Exploring the roles of saturating and supersaturating contrast-response functions in conjunction detection and contrast coding

Keith A. May  
Department of Computer Science, UCL, London, UK

Li Zhaoping  
Department of Computer Science, UCL, London, UK

J. W. Peirce (2007, p. 1) has proposed that saturating contrast-response functions in V1 and V2 may form “a critical part of the selective detection of compound stimuli over their components” and that supersaturating (non-monotonic) functions allow even greater conjunction selectivity. Here, we argue that saturating and supersaturating contrast-response functions cannot be exploited by conjunction detectors in the way that Peirce proposes. First, the advantage of these functions only applies to conjunctions with components of lower contrast than the equivalent non-conjunction stimulus, e.g., plaid (conjunctions) vs. gratings (non-conjunctions); most types of conjunction do not have this property. Second, in many experiments, conjunction and non-conjunction components have identical contrast, sampling the contrast-response function at a single point, so the function’s shape is irrelevant. Third, Peirce considered only maximum-contrast stimuli, whereas contrasts in natural scenes are low, corresponding to a contrast-response function’s expansive region; we show that, for naturally occurring contrasts, Peirce’s plaid detector would generally respond more weakly to plaid than to gratings. We also reassess Peirce’s claim that supersaturating contrast-response functions are suboptimal for contrast coding; we argue that supersaturation improves contrast coding, and that the multiplicity of supersaturation levels reflects varying trade-offs between contrast coding and coding of other features.

Keywords: saturation, supersaturation, conjunction, AND gate, multiplication, contrast coding

Introduction

Using a novel and elegant adaptation paradigm, Peirce and colleagues have obtained evidence that the human visual system contains detectors for at least two types of conjunction: plaid (conjunctions of sine-wave gratings of different orientation; Peirce & Taylor, 2006) and curves (conjunctions of spatially adjacent elements of similar orientation; Hancock & Peirce, 2008). But, little is known about how these conjunction detectors are implemented in the brain.

“Conjunction detection” is the detection of specific combinations of features. Formally, a conjunction detector should operate like an AND gate, responding only when all the components of the conjunction are present. Many properties of the AND operation are analogous (or isomorphic) to those of multiplication: Like an AND gate, multiplication gives a non-zero response if and only if all the input units are non-zero (see, for example, Schnupp & King, 2001). Because of this, several researchers have proposed conjunction detectors that multiply the outputs of units tuned to the components of the conjunction (e.g., Gheorghiu & Kingdom, 2009; May & Hess, 2010; Poirier & Wilson, 2006).

The difficulty with providing a neural implementation of a multiplicative conjunction detector is that feed-forward synaptic inputs to visual neurons are generally considered to be additive (Carandini & Heeger, 1994; Jagadeesh, Wheat, & Ferster, 1993; Movshon, Thompson, & Tolhurst, 1978), which is more like a logical OR operation, giving a non-zero response when any of the inputs are non-zero (Schnupp & King, 2001). However, Peirce (2007) argued that the saturating or supersaturating response non-linearities shown by V1 and V2 cells allow a simple summing circuit to behave like an AND gate, giving considerably stronger responses to conjunctions than non-conjunctions.

The specific example that Peirce gave to illustrate this idea was detection of a maximum-contrast plaid (i.e., the sum of two gratings), compared with detection of the equivalent non-conjunction stimulus, a maximum-contrast grating aligned with one of the plaid components. In Peirce’s proposed conjunction detector, there are two input neurons that are identical, except that one is tuned to the orientation of one plaid component and the other is tuned to the orientation of the other plaid component. Each neuron responds with a mean firing rate of \( r(c) \), where \( r \) is the contrast-response function, and \( c \) is the contrast of the stimulus component to which the neuron is tuned. The outputs of the two input neurons are then summed to give the output of the plaid detector. For a maximum-contrast plaid, the contrast, \( c \), of each component is 0.5, while, for a maximum-contrast grating, the contrast of the single
component is 1 (see Appendix A for an explanation of this). The output of the plaid detector in response to the plaid is therefore \( r(0.5) + r(0.5) = 2r(0.5) \), while the output in response to the grating is \( r(1) + r(0) = r(1) \), assuming zero response to zero contrast.

Peirce defined a conjunction selectivity index (CSI), which gives, for Peirce’s (2007) theoretical plaid detector described above, the ratio of plaid to grating response minus 1:

\[
\text{CSI} = \frac{2r(0.5)}{r(1)} - 1. 
\]

The CSI takes a positive value if the response to the plaid is greater than the response to the grating, a negative value if the plaid response is lower than the grating response, and zero if there is no difference. If the response of each input neuron is proportional to contrast, then \( 2r(0.5) = r(1) \), so the plaid response equals the grating response (CSI = 0). Most of the V1 and V2 neurons that Peirce (2007) examined had contrast-response functions that either saturated (i.e., started to level off at high contrasts) or supersaturated (i.e., the response initially increased with increasing contrast but then reached a peak and declined with further increases in contrast). These neurons tended to give \( 2r(0.5) > r(1) \), so the plaid response was greater than the grating response (CSI > 0). On this basis, Peirce concluded that their contrast-response functions facilitate conjunction detection.

One problem with this argument is that it depends on the assumption that the components of the conjunction have lower contrast than the equivalent non-conjunction stimulus: The advantage of a saturating or supersaturating contrast-response function then arises because it gives a disproportionately high response to low inputs. But, for most conjunctions of simple features, such as a conjunction of edge elements to form a curved edge segment or corner, we would not expect the contrast of the components to depend on whether or not they were part of a conjunction. Indeed, in the majority of studies on integration of oriented elements into contours or curves, all the elements have had the same contrast, whether or not they were part of a conjunction. For example, in Hancock and Peirce’s (2008) Experiment 2, subjects adapted to maximum-contrast grating patches that were presented either in isolation or together as part of a curved contour segment; a difference between these adaptation conditions in perceived shape of a test contour indicated the presence of curve detectors selective for a specific conjunction of components. Hancock and Peirce (p. 8) argued that a sum of saturating responses could create an AND gate for detection of these curve segments, although Peirce (2007, p. 8) suggested that such a mechanism would be merely advantageous, but not necessary, for detection of this kind of conjunction. It is our view that, in a simple summing circuit like this, the shape of the contrast-response function can play no role whatsoever in detection of this kind of conjunction. Consider a conjunction with components of contrast \( c \). If we assume that each input neuron responds to only one of the stimulus components, then the response of the conjunction detector to a single component presented in isolation is \( r(c) \), and the response to the conjunction stimulus is \( r(c) + r(c) = 2r(c) \). The ratio of conjunction response to isolated-component response is 2, whatever form the contrast-response function takes: Since all the components have the same contrast in both conjunction and non-conjunction stimuli, they all sample the contrast-response function at the same point, so its shape is irrelevant to detection of the conjunction in Peirce’s model.

