Precursors to Number: Making the Most of Continuous Amount

Peter Drake
pedrake@cs.indiana.edu
Department of Computer Science, Indiana University, Bloomington; Lindley Hall 215
Bloomington, IN 47405 USA
Kelly Mix
kmix@indiana.edu
Department of Psychology, Indiana University, Bloomington; 1101 E. 10th Street
Bloomington, IN 47405 USA
Melissa Clearfield
mclearfi@indiana.edu
Department of Psychology, Indiana University, Bloomington; 1101 E. 10th Street
Bloomington, IN 47405 USA
May 6, 2000

Abstract

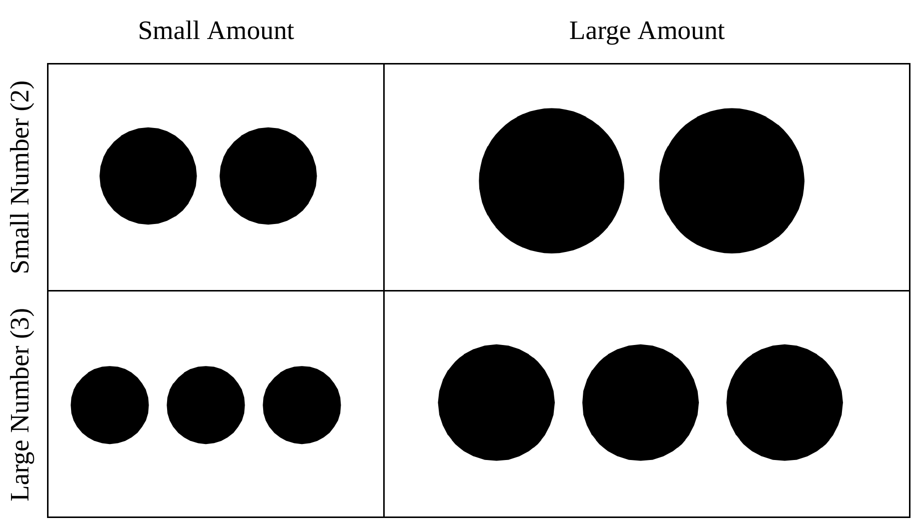
How does our understanding of number develop? There is evidence suggesting that even infants have primitive concepts of “more”, “less”, and “the same”. Some researchers have concluded that humans have an innate number sense, present from birth. In this paper, we present a two-part model which explains these results in terms of continuous amount. The first part is a quantitative model addressing the results of infant habituation studies. The second, more tentative part of the model addresses object individuation, subitizing, and number estimation.

Amount vs Number

In this paper, we use the word *amount* to refer to the total area of the objects in view; this is a continuous quantity. *Number*, a discrete quantity, refers to how many objects are present. As shown in Table 1, these two aspects can be varied independently.

A complete model would have to take into account other features, such as total contour length (edge length), shape, and color. Except where otherwise stated, we disregard these details.

Table 1: Amount vs Number. Both pictures in each column have the same total area.



Habituation Studies on Infant Numerical Abilities

Three studies are addressed directly by the first part of our model: Starkey & Cooper (1980), Antell & Keating (1983), and Clearfield & Mix (1999). All three studies use the same habituation paradigm, described below.

An infant is shown a series of images of black circles or squares on a white background, such as those in Table 1.

The infant is shown several more images. They may differ in arrangement, but they are the same on some critical dimension, such as the number of dots. If the infant habituates (stops looking at new images as long), this is taken as evidence that the infant detected the invariant property and became bored with it.

After habituation, the infant is shown a test image which differs on the critical dimension. If the infant dishabituates (spends significantly more time looking at the test image), he or she presumably noticed that the property changed—the test image is new and exciting. If the infant does not dishabituate, he or she presumably did not notice anything special about the test image.

In Starkey & Cooper's study, 22-week-old infants who had been habituated on images of 2 dots dishabituated when tested on images of 3 dots (and vice versa). Infants, it appeared, can tell the difference between 2 and 3.

To discount the possibility that the infants were simply reacting to amount, Starkey & Cooper tried 4 vs 6 dots. The relative difference between the images in this condition is the same as in the 2 vs 3 condition: the larger number has 1.5 times as much area as the smaller one. If infants are using amount of area, they should be at least as likely to dishabituate in the 4 vs 6 condition.

