dehaene et al. 2008). Understanding the links between these two distinguishable functions is, of course, complicated in humans by the myriad ways we learn to associate them, so our monkey behavior can provide a simplified model. Our monkeys' associations between numerals and numerosities show some of the same hallmarks of the link between humans' approximate and symbolic number systems. First, of course, is the difference in their variance. Second, even though the monkeys did not show increasing inaccuracy with increasing numeral magnitude, they did show slower reaction times for smaller numerical differences, as do humans (Moyer and Landauer 1967). Third, the monkeys systematically undervalued the larger numerosity dot arrays compared to the numerals, despite the repeated association of each with equivalent rewards. This is analogous to the observation that educated adult humans, despite decades of training associating numerosity and numbers, grossly underestimate the number of dots in a display, for numerosities between 10 and 100 (Izard and Dehaene 2008).

One explanation for both our monkeys' behavior, and much about human numerosity behavior, is that numerosity is a perception, a sensory process (Burr and Ross 2008; Ross and Burr 2008), and, like other fundamental sensory processes, it shows magnitude-dependent variability. The monkeys' numeral behavior did not show scalar variability, and therefore likely does not depend on numerosity comprehension. Instead we propose that the degree of discriminability of symbols governs the linearity and accuracy of their scalar associations. A reviewer argued that this difference could be regarded as a 'methodological quirk' in that each dot numerosity could be represented by a variety of configurations with occasional overlap of dots, whereas each numeral was always the same. When anyone, animal or human, estimates the number of objects they are less accurate for larger numerosities, and this is presumably because as you increase the number of objects, among other things, the objects do pile up in different configurations and can occlude each other. So, yes, the number of possible configurations and possible overlap doubtless contribute to the scalar variability we observed, but the scalar variability of numerosity perception has been exhaustively shown not to be a methodological 'quirk', but rather a fundamental feature of the way humans and animals estimate the numbers of objects (Dehaene et al. 2004). One could make the

presentation of numbers of dots less variable, by always using the same dot configuration for each numerosity, but then one would no longer be presenting numbers of objects, but rather symbols. Indeed it has been argued that number of possible configurations between large and small numbers of objects may underlie the phenomenon of subitizing—two objects always form a line, three objects always form a triangle, but five objects can have numerous configurations (Mandler and Shebo 1982).

Adult humans estimating large numbers of items show increasing errors for proportionately smaller differences, but our accuracy using numerals is almost entirely independent of number magnitude, i.e. it is linear. Very young children, however, do show reduced accuracy in comparing increasingly large numerals (Sekuler and Mierkiewicz 1977). These results have been interpreted as indicating that humans have an innate compressed, or logarithmic, scaling of numerosity, which later becomes associated with a more linear symbolic representation (Dehaene et al. 2008). Indeed young children place 30 near the middle of a 1–100 number line; but they shift between second and fourth grade to a linear mapping (Siegler and Opfer 2003). This suggests that either education or maturation causes a compressed to linear shift in number mapping. Because remote Amazonians untrained in the use of numerals above 5 retain compressed numerosity maps as adults (Dehaene et al. 2008), it has been concluded that education must be responsible for the compressed to linear shift in humans. It has been proposed that experience with measurement or with addition and subtraction brings about the realization that all consecutive numbers are separated by the same interval, and the realization that the number scale is therefore linear. However, because our otherwise uneducated macaque monkeys show linear discrimination when using symbolic representations, we propose that a linear map is not imposed on children by the repeated use of linear tools, like rulers, or by learning addition, but rather that linear mapping arises from the equal distinguishability of one symbol from another.

Our results also suggest an explanation for why educated adults systematically underestimate large numbers of items. The monkeys undervalued the dots relative to the numerals at larger numerosities, even though both were associated with the same reward scale. Previous work has shown that scalar variability in a sensory modality results in a perceptual scale that is logarithmic, or otherwise compressed with increasing magnitude (Fechner 1860; Stevens 1961), but the scales are either arbitrary or relative in magnitude. Here we show, using a simple Monte Carlo simulation, that given two sensory inputs and a mutual finite value calibration, the modality with the larger scalar variance will be increasingly undervalued with magnitude.

In monkeys, as in educated humans, the correlation between approximate numerosity perception and more precisely discriminable symbolic representations can dissociate in characteristic and predictable ways, depending on the perceptual discriminability of the representations.