So the proposal that saturating contrast-response functions can be exploited by a summing conjunction detector applies only to plaid-type conjunctions, in which the components have lower contrast than the equivalent non-conjunction stimulus. But, even then, the advantage of a saturating function appears only at high contrasts, where the function saturates. Peirce considered only the case of maximum-contrast stimuli, which (for a saturating contrast-response function) maximize the predicted ratio of plaid to grating response, leading to a high CSI as defined in Equation 1. Natural images are biased toward low contrasts, so stimulus components with contrasts close to maximum rarely occur in the natural environment. Contrast distributions in natural images have been reported to peak at zero (Brady & Field, 2000; Tadmor & Tolhurst, 2000), but others have reported a non-zero peak of around 0.1 (Clatworthy, Chirimuuta, Lauritzen, & Tolhurst, 2003). As pointed out to us by D. J. Tolhurst (personal communication), the apparent inconsistency of Clatworthy et al.’s result with the others can be explained by the fact that Clatworthy et al. used equal-width histogram bins on a log contrast axis, whereas the others used equal-width bins on a linear contrast axis; see Appendix B for verification of this explanation. Whichever distribution we take as the naturally occurring one, it is clear that most components of natural images have contrasts much closer to zero than to 1. Therefore, to assess the validity of Peirce’s proposed conjunction detection scheme, we need to examine how it performs across the whole range of contrasts, especially low ones. We show that Peirce’s proposed conjunction detector favors conjunctions only when the contrast is very high; with contrast levels typical of those found in natural scenes, it usually shows a weaker response to plaid than to gratings of the same Michelson contrast.

The limited success of Peirce’s conjunction detector raises two questions: First, how are conjunctions detected, and second, what are the roles of saturating
and supersaturating contrast-response functions in the visual cortex, if they do not facilitate conjunction detection? These two issues are explored in the Discussion section.

**Methods**

To examine how Peirce’s (2007) conjunction detector performs across the whole range of contrasts, we generalized the CSI to plaids and gratings of any contrast. The generalized CSI gives the response of the conjunction detector to a plaid of Michelson contrast \( c_M \), divided by the response to a grating of Michelson contrast \( c_M \), minus 1:

\[
\text{CSI}(c_M) = \frac{2r(c_M/2)}{r(c_M) + r(0)} - 1. \tag{2}
\]

Peirce’s (2007) CSI (Equation 1) is equivalent to our CSI(1), with \( r(0) = 0 \). In the generalized CSI, it is important to allow for the possibility that \( r(0) \neq 0 \), so that the estimated ratio of plaid response to grating response remains accurate at low contrasts for neurons that have a non-zero response to zero contrast.

To obtain some representative contrast-response functions \( r \), we fitted Peirce’s (2007) modified Naka–Rushton function to the data given in Peirce’s Figure 2, which shows responses to a range of contrasts for six neurons with different levels of saturation, from weak saturation to extreme supersaturation. Given that these neurons were selected by Peirce, they were unlikely to show any selection bias in our favor.

Peirce’s modified Naka–Rushton function provides an excellent fit to the contrast-response function of both supersaturating and non-supersaturating neurons. It has the following form, where \( c \) is the contrast of the stimulus component to which the cell is tuned:

\[
r(c) = \frac{Ac^q}{c_50^q + c^q} + B. \tag{3}
\]

\( B \) is the baseline firing rate at contrast \( c = 0 \); \( A \) controls the amplitude; \( q \) controls the steepness of the initial rise in response; \( c_50 \) controls the position of the response function along a log contrast axis; and \( s \) controls the level of supersaturation. With \( 0 < s < 1 \), the neuron never completely saturates: The function asymptotes toward \( r = Ac^q(1-s) + B \) for large \( c \). With \( s = 1 \), there is saturation but no supersaturation: The function reduces to the standard Naka–Rushton function, which rises with increasing contrast toward an asymptote of \( r = A + B \); in this case, \( c_50 \) is the semi-saturation contrast, the contrast at which the response is halfway between \( A \) and \( B \). With \( s > 1 \), the function supersaturates, rising to a peak at a contrast of

\[
\text{cpeak} = c_50/(s-1)^{(1/q)}, \tag{4}
\]

and then decaying with further increases in contrast toward an asymptote of \( r = B \). Although, strictly speaking, \( c_50 \) in Equation 3 is the “semi-saturation contrast” only when \( s = 1 \), we will use this phrase more loosely to refer to this parameter whatever the value of \( s \).

We fitted the five parameters of Equation 3 to each neuron using simplex minimization (Nelder & Mead, 1965) of the squared difference between the function and the data. \( B \) was constrained to be non-negative. The resulting contrast-response functions, \( r \), were substituted into Equation 2 to calculate the CSI for each neuron across the whole range of contrasts, 0 to 1.

**Results**

In Figure 1, the left column shows the contrast-response function of Equation 3 fitted to each of the six neurons shown in Peirce’s (2007) Figure 2. The fitted parameters are given in Table 1. The right column of Figure 1 shows, for each neuron, the CSI as a function of contrast, \( c \). In each case, the CSI is positive for high contrasts but negative for low contrasts. The critical contrast, \( c_0 \), at which the CSI crosses zero, is given by

\[
c_0 = c_50 \left( \frac{2^q - 2}{2 - 2^{1-s}} \right)^{1/(sq)}. \tag{5}
\]

The \( c_0 \) values for the neurons (given in Table 1 and indicated by the red vertical lines in Figure 1) range from 0.0980 to 0.623, with a mean of 0.336. Whether we sample our contrasts from a natural distribution that peaks at zero (Brady & Field, 2000; Tadmor & Tolhurst, 2000) or close to 0.1 (Clatworthy et al., 2003), this set of contrast-response functions will usually give negative CSI values, indicating weaker responses to plaids than to gratings of the same Michelson contrast for Peirce’s (2007) plaid detector.