On the other hand, older children and adults have a much easier time enumerating sets of up to 3 or 4 objects than larger sets. Enumerating large sets requires use of an explicit, learned counting procedure; smaller sets can be enumerated quickly and subconsciously, through a process called "subitizing". The nature of subitizing remains controversial. In any case, if infants are subitizing, the 2 vs 3 condition should be more likely to produce dishabituation.

In fact, Starkey & Cooper's subjects did *not* dishabituate in the 4 vs 6 condition. It was therefore suggested that subitizing may be innate.

Antell & Keating replicated these results in newborns (less than a week old).

Clearfield & Mix's study reexamined the amount hypothesis. By changing the sizes of the objects (squares) in their displays, they were able to independently vary the number of objects and the total amount of area and contour length. (This study used 6- to 8-month-old infants, and only the 2 vs 3 condition was considered.)

They found that if the number remained the same, but the total area and contour length changed significantly, the infants dishabituated. Moreover, the infants *did not* dishabituate if the test image had approximately the same total area and contour length as the habituation image, *even if the number of objects was different*. In other words, infants do not appear to distinguish between 2 large objects and 3 small ones.

One problem remains: if infants are using amount to discriminate quantities, why don't they dishabituate in the 4 vs 6 condition?

Clearfield & Mix proposed that the 4 vs 6 displays might simply contain too much visual complexity. There is no question of comparison; the infants are overwhelmed by the displays.

The current model focuses instead on the way amount is represented internally[1] by the infants.

If the perceived magnitude grows linearly with the actual amount, then the 4 vs 6 condition should be more likely to produce dishabituation than the 2 vs 3 condition, because the absolute difference in amount of stuff is larger in the former condition.

Fechner (1860) suggested that the perceived amount grows as the logarithm of the actual amount. This almost does the trick, but not quite: whenever ratios are the same (as in the 2 vs 3 and 4 vs 6 conditions), differences of logarithms are equal. This would make the two conditions equally likely to produce dishabituation.

To explain the greater difficulty of the 4 vs 6 condition, we need a perception function which grows *more slowly* than a logarithm. One such function is the sigmoidal "squashing" activation function commonly used in connectionist "neural network" models (Rumelhart, McClelland, et al., 1986).

A Quantitative Model

We presume that the photoreceptors (rods and cones) in the retina provide (indirect) input to a neuron (or, more likely, a group of neurons) which codes for the total area of the objects in view. If the total area is larger, this *area unit* is more active; if the area is smaller, it is less active.

This does not require that the image be individuated into objects or preprocessed in any other way. The area unit is actually recording the total amount of light (or lack thereof) received by the retina. For the studies mentioned in the previous section, which use simple, black objects on a white background, this is equivalent to the total area.

The activity of the area unit does not vary linearly with its input. Instead, it varies according to a function of the form:

$$f(x) = \frac{1}{1 + e^{-\alpha x}}$$

In this equation, x is the total input to the area unit, and α is a parameter of the model.

The consequence of all this is that the same difference is perceived as being smaller if the absolute amounts involved are larger. This is the well-known magnitude effect, and is shown in Figure 1.

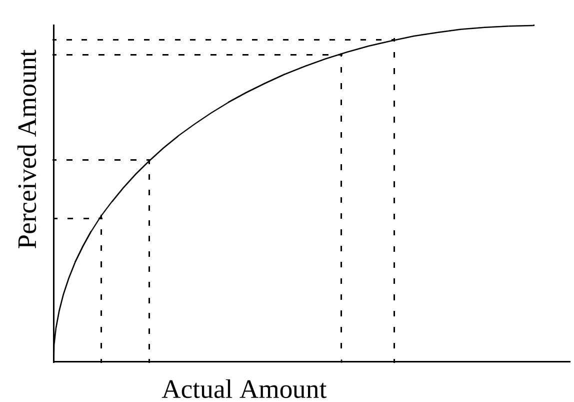


Figure 1: The magnitude effect as a result of a squashing function. Differences are perceived as smaller if the absolute magnitudes involved are larger.

[1] By "represented internally", we mean "represented as a pattern of neural activation". We do not mean to imply that infants are deliberately manipulating abstract, symbolic representations of amount.

In the context of the infant studies, our model predicts that the infant dishabituates if the perceived difference in area (between the habituation and test images) exceeds some threshold.

Just as the total activity of photoreceptors reflects the total visible area, the total activity of center-surround "edge detector" retina cells (ganglion cells) reflects the total visible contour length. A *contour unit*, analogous to the area unit, takes input from these neurons. A sufficiently large perceived difference here also induces dishabituation.