References

Beran MJ, Harris EH, Evans TA, Klein ED, Chan B, Flemming TM, Washburn DA (2008) Ordinal judgments of symbolic stimuli by capuchin monkeys (*Cebus apella*) and rhesus monkeys (*Macaca mulatta*): the effects of differential and nondifferential reward. J Comp Psychol 122:52–61

Boysen S, Hallberg K (2000) Primate numerical competence: contributions toward understanding nonhuman cognition. Cogn Sci 24:423–443

Brannon EM, Terrace HS (1998) Ordering of the numerosities 1 to 9 by monkeys. Science 282:746–749

Brannon EM, Terrace HS (2000) Representation of the numerosities 1–9 by rhesus macaques (Macaca mulatta). J Exp Psychol Anim Behav Process 26:31–49

Burr D, Ross J (2008) A visual sense of number. Curr Biol 18:425–428

Cantlon JF, Brannon EM (2007) How much does number matter to a monkey (Macaca mulatta)? J Exp Psychol Anim Behav Process 33:32–41

Cantlon J, Cordes S, Libertus M, Brannon E (2009) Comment on "Log or Linear? Distinct intuitions of the number scale in Western and Amazonian Indigene Cultures". Science 323:38b

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

Dehaene S, Mehler J (1992) Cross-linguistic regularities in the frequency of number words. Cognition 43:1–29

Dehaene S, Molko N, Cohen L, Wilson AJ (2004) Arithmetic and the brain. Curr Opin Neurobiol 14:218–224

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

Dehaene S, Izard V, Pica P, Spelke E (2009) Response to Comment on "Log or Linear? Distinct Intuitions of the Number Scale in Western and Amazonian Indigene Cultures". Science 323:38c

Fechner GT (1860) Elemente der psychophysik. Breitkopf und Hèartel, Leipzig

Ferster CB (1964) Arithmetic behavior in chimpanzees. Sci Am 210:98–106

Gibbon J (1977) Scalar expectancy theory and Weber's law in animal timing. Psychol Rev 84:279–325

Harris EH, Beran MJ, Washburn DA (2007) Ordinal-list integration for symbolic, arbitrary, and analog stimuli by rhesus macaques (*Macaca mulatta*). J Gen Psychol 134:183–197

Hauser MD, Carey S, Hauser LB (2000) Spontaneous number representation in semi-free-ranging rhesus monkeys. Proc Biol Sci 267:829–833

Inoue S, Matsuzawa T (2007) Working memory of numerals in chimpanzees. Curr Biol 17:R1004–R1005

Izard V, Dehaene S (2008) Calibrating the mental number line. Cognition 106:1221–1247

Krueger LE (1982) Single judgments of numerosity. Percept Psychophys 31:175–182

Mandler G, Shebo BJ (1982) Subitizing: an analysis of its component processes. J Exp Psychol Gen 111:1–22

Matsuno T, Kawai N, Matsuzawa T (2006) Color recognition in chimpanzees. Cognitive development in chimpanzees. Springer, Tokyo

Matsuzawa T (1985) Use of numbers by a chimpanzee. Nature 315:57–59

Moyer RS, Landauer TK (1967) Time required for judgements of numerical inequality. Nature 215:1519–1520

Murofushi K (1987) Numerical matching behavior by a chimpanzee (Pan troglodytes): subitizing and analogue magnitude estimation. Jpn Psychol Res 39:140–153

Pepperberg IM, Gordon JD (2005) Number comprehension by a grey parrot (Psittacus erithacus), including a zero-like concept. J Comp Psychol 119:197–209

Ross J, Burr D (2008) The knowing visual self. Trends Cogn Sci 12:363–364

Sekuler R, Mierkiewicz D (1977) Children's judgments of numerical inequality. Child Dev 48:630–633

Siegler RS, Opfer JE (2003) The development of numerical estimation: evidence for multiple representations of numerical quantity. Psychol Sci 14:237–243

Stevens SS (1961) The psychophysics of sensory function. Sensory communication. W. A. Rosenblith. MIT press, Cambridge

Terrace HS, Son LK, Brannon EM (2003) Serial expertise of rhesus macaques. Psychol Sci 14:66–73

Treichler FR, Van Tilburg D (1999) Training requirements and retention characteristics of serial list organization by macaque monkeys. Anim Cogn 2

Treichler FR, Raghanti MA, Van Tilburg DN (2003) Linking of serially ordered lists by macaque monkeys (Macaca mulatta): list position influences. J Exp Psychol Anim Behav Process 29:211–221

Washburn DA, Rumbaugh DM (1991) Ordinal judgments of numerical symbols by macaques (Macaca mulatta). Psychol Sci 2:190–193

Weber EH (1843) De pulsu, resorptione, auditu et tactu annotationes anatomicae et physiologicae. C.F. Koehler, Leipzig

Whalen J, Gallistel C, Gelman R (1999) Nonverbal counting in humans: the psychophysics of number representation. Psychol Sci 10:130–137