One might argue that this is an unfair test of the plaid detector because these neurons had been measured in laboratory conditions in which they were exposed to unnaturally high contrasts, going right up to maximum contrast; in natural vision, divisive contrast gain control mechanisms would compress each contrast-response function horizontally so that the contrast, \( c_0 \), corresponding to its point of inflection (i.e., point of steepest positive gradient) was close to the mean contrast of natural images (Ohzawa, Sclar, & Freeman, 1985), and this would mean...
that it saturated at lower contrasts than in the laboratory. However, gain control would also reduce \( c_0 \) (the critical contrast at which the CSI switches from positive to negative) by the same factor; this is because the CSI function is calculated from the contrast-response function and inherits any horizontal compression applied to the latter. Thus, the ratio \( c_0/c_1 \) would be unchanged by divisive gain control.

The point of inflection of each neuron is given in Table 1 and indicated in Figure 1, both for linear and log contrast axes. With only one exception (neuron f, log axes), \( c_0 \) is higher than \( c_1 \). This means that, even if the contrast-response functions of these neurons did compress horizontally so that their points of inflection occurred at the prevailing mean contrast level, \( c_0 \) would still be higher than the mean contrast, so most contrasts would still fall below \( c_0 \), giving a negative CSI for most natural stimuli. So, with or without contrast gain control, it is clear that, for naturally occurring contrast levels, Peirce’s (2007) plaid detector would give a weaker response to plaids than to gratings of the same Michelson contrast most of the time.

**Discussion**

We have shown that, in Peirce’s proposed conjunction detection circuit, which sums the outputs of V1 or V2 neurons, the shapes of typical contrast-response functions do not facilitate conjunction detection for most ecologically plausible inputs. When the conjunction components and non-conjunction stimuli all have the same contrast, they sample the contrast-response function at a single point, so the shape of the function is irrelevant. Alternatively, for plaid-type conjunctions, in which the components have a lower contrast than the equivalent non-conjunction stimulus, the summing plaid detector would usually give a weaker, not stronger, response to plaid than to gratings of the same contrast when they are presented at naturally occurring contrast levels. Furthermore, as noted by Peirce (2007), the CSI does not take account of cross-orientation suppression (Bonds, 1989; Freeman, Durand, Kiper, & Carandini, 2002; Morrone, Burr, & Maffei, 1982; Priebe & Ferster, 2006) and, therefore, greatly overestimates the response to the plaid; inclusion of the effect of cross-orientation suppression would extend even further the range of contrasts over which the plaid response was lower than the grating response.

In Boolean logic, a conjunction is an AND operation, and a conjunction detector should respond when both components are present. In Peirce’s circuit, the detector responds when both components are present and the response is therefore given by the product of the components’ responses. This is equivalent to the product of the probabilities of the individual components being present, which is the definition of the conjunction of two events. Therefore, the response of the conjunction detector is given by

\[
\text{Response} = \prod_{i=1}^{N} \text{Response}_i
\]

where \( N \) is the number of components and \( \text{Response}_i \) is the response of the component. This is in contrast to the summing plaid detector, which responds when either component is present, and the response is given by

\[
\text{Response} = \sum_{i=1}^{N} \text{Response}_i
\]

The blue and green vertical lines in each panel indicate the contrasts of the points of inflection, \( c_{I,\text{linear}} \) and \( c_{I,\text{log}} \), respectively (as defined in Table 1). The red vertical line indicates the contrast at which the CSI switches from positive to negative (\( c_0 \) in Equation 5).
components of the conjunction are presented but not to a single component on its own. The output of Peirce’s circuit in response to a conjunction is the sum of the responses to the individual components, which is more like an OR gate than an AND gate. To achieve AND-like behavior, it would be better to multiply the outputs of the input neurons (Schnupp & King, 2001).

There are already several lines of evidence for multiplication or AND gating in conjunction detection. Gheorghiu and Kingdom (2009) presented psychophysical evidence for multiplicative units for processing curves (i.e., conjunctions of edge elements over space), and van Santen and Sperling (1984) reported psychophysical evidence for multiplication in motion detection, which can be considered to be detection of a conjunction over space and time. van Santen and Sperling (1985) also showed that several motion detection models are formally equivalent to van Santen and Sperling’s (1984) elaborated Reichardt detector, a key feature of which is multiplication of a signal from one spatial location with the delayed (or, more precisely, temporally filtered) signal from an adjacent spatial location.

There are plenty of reports of neurons behaving in a multiplicative or AND-like manner. Hubel and Wiesel (1962) reported that some cells in V1 “did not respond to stimulation of either eye alone but could be activated only by simultaneous stimulation of the two eyes” (Hubel & Wiesel, 1962, p. 124). Barlow and Levick (1965) showed that the behavior of direction-selective ganglion cells in the rabbit retina could be described as an AND–NOT operation. Many reports of multiplicative behavior have been qualitative, but Peña and Konishi (2001) provided quantitative support for multiplication in space-specific neurons in the owl. These neurons respond to sound from a particular direction, selectively responding to the conjunction of a particular horizontal position [measured from the interaural time difference (ITD)] and vertical position [measured from the interaural level difference (ILD), due to an asymmetry in the left and right ear openings and flaps (Knudsen & Konishi, 1979; Payne, 1971)]. Peña and Konishi measured membrane potential as a function of different combinations of ITD and ILD, and found that membrane potential was well described as a function of the product of ITD and ILD inputs, not the sum.

The following discussion is split into two parts. In the first part, we examine several different neuronal mechanisms that might be responsible for conjunction detection. None of the successful conjunction detection mechanisms that we looked at depend on a saturating or super-saturating non-linearity, so, in the second part, we examine the roles that these non-linearities might have.

**How are conjunctions detected?**

Peirce’s (2007) summing circuit for conjunction detection seems to have been motivated by the largely additive nature of feed-forward synaptic inputs to visual neurons (Carandini & Heeger, 1994; Jagadeesh et al., 1993; Movshon et al., 1978). In this section, we consider what mechanisms might create the kind of multiplicative or AND-like non-linearities required for detecting conjunctions. For a more detailed review of the different ways to multiply using neurons, see Koch and Poggio (1992).