In order to make predictions, the model must be defined formally. Let $AREA_{hab}$ be the total area (in square radians of visual angle) in the habituation images, and $AREA_{test}$ be the total area in the test image. Similarly, let $CONT_{hab}$ be the total contour length (in radians) in the habituation image, and $CONT_{test}$ be the total contour length in the test image.

The model predicts that the infant will dishabituate if and only if

$$|f(AREA_{test}) - f(AREA_{hab})| > \tau_{AREA}$$

OR

$$|g(CONT_{test}) - g(CONT_{hab})| > \tau_{CONT}$$

where

$$f(x) = \frac{1}{1 + e^{-\alpha_{AREA}x}}$$

is the activation function for the area unit and

$$g(x) = \frac{1}{1 + e^{-\alpha_{CONT}x}}$$

is the activation function for the contour unit.

The model has four parameters: the thresholds τ_{AREA} and τ_{CONT} and the activation function sharpnesses α_{AREA} and α_{CONT}.

Setting the Parameters

We begin by calculating the total area and contour length of the stimuli from each study. The distance from the infants to the screen was different in each study, so we first convert all lengths into radians of visual angle. The data are shown in Table 2.

Table 2: Data from habituation studies. Areas are in square radians of visual angle; contour lengths are in radians. The test image is assumed to be larger (in terms of area and/or contour length) than the habituation image; because of the absolute value in the formula, the model would make identical predictions for the converse condition.

	Dishabituation Condition				No Dishabituation Condition			
	Habituation		Test		Habituation		Test	
Study	Area	Contour	Area	Contour	Area	Contour	Area	Contour
Starkey & Cooper	0.00044	0.10	0.00065	0.16	0.00087	0.21	0.0013	0.31
Antell & Keating	0.0031	0.28	0.0046	0.42	0.0062	0.56	0.0093	0.84
Clearfield & Mix	0.0089	0.53	0.020	0.80	0.0089	0.53	0.0059	0.53

We were able to find parameters for which the model gives the correct predictions for all of these studies. In other words, at least one of the thresholds is exceeded in each dishabituation condition, and neither are exceeded in any no-dishabituation condition.

One satisfactory set of parameters is:

$$\alpha_{AREA} = 3000$$
$$\tau_{AREA} = 0.075$$
$$\alpha_{CONT} = 3.5$$
$$\tau_{CONT} = 0.075$$

Admittedly, there is some degree of coincidence involved in our being able to find parameters consistent will all three studies. Variables such as lighting level, size of the card on which the object appear, and subject age may affect these values. Still, it is satisfying that none of the studies disagree qualitatively with the model, and intriguing that the same value can be used for both thresholds.

Problems With the Quantitative Model

In addition to the variables just mentioned, we have some reservations about the quantitative model.

The model assumes that each image is registered in a single eye fixation. In fact, infants move their eyes around quite a bit while looking at an image. There are two ways this may not matter. First, if the infant is keeping a running average of the area and contour length in the image, minor eye movements should have little effect. Second, the infant may be building an internal map of the image, and then extracting area and contour information from this “mind’s eye” view.

Another problem arises from studies on visual complexity. Karmel (1969) has given evidence that infants prefer to look at pictures with a certain amount (varying with age) of contour length. The total contour lengths in question are so huge (tens of radians) that, after passing through the activation function in our model, they would be indistinguishable. If, as our model predicts, these images are indistinguishable, how could infants have a preference? Karmel’s data come from a significantly different paradigm, and additional factors such as visual frequency may be coming into play. Still, these results will eventually have to be addressed.

Estimation and Subitizing

The model presented thus far may go a long way to explaining infant habituation data, but it can’t be the whole story for adults. In addition to the explicit, sequential counting procedure, we are able to estimate number. This estimation ability appears to operate in parallel—unlike counting, it doesn’t take twice as long when the number of objects is twice as large. The magnitude effect appears here, too: estimation of large numbers is less accurate. Estimation is only precise within the subitizing range, up to 3 or 4 objects.

Before number can even be estimated, it is necessary to individuate objects. In this section, we propose a model of object individuation which can underlie the estimation ability.

Temporal Synchrony

Animal studies (Eckhorn et al, 1988) have suggested an intriguing hypothesis about the visual system: in certain parts of the brain, cells which are responding to the same object fire at the same time. This is shown in Figure 2.