**Multiplication by summing logarithms**

As several authors have noted (e.g., Gheorghiu & Kingdom, 2009; Koch & Poggio, 1992; Tal & Schwartz, 1997), multiplication can be achieved by summing logarithms:

\[ a \times b = \exp(\ln a + \ln b). \]  

The logarithm could be approximated by the compressive, saturating, portion of a neuron’s contrast-response function, and exponentiation could be approximated by the

---

**Table 1.** Data for the six neurons in Figure 1. \( c_0 \) is the contrast at which the CSI changes from positive to negative, calculated according to Equation 5. \( c_{\text{linear}} \) is the contrast corresponding to the point of inflection on the rising portion of the contrast-response function when plotted on a linear contrast axis, i.e., the first peak of positive gradient that we encounter as contrast is increased from zero; \( c_{\text{log}} \) is the contrast corresponding to the point of inflection when the contrast-response function is plotted on a log contrast axis.

<table>
<thead>
<tr>
<th>Neuron</th>
<th>A</th>
<th>B</th>
<th>( c_{50} )</th>
<th>q</th>
<th>s</th>
<th>( c_0 )</th>
<th>( c_{\text{linear}} )</th>
<th>( c_{\text{log}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>33.0</td>
<td>1.66</td>
<td>0.363</td>
<td>2.23</td>
<td>0.93</td>
<td>0.623</td>
<td>0.248</td>
<td>0.430</td>
</tr>
<tr>
<td>b</td>
<td>82.2</td>
<td>4.50</td>
<td>0.150</td>
<td>3.75</td>
<td>0.862</td>
<td>0.379</td>
<td>0.147</td>
<td>0.195</td>
</tr>
<tr>
<td>c</td>
<td>19.4</td>
<td>4.21</td>
<td>0.234</td>
<td>3.11</td>
<td>1.06</td>
<td>0.400</td>
<td>0.182</td>
<td>0.219</td>
</tr>
<tr>
<td>d</td>
<td>27.6</td>
<td>2.00 ( \times 10^{-5} )</td>
<td>0.0639</td>
<td>2.33</td>
<td>1.04</td>
<td>0.0980</td>
<td>0.0420</td>
<td>0.0599</td>
</tr>
<tr>
<td>e</td>
<td>16.0</td>
<td>9.61 ( \times 10^{-5} )</td>
<td>0.235</td>
<td>3.81</td>
<td>1.12</td>
<td>0.397</td>
<td>0.193</td>
<td>0.214</td>
</tr>
<tr>
<td>f</td>
<td>1.26</td>
<td>2.53</td>
<td>0.306</td>
<td>1.05</td>
<td>3.32</td>
<td>0.119</td>
<td>0.0586</td>
<td>0.140</td>
</tr>
</tbody>
</table>
expansive portion. If we can implement multiplication (and therefore AND-like behavior) by summing logarithms, then why can this not be achieved by Peirce’s model, which has a similar architecture (i.e., it applies a compressive non-linearity to each input, and then adds them together)? The reason is the difference in response to zero contrast. The critical feature of multiplication that makes it similar to an AND gate is that the output is zero when any of the inputs are zero; and the critical feature of the logarithm that achieves this is that the logarithm approaches $-\infty$ as the input tends to zero: If any input approaches zero, the sum of logs of the inputs will approach $-\infty$, and so, when the exponential function is applied to this sum, the output approaches zero. But, Peirce’s model uses realistic contrast-response functions, in which the response to zero contrast is close to zero. A neuron’s firing rate cannot drop below zero, i.e., log(1), and there is no plausible way of generating a $-\infty$ response to zero input, and so although the sum of logs could be used to multiply above-zero quantities with neurons, it cannot handle multiplication by zero, making it useless for implementing an AND-like conjunction detector.

**The Babylonian trick**

A more promising algorithm for multiplication is described by Equation 7, which is sometimes referred to as the “Babylonian trick” (Ferraro & Caeli, 1998; Gheorghiu & Kingdom, 2009; Zetzsche & Barth, 1990):

$$a \times b = \frac{1}{4} \left[ (a + b)^2 - (a - b)^2 \right].$$

(7)

The addition and subtraction can be carried out by synaptic excitation and inhibition. The squaring operation could be implemented by the expansive portion of the contrast-response function because, in Equation 3, the response, $r$, is close to being proportional to $c^b$ when $c$ is small compared with $c_{50}$, and the distribution of fitted $q$ values in V1 peaks at around 2 using the standard Naka–Rushton function (Albrecht & Hamilton, 1982; Geisler & Albrecht, 1997; Sclar, Maunsell, & Lennie, 1990). If this is the approach that the brain uses to perform multiplication for conjunction detection, then the shape of the contrast-response function does play a critical role, but it is the expansive portion of the function that is critical, not the saturating portion.

To be implemented by real neurons, the Babylonian trick needs a slight modification. Since the neural output cannot be negative, it is more accurately described as “half-squaring,” i.e., half-wave rectification (setting negative values to zero) followed by squaring (Heeger, 1992). This means that the $(a - b)^2$ term in Equation 7 cannot be evaluated by a single neuron when $a$ is free to vary above or below $b$. This problem can be easily dealt with by splitting the $(a - b)^2$ term into two half-squared terms, $|a - b|^2$ and $|b - a|^2$, where $|x| = \max(x, 0)$, $|a - b|^2$ carries the $(a - b)^2$ signal when $a > b$, and $|b - a|^2$ carries this signal when $a < b$. Assuming $a$ and $b$ to be positive, excitatory, inputs, the Babylonian trick then becomes

$$a \times b = \frac{1}{4} \left[ |a + b|^2 - |a - b|^2 - |b - a|^2 \right].$$

(8)

If we assume that $a$ and $b$ are the outputs of an initial layer of neurons, each tuned to one of the components of the conjunction, then the computation of Equation 8 would require two additional stages of neuronal processing: Layer 2 computes half-squared linear functions of the inputs $(|a + b|^2, |a - b|^2, |b - a|^2)$, and layer 3 computes a linear function of layer 2’s responses. This circuit produces a strong response when both stimulus components are present at sufficient contrast, and zero response when at least one component is missing; and this will still be true even if we introduce physiologically plausible non-linearities to the outputs of layers 1 and 3 (although in that case, the output will no longer be strictly proportional to $a \times b$).

**Multiplication by direct detection of conjunctions**

The previous section outlined a way in which neurons could carry out multiplication for the purpose of conjunction detection. Srinivasan and Bernard (1976) did the reverse, proposing a neural multiplication model that works by detecting conjunctions directly. Their circuit consists of two input neurons, $A$ and $B$, which make excitatory synapses onto an output neuron, $C$. $C$ has a threshold such that a spike from one of the input neurons is insufficient to trigger a spike in the output, but, if both input neurons produce a spike within a sufficiently short time period, their excitatory postsynaptic potentials sum to produce an above-threshold excitation, triggering a spike in the output neuron. Assuming that $A$ and $B$ fire independently, if the probability that input neuron $A$ spikes within a short time interval is $P(A)$, and the corresponding probability for neuron $B$ is $P(B)$, then the probability that they both spike within the same brief interval (thereby triggering a spike in the output neuron) is $P(A)P(B)$. Since the probability of getting a spike within a short time interval is proportional to the firing rate (Dayan & Abbott, 2001, p. 10), the output firing rate is proportional to the product of the two input firing rates. Using the same basic principle (probability multiplication), Solomon, Chubb, John, and Morgan (2005) demonstrated that multiplicative behavior could occur without explicit multiplication in the kind of psychophysical motion task that van Santen and Sperling (1984) had offered as evidence of multiplication in motion processing.
Synaptic interactions

The mechanisms described so far have assumed that synaptic inputs to a neuron sum linearly, so that a multiplication or AND operation must be generated by applying a non-linearity to the output. But, if two synapses are physically close to each other on the dendritic tree, or fall on the same direct path to the soma of the postsynaptic cell, the synaptic inputs can interact in highly non-linear ways (Koch, Poggio, & Torre, 1983, 1986; Segev & Parnas, 1983; Torre & Poggio, 1978), allowing the implementation of a variety of logical operations (Shepherd & Brayton, 1987). Torre and Poggio’s model showed that nearby excitatory and inhibitory synapses could interact to produce a multiplicative term that could serve to implement the AND–NOT gate proposed by Barlow and Levick (1965) to account for motion direction selectivity in the rabbit retina. Torre and Poggio’s model has subsequently received empirical support from several sources (Koch et al., 1986; Taylor, He, Levick, & Vaney, 2000).

Another likely locus of multiplication or AND gates is the receptors for N-methyl-D-aspartate (NMDA). Synaptic input to an NMDA receptor has little effect when the postsynaptic cell is hyperpolarized, but has an increasingly large effect on membrane conductance as the cell becomes more depolarized, which leads to a multiplicative, rather than additive, effect of NMDA receptor activation on membrane potential (Dingledine, 1983). This non-linearity occurs because of a voltage-dependent block of NMDA receptors by Mg²⁺ ions, which diminishes as the cell becomes depolarized (Mayer, Westbrook, & Guthrie, 1984; Nowak, Bregestovski, Ascher, Herbet, & Prochiantz, 1984).

Recurrent networks

Salinas and Abbott (1996) showed that multiplicative behavior could arise from a recurrently connected network of linear units. They were specifically modeling neurons in parietal area 7a for which the response to visual input is multiplied by a function of the gaze direction (Andersen, Bracewell, Barash, Gnadt, & Fogassi, 1990; Andersen, Essick, & Siegel, 1985), but their model could, in principle, be used to detect conjunctions.

The roles of saturating and supersaturating contrast-response functions

Possible roles of saturating contrast-response functions

Several possible roles of saturating (but not supersaturating) contrast-response functions have been proposed in the literature. As noted earlier, saturating functions could be used to approximate a sum of logs for multiplication of above-zero quantities (Tal & Schwartz, 1997), although this mechanism could not be used to detect conjunctions because it cannot handle multiplication by zero. In a different approach, Geisler and Albrecht (1995) argued that saturating contrast-response functions allow accurate decoding of all stimulus properties, except contrast, across a wide range of stimulus contrasts.

Others have used information theory to show that a saturating response may benefit contrast coding too. If we ignore neural costs, the goal of contrast coding should be to maximize the mutual information between the stimulus contrast and the neural response, i.e., the amount of information that the neural response tells us about the stimulus contrast. Mutual information in this case is the reduction in entropy (i.e., unpredictability) of the stimulus contrast that we achieve by receiving the neural response. Mutual information can be expressed as the entropy of the response minus the entropy of the noise so, assuming a constant noise entropy, we maximize the mutual information by maximizing the response entropy. For a single neuron, this is achieved by having a flat distribution of responses, so that the whole dynamic range of the neuron is used with equal probability (any deviation from a flat response distribution makes the neuron’s response more predictable, so it has a lower entropy). The contrast-response function that flattens the response distribution has the shape of the cumulative probability distribution of contrasts in the environment, which is a saturating function (Laughlin, 1981). Several types of cell have been found to have this property [large monopolar cells in the compound eye of the fly (Laughlin, 1981), X and Y cells in retina and lateral geniculate nucleus (LGN) of the cat (Tadmor & Tolhurst, 2000), and M cells in macaque LGN (Tadmor & Tolhurst, 2000)].

Most saturating V1 cells do not match the cumulative probability distribution of contrasts, as their contrast-response functions are too steep (Clatworthy et al., 2003). However, Gottschalk (2002) argued that typical V1 saturating contrast-response functions may optimize contrast coding if we consider the neural costs as well as mutual information. He defined a performance measure in which the costs of synaptic transmission and neural spiking were subtracted from the mutual information between the stimulus contrast and a neuron’s response (assuming a Gaussian stimulus distribution and additive Gaussian noise). Given certain restrictions on the forms of the neural cost terms, the gain function that maximized this performance measure was a Naka–Rushton function with exponent $q = 2$; the maximum response and semi-saturation contrast were determined by parameters of the cost terms. Similar functions, with steeper or shallower slopes, were obtained by altering the exponent on one of the neural cost terms.
Possible roles of supersaturating contrast-response functions

It is possible that at least some occurrences of supersaturation found in electrophysiological recordings merely reflect inadvertent activation of the suppressive surround that lies outside of the classical receptive field. In view of the limited time available to characterize a cell before losing it, it may not always be possible to make absolutely sure that the stimulus is fully within the receptive field. If the stimulus encroaches slightly onto the suppressive surround, the neurons within the surround might only be stimulated above threshold when the recorded cell has saturated. In this case, as the stimulus contrast is increased, the cell will first saturate, and then start to be inhibited as the suppressive surround rises above threshold (we thank Isabelle Mareschal for this suggestion). If this explanation is correct, then supersaturation may not have any functional role itself, being just an artifactual consequence of iso-feature suppression in the visual cortex, which is thought to underlie the computation of saliency for attentional selection (Jingling & Zhaoping, 2008; Knierim & Van Essen, 1992; Koene & Zhaoping, 2007; Li, 1999, 2002; Nothdurft, Gallant, & Van Essen, 1999; Zhaoping, 2008; Zhaoping & May, 2007; Zhaoping, May, & Koene, 2009).