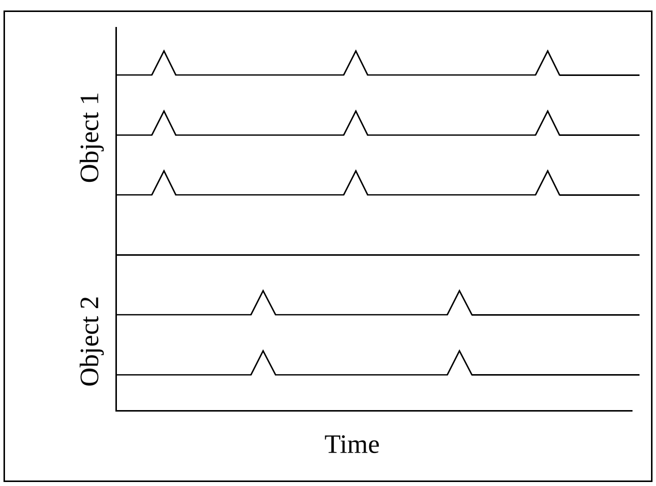


Figure 2: Temporal synchrony. Three cells are responding to the first object, two to the second. Cells responding to the same object fire at the same time.

Because the neural hardware is inherently noisy, there is a limit to the number of synchronized phases that can be kept distinct. This has implications regarding parallel vs sequential visual search, attention, variable binding, short-term memory capacity (the magical number 7 +/– 2), and other areas of cognitive science. In the rest of this section, we explore how temporal synchrony may aid in number estimation.

Subitizing

Temporal synchrony provides a simple explanation for the subitizing phenomenon. Suppose there is a unit which fires whenever any other unit fires. This *subitizing unit* repolarizes faster than the other units, but not so fast that it fires more than once in response to a synchronized pulse. This is shown in Figure 3.

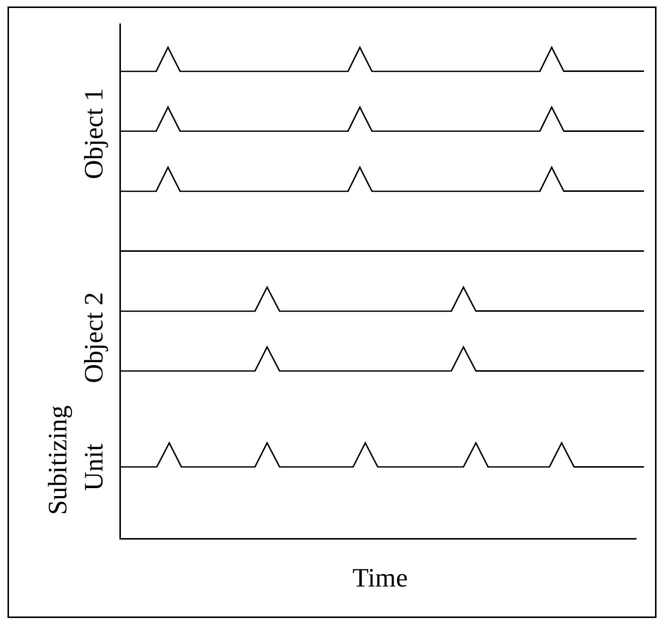


Figure 3: Subitizing with temporal synchrony. The subitizing unit fires whenever any other unit fires.

The frequency with which the subitizing unit fires encodes the number of objects visible. Beyond 3 or 4 objects, the phases begin to blur together; the subitizing unit fires at its maximum rate and the subject perceives “many” objects.

Estimation

Numbers beyond the subitizing range can still be estimated with the help of temporal synchrony. If a unit accepts input only from others firing at a particular phase, it only receives input regarding one object. Even if there are too many objects to subitize, this can provide some useful information: the size of a typical object. If this amount is used to scale the total amount of area visible (effectively dividing the total by the size of one object), a continuous representation of the number of objects is produced. This is not a terribly accurate mechanism (it suffers from the magnitude effect), but it is much faster than counting.

Related Work

An alternative model of nonverbal numerical abilities is the accumulator model proposed by Meck & Church (1983; Gallistel & Gelman, 1992). This model proposes an accumulator which integrates over time. As each object is perceived, the activity of the accumulator increases by a fixed amount. The accumulator's activity then serves as a representation of number.

The imprecision of larger numbers is explained as variability in the pulses passed to the accumulator. The more pulses there have been, the less accurate the resulting value in the accumulator.

Our model differs in two ways. First, we explain the lower precision of larger amounts with the squashing function. We have difficulty conceiving of a neurally plausible accumulator which is capable of both taking on very large values and providing precision for small values.

A second difference is that our model is strictly parallel, while the accumulator model is sequential: the stimuli are "fed into" the accumulator one after another. In a static image, this would require a pointing strategy, with the infant carefully "counting" each item exactly once. Since this is not a trivial task even for 3-year-olds (Fuson, 1988), it is difficult to believe that infants would have this ability. Our model does not ask so much; indeed, the quantitative model does not even require the infant to break the image down into separate objects.

Conclusions and Future Work

We have presented a two-part model of protonumerical abilities. The model reproduces human data on numerical perception without any explicit counting. The abilities granted by the model may provide useful grounding to children as they learn conventional counting.

The quantitative model predicts that infants will dishabituate to a sufficiently large change in either total area or total contour length. The exact meaning of "sufficiently large" depends on four parameters, and we have found values for these parameters which are consistent with several existing studies.

The quantitative model makes an interesting, counterintuitive prediction: infants will *not* dishabituate in Starkey & Cooper's 2 vs 3 condition with dots of certain sizes (e.g., extremely large ones). Conversely, the model predicts that infants *will* dishabituate in the 4 vs 6 condition for other dot sizes.

More specific predictions can be cautiously made based on the particular parameter values we found. The perceived difference between images is graphed as a function of stimulus size in Figure 4. Where this magnitude exceeds the threshold, habituation is predicted. Specific predictions are given in Table 3. We have begun empirical studies to test these predictions.

The second, more tentative part of the model accepts the temporal synchrony hypothesis of object individuation. Each visible object (or some of them, if there are too many) is bound to a particular phase. Within the subitizing range, the density of the phases indicates the number of objects. Beyond this range, the amount of area present at a particular phase indicates the size of an individual object, which can in turn be used to estimate the number of objects.

The second part of the model makes a less surprising prediction: estimating the number of objects visible should be difficult if the objects vary greatly in size.

We are now working on a connectionist implementation of our model, based on Gasser and Colunga's (1997) Playpen model of object individuation and spatial relations.

Acknowledgments

We wish to thank the following for their comments and encouragement: Mike Gasser, Deborah Alterman, Paul Purdom, Heather Drake, and Dan Friedman.

References

Antell, S. E., & Keating, D. P. (1983). Perception of numerical invariance in neonates. *Child Development*, *54*, 695-701.

Clearfield, M. W., & Mix, K. S. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychological Science*, *10*, 408-411.

Eckhorn, R., et al. (1988). Coherent Oscillations: A Mechanism of Feature Linking in the Visual Cortex? *Biological Cybernetics*, *60*, 121-130.

Fechner, G. (1860). In. D. H. Howes & E. G. Boring (Eds.), *Elements of Psychophysics, volume 1*. New York: Holt, Rinehart, and Winston. Cited in Kandel, Schwartz, & Jessell, 1991.

Fuson, K. C. (1988). *Children's Counting and Concepts of Number.* New York: Springer-Verlag.

Gasser, M. & Colunga, E. (1997). Playpen: Toward an Architecture for Modeling the Development of Spatial Cognition. Technical Report #195, Indiana University Cognitive Science Program.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, *44*, 43-74.

Kandel, E. R., Schwartz, J. H., & Jessell, T. M., Eds. (1991). *Principles of Neural Science, 3rd edition*. New York: Elsevier.

Karmel, B. Z. (1969). The effect of age, complexity, and amount of contour on pattern preferences in human infants. *Journal of Experimental Child Psychology, 7,* 339-354.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes, 9,* 320-334.

Rumelhart, D. E., McClelland, J. L., et al. (1986). *Parallel Distributed Processing*. Cambridge: MIT Press.

Starkey, P., & Cooper, R. (1980). Perception of numbers by human infants. *Science, 210*, 1033-1034.