Alternatively (or additionally), supersaturation may be a real phenomenon, not just an artifact as suggested in the previous paragraph. Only a few studies have examined supersaturation in any detail (Ledgeway, Zhan, Johnson, Song, & Baker, 2005; Li & Creutzfeldt, 1984; Peirce, 2007), and, of these, only Peirce (2007) gave much consideration to the role that this phenomenon might play. Here, we explore the effect of supersaturation on contrast coding.

As noted in the previous section, if we consider contrast coding by a single neuron, different approaches yield different “optimal” contrast-response functions, yet the derived functions all have two things in common: They are all saturating functions that increase monotonically. Since a supersaturating contrast-response function is non-monotonic, it cannot match any of these “optimal” functions, and because of this it has been argued that supersaturation is suboptimal for contrast coding (Peirce, 2007). But, the optimality constraints discussed in the previous section (which yielded monotonic contrast-response functions) considered only the information in a single neuron’s response. In the cortex, the limited dynamic range of each neuron suggests that the contrast code is distributed across a population of neurons (Albrecht & Hamilton, 1982; Chirimuuta, Clatworthy, & Tolhurst, 2003; Clatworthy et al., 2003; Teo & Heeger, 1994). In this case, supersaturation actually improves contrast coding. One reason for this is that a neuron can only contribute to discrimination of contrasts that fall on sloped regions of its response function; by supersaturating, a cell adds a downward slope to its response function, thereby adding an extra range of contrasts over which it can discriminate, without impairing its ability to discriminate contrasts over the initial rising slope. Although the response from an individual supersaturating neuron will be ambiguous as to whether the contrast falls above or below the peak of sensitivity, the signal can be disambiguated by taking into account the responses from other neurons in the population; similar considerations apply to population decoding of orientation (Coltheart, 1971) and spatial frequency (Blakemore, Nachmias, & Sutton, 1970; Blakemore & Sutton, 1969), which are both encoded in the cortex using non-monotonic response functions (Hubel & Wiesel, 1959, 1962, 1968; Maffei & Fiorentini, 1973).

The benefit of supersaturation in a population code is illustrated in Figure 2. The upper panel shows the contrast-response functions of a set of supersaturating neurons, as described by Peirce’s modified Naka–Rushton function (Equation 3) with zero baseline ($B$), exponent $q = 3$, and supersaturation parameter $s = 2$. This value of $s$ is a special case, giving a contrast-response function that is even-symmetric about $c_{\text{50}}$ on a log contrast axis, with a similar shape to a Gaussian. The neurons in this panel differ only in $c_{\text{50}}$, which shifts the contrast-response function left or right on a log contrast axis, but otherwise leaves it unchanged.

The vertical blue line in Figure 2 indicates a test contrast of 0.1 in linear contrast units. The only neurons that can play a significant role in discriminating between this and very similar contrasts are those for which the test contrast falls on a steeply sloped region. For the purposes of this informal illustration, we define “steep” to mean “greater than 10% of the maximum slope.” In the upper panel of Figure 2, there are ten neurons for which the test contrast falls on a steep region, indicated in red: five with peaks to the right of the test contrast and another five with peaks to the left. These two sets of neurons make equal contributions to discrimination of the test contrast.

The lower panel shows what happens when we remove the supersaturating region from each contrast-response function; these contrast-response functions were made by taking each function in the upper panel and replacing the downward-sloped region to the right of the peak with a flat region so that the response remains at maximum, giving a saturating function. Now there are only five neurons for which the test contrast falls on a steep region: The five neurons that peak to the right of the test contrast still make the same contribution as before, but the five neurons peaking to the left of the test contrast now contribute nothing to discrimination of the test contrast, yet they are firing at their maximum rate. In the supersaturating scheme in the upper panel, the weakly informative neurons (in black) have a very low firing rate; supersaturation would, therefore, give rise to more accurate contrast discrimination, while consuming less energy.

It is unusual to find a neuron that shows extreme supersaturation with $s$ as high as 2. We include this example only to demonstrate how much better contrast
coding would be if the visual system were like this, as a counterevidence to Peirce’s claim that supersaturation is suboptimal for contrast coding. More typical values of $s$ are around 1.05 or 1.1, like neurons c, d, and e in Figure 1 and Table 1. Figures 3 and 4 show populations of neurons identical to Figure 2, except with $s = 1.1$ in Figure 3 and $s = 1.05$ in Figure 4. We now show how to evaluate all of these coding schemes formally.

So far, we have presented an informal argument based on the slopes of the contrast-response functions. Another approach is to consider that, for accurate decoding of contrast, we need the probability of the obtained population response to be narrowly distributed across contrast, so that only a narrow range of contrasts could plausibly have given rise to the response. By adding the downward slopes to the response functions in Figures 2 to 4, we allow more neurons to change their average response as a result of a change of contrast, so that a small change of contrast would be more likely to give rise to a different population response; this has the effect of narrowing the probability distribution of the obtained population response.

This idea can be formalized in terms of the Fisher information, which sets an upper bound on the accuracy with which a signal can be decoded. The Fisher information, $I_F$, for contrast $c$ is given by

$$I_F(c) = \int d\mathbf{r} p(\mathbf{r}|c) \left( -\frac{\partial^2 \ln p(\mathbf{r}|c)}{\partial c^2} \right),$$

where $\mathbf{r}$ is a vector representing the responses of all the neurons, and $p(\mathbf{r}|c)$ is the probability of population response $\mathbf{r}$, given contrast $c$. The second derivative of the log likelihood in Equation 9 is a measure of curvature of the probability distribution, which in turn reflects how narrowly the probability is distributed across contrast. The integral in Equation 9 calculates the expected average curvature over many trials in response to contrast $c$. For a neural population with a sufficiently high spike count, the Fisher information is very close to the reciprocal of the variance of the contrast estimate when decoding the population response using maximum-likelihood estimation (Dayan & Abbott, 2001, p. 109).