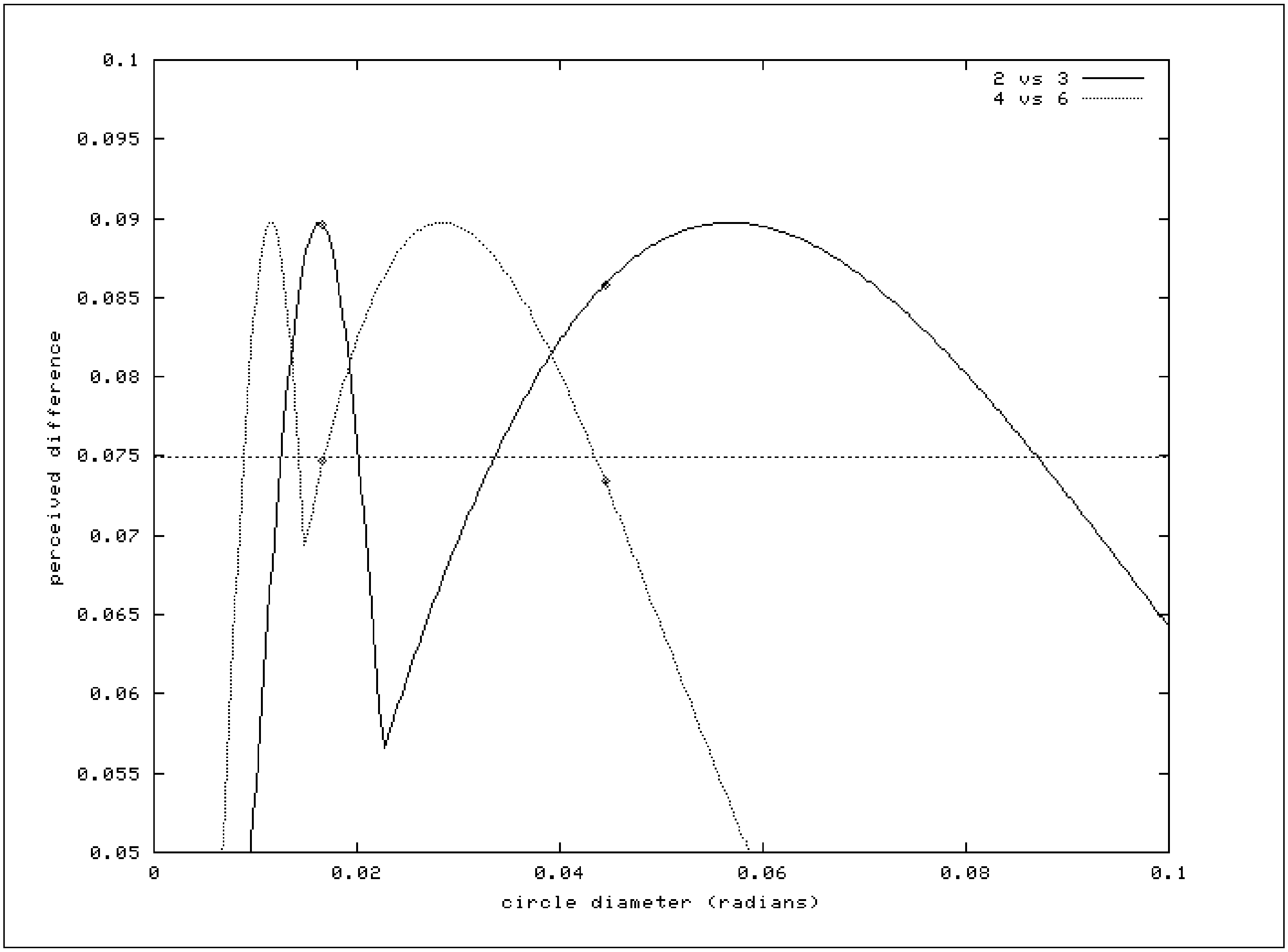


Figure 4: Perceived difference between images as a function of circular stimulus diameter. The M-shaped curves result from taking the maximum of the perceived area difference and the perceived contour difference. The marked points are the data from Starkey & Cooper (around 0.017 radians) and Antell & Keating (around 0.044 radians).

Table 3: Predictions of the model.

Table 3: Predictions of the model.

Circle Diameter		Dishabituation?	
radians	cm @ 60cm	2 vs 3	4 vs 6
< 0.0088	< 0.53	no	no
0.0099 - 0.012	0.53 - 0.72	no	yes
0.012 - 0.014	0.72 - 0.84	yes	yes
0.014 - 0.017 (includes Starkey & Cooper)	0.84 - 1.0	yes	no
0.017 - 0.020	1.0 - 1.2	yes	yes
0.020 - 0.034	1.2 - 2.0	no	yes
0.034 - 0.043	2.0 - 2.6	yes	yes
0.043 - 0.087 (includes Antell & Keating)	2.6 - 5.2	yes	no
> 0.087	> 5.2	no	no