We can use the Fisher information to evaluate the coding schemes illustrated in Figures 2 to 4. For a single Poisson-spiking neuron, it can be shown that the Fisher information is given by $I_F(c) = T\bar{r}'(c)^2 / r(c)$, where $T$ is the trial duration, $r$ is the contrast-response function (measured in spikes per unit time), and $r'$ is its first derivative with respect to contrast (Dayan & Abbott, 2001, Chapter 3). For a population of independent, Poisson-spiking neurons, the Fisher information is found by summing the Fisher information across the population:

$$I_P(c) = \sum_i r_i'(c)^2 / r_i(c).$$

Equation 10 supports the intuitive notion that, for accurate contrast decoding, we need the slope, $r'$, of the contrast-response function to be steep at the test contrast. In the upper panel of Figure 2, each neuron peaking to the right of the test contrast has a corresponding neuron that peaks the same distance from the test contrast but to the left. The pair have the same mean response and gradient magnitude as each other at the test contrast, so they make equal contributions to the sum in Equation 10. The move from supersaturation (upper panel) to saturation (lower panel) eliminates the contribution from the neurons that peak to left of the test contrast, and this halves the Fisher information.

We now use Equation 10 to evaluate the coding schemes shown in Figures 2 to 4. As noted by Peirce (2007), there is no simple relationship between the parameters of Equation 3 and the maximum response, so
we first introduce a more convenient, normalized, version of Peirce’s modified Naka–Rushton equation, in which the amplitude parameter, $A$, is replaced with $r_{\text{max}}$, the peak response difference from baseline:

$$r = r_{\text{max}} c_q (s - 1)^{1/s - 1} \frac{c^q}{c_{s0} + c_{sq}} + B. \quad (11)$$

The peak value of $r$ is $r_{\text{max}} + B$. It is not appropriate to use Equation 11 when $s < 1$ (i.e., non-saturating functions) because, in this case, there is no peak response (other than

Figure 3. The same as Figure 2 but with supersaturation parameter $s = 1.1$.

Figure 4. The same as Figure 2 but with supersaturation parameter $s = 1.05$.

Figure 5. Fisher information (top row of panels) and population spike rate (bottom row) as functions of test contrast for the coding schemes shown in Figures 2, 3, and 4. In each panel, super-saturating contrast-response functions are compared with the corresponding saturating functions, formed by replacing the supersaturating region with a flat region of response $r_{\text{max}}$. Each column of panels shows this comparison for one level of the supersaturation parameter ($s = 1.05, 1.1, 2$). Fisher information was found by finding the Fisher information for each neuron in the population, as described in the text, and then summing across all neurons, as in Equation 10. The population spike rate is the sum of the mean spike rate $r(u)$ across the population of neurons. The plotted values are normalized to make them independent of $r_{\text{max}}$ and the stimulus duration, $T$. 
the one imposed by the physical impossibility of exceeding a Michelson contrast of 1). In Equation 11 the contrast, $c$, is in linear units. The near miss to Weber’s law for suprathreshold contrast discrimination (Bird, Henning, & Wichmann, 2002; Nachmias & Sansbury, 1974; Swift & Smith, 1983) and the fact that cortical contrast adaptation is divisive (Ohzawa et al., 1985) suggest that the internal representation of contrast is closer to a log than linear scale, so we now express Equation 11 in terms of $u = \log_b(c)$:

$$r = r_{\max} b^{qz(s-1)} (s-1)^{1/s-1} s \frac{b^{qz}}{b^{qz} + b^{qzu}} + B,$$

(12)

where $b$ is the base of the logarithm, and $z = \log_b(c_{\text{std}})$. The first derivative of $r$ is given by

$$\frac{dr}{du} = r_{\max} b^{qz(s-1)} (s-1)^{1/s-1} s q \ln(b) \times b^{qzu} \left( \frac{b^{qz} + (1-s)b^{qzu}}{(b^{qz} + b^{qzu})^2} \right),$$

(13)

and the Fisher information for this neuron in response to log contrast, $u$, is given by $I_F(u) = (dr/du)^2 / r$, which, for a baseline, $B$, of zero, simplifies to

$$I_F(u) = Tr_{\max} b^{qz(s-1)} (s-1)^{1/s-1} s(q \ln(b))^2 \times b^{qzu} \left( \frac{(b^{qz} + (1-s)b^{qzu})^3}{(b^{qz} + b^{qzu})^5} \right).$$

(14)

Within each of the populations shown in Figures 2, 3, and 4, all the neurons are identical except for $z$, so the population Fisher information at test contrast $u$ is found by summing Equation 14 across all the $z$-values in the population. This gives the Fisher information for the supersaturating coding schemes given in the top panels of Figures 2, 3, and 4. For the saturating schemes in the lower panels, the procedure is the same, except that, for each neuron, we set the Fisher information to zero for any contrast above the peak (given by Equation 4), where the gradient of the contrast-response function is zero.

In our calculations, we used a logarithm with base $b = 10$. For simplicity, we assumed that the log semi-saturation contrasts, $z$, were uniformly distributed across the log contrast axis, from $-1.95$ to $-0.05$ in log contrast steps of 0.1. This range of semi-saturation contrasts (i.e., contrasts of around 0.01 to 1 in linear units) approximately reflects the range found physiologically (Chirimuuta et al., 2003; Clatworthy et al., 2003; Sclar et al., 1990), although monkeys do seem to have a separate population of neurons with semi-saturation contrasts clustered around a value greater than 1 (Clatworthy et al., 2003).

In Figure 5, we plot the Fisher information and population spike rate as functions of test contrast for three different levels of the supersaturation parameter ($s = 1.05, 1.1, \text{ and } 2$) and compare these with the values obtained from the corresponding saturating schemes; the right, middle, and left columns of panels in Figure 5 correspond to the population coding schemes shown in Figures 2, 3, and 4, respectively. The unrealistically strong supersaturation scheme ($s = 2$) is highly beneficial in terms of both decoding accuracy (indicated by Fisher information) and energy usage (indicated by the population spike rate). For the more realistic levels of supersaturation ($s = 1.05$ or 1.1), the improvement in contrast decoding accuracy is very slight, but the supersaturating schemes show significant savings in energy usage for moderate and high contrasts.

All the values plotted in Figure 5 assume the same uniform distribution of log semi-saturation contrasts, $z$, described above. However, our conclusions do not depend strongly on this assumption or assumptions about the level of supersaturation. Adding a supersaturating region (i.e., a region of gradient less than zero) to the contrast-response function of any saturating neuron will improve the contrast decoding accuracy for contrasts falling within that region, while reducing metabolic costs. In the special case of $s = 2$, the downward slope of the contrast-response function is an exact mirror image of the upward slope, and supersaturation improves contrast coding accuracy to the same extent as adding another saturating neuron, while consuming far less energy.

We have shown that the inclusion of a supersaturating region in the contrast-response function improves both the accuracy and energy efficiency of a population contrast code. Therefore, the question is not “Why do some neurons supersaturate?” but, rather, “Why don’t they all supersaturate strongly, like in the upper panel of Figure 2?” The answer is probably that, as noted by Chirimuuta et al. (2003, p. 1259) and Geisler and Albrecht (1995), neurons do not just code for contrast; they also code for other stimulus attributes, such as orientation and spatial frequency, and the requirements of these different roles may come into conflict. If all neurons supersaturated strongly with contrast, then, for any stimulus contrast, only a small proportion of neurons could respond strongly enough to carry an accurate code for the other stimulus attributes, leading to poor performance on estimation of those attributes. With a saturating (but not supersaturating) scheme, most neurons respond close to maximum over much of the contrast range in response to their preferred stimulus orientation, spatial frequency, etc., and this would improve the accuracy with which those other attributes were coded (Geisler & Albrecht, 1995). According to this view, the multiplicity of different super-saturation levels is the result of varying trade-offs between accurate, energy-efficient, contrast coding, and accurate coding of everything else. This conclusion is consistent with the finding that strongly supersaturating...
neurons are much more prevalent in the auditory than the visual cortex (Brugge & Merzenich, 1973; Phillips, Semple, Calford, & Kitzes, 1994). The auditory signal has lower dimensionality than the visual signal, so coding of signal level has fewer competing stimulus attributes to trade off against; we would, therefore, expect the auditory cortex to have a greater tendency toward supersaturation, which benefits only coding of signal level.

Conclusions

Peirce (2007) argued that the contrast-response functions of V1 and V2 cells are suboptimal for contrast coding, especially functions that supersaturate, and proposed that they instead facilitate conjunction detection. We take the opposite view: The contrast-response functions of V1 and V2 cells cannot facilitate conjunction detection in the way that Peirce proposes, and supersaturation actually improves both the accuracy and energy efficiency of contrast coding mechanisms. We argue that supersaturation in the visual cortex is generally weak because, although it benefits contrast coding, it impairs coding of everything else.

Appendix A

Here we explain why the component contrast is half the stimulus Michelson contrast for a plaid, and equal to the stimulus Michelson contrast for a grating. First, we need to clarify the relationship between two definitions of contrast: the contrast of a stimulus component, and the Michelson contrast of an entire stimulus. Michelson contrast, \( c_M \), of a stimulus is defined as follows:

\[
  c_M = \frac{L_{\text{max}} - L_{\text{min}}}{L_{\text{max}} + L_{\text{min}}},
\]

where \( L_{\text{max}} \) and \( L_{\text{min}} \) are the maximum and minimum stimulus luminance, respectively. Michelson contrast varies between 0 (when \( L_{\text{min}} = L_{\text{max}} \)) and 1 (when \( L_{\text{min}} = 0 \)). The general formula for a plaid is

\[
L(x, y) = L_0[1 + c_1 g_1(x, y) + c_2 g_2(x, y)].
\]

\( L(x, y) \) is the luminance at position \((x, y)\); \( L_0 \) is the background luminance; \( c_1 \) and \( c_2 \) are the contrasts of the two components of the plaid and specify the amplitude of each component as a proportion of the background luminance; \( g_1 \) and \( g_2 \) are (usually) sinusoidal functions of space, e.g., \( \sin(2\pi fx) \), for a vertical component with spatial frequency \( f \), and \( \sin(2\pi fy) \) for a horizontal component. If one component has zero contrast, e.g., \( c_2 = 0 \) and \( c_1 = c \), then Equation A2 reduces to the equation for a grating:

\[
L(x, y) = L_0[1 + c g_1(x, y)].
\]

\( g_1 \) and \( g_2 \) vary between 1 and -1, so, for the grating in Equation A3, we have \( L_{\text{max}} = L_0(1 + c) \) and \( L_{\text{min}} = L_0(1 - c) \); substituting these values into Equation A1 gives \( c_M = c \), so the Michelson contrast of the stimulus is given by \( c \) in Equation A3. For a plaid with equal-contrast components \((c_1 = c_2 = c)\), Equation A2 gives \( L_{\text{max}} = L_0(1 + 2c) \) and \( L_{\text{min}} = L_0(1 - 2c) \), giving \( c_M = 2c \), so the Michelson contrast of the plaid stimulus is twice the contrast of each component.

Appendix B

Using equal-width histogram bins on a linear contrast axis, Brady and Field (2000) and Tadmor and Tolhurst (2000) both reported that the contrast distributions in natural images peak at zero contrast and show approximately exponential decay, i.e., the probability, \( P(c) \), of contrast \( c \) is given by \( P(c) \propto \exp(-c/\lambda) \); but using equal-width bins on a log contrast axis, Clatworthy et al. (2003) found that the contrast distribution peaks close to a contrast of 0.1. This difference can be explained by the fact that switching from equal-width bins on a linear contrast axis to equal-width bins on a log contrast axis is equivalent to multiplying each value of the distribution by the contrast, \( c \), as proved below.

Assume that log units, \( u \), are given by \( u = \ln(c) \). Then, \( c = \exp(u) \). If the width of a bin in log units is \( \delta u \), and the width of the corresponding bin in linear units is \( \delta c \), then the ratio of bin widths (linear/log) as \( \delta u \to 0 \) is given by

\[
\frac{dc}{du} = \exp(u) = c.
\]

This means that, for each equal-width bin on a log axis, the bin covering the corresponding contrast range on a linear axis will have width proportional to \( c \). So, to calculate the shape of the probability distribution for equal-width bins on log axes, we need to take the probability distribution for equal-width bins on linear axes and multiply each value by the variable bin width, i.e., by \( c \), giving \( P(c) \propto c \exp(-c/\lambda) \). This peaks at \( c = \lambda \), and setting \( \lambda = 0.1 \) gives a distribution very similar to that of Clatworthy et al.
Acknowledgments

This work was supported by a grant to Li Zhaoping from the Gatsby Charitable Foundation and BBSRC Cognitive Science Foresight Grant BB/E002536/1. We thank the Eyeball Kids for a helpful discussion and Mark Georgeson, Isabelle Mareschal, and Joshua Solomon for commenting on previous drafts of this paper.

Commercial relationships: none.
Corresponding author: Keith A. May.
Email: keith@keithmay.org.
Address: Department of Computer Science, UCL, Gower Street, London WC1E 6BT, UK.

References


