

Minireview

# Receptive versus perceptive fields from the reverse-correlation viewpoint

Peter Neri \*, Dennis M. Levi

*School of Optometry, University of California, 360 Minor Hall, Berkeley, CA 94720-2020, USA*

Received 28 September 2005; received in revised form 13 January 2006

## Abstract

This brief review article brings together a series of related experiments in psychophysics and physiology that show striking similarities between measurements in human observers and in single neurons. We consider seven pairs of primary research articles, each pair consisting of one paper in physiology and one in psychophysics, and we highlight common features between receptive and perceptive fields obtained using reverse correlation. We conclude by discussing how to assess the validity of perceptive fields as predictors of human responses, and by deriving a novel expression for the maximum trial-by-trial predictability attainable by any model for any psychophysical task.

© 2006 Elsevier Ltd. All rights reserved.

*Keywords:* Noise image classification; Visual psychophysics; Behavioural receptive fields

## 1. Introduction

Over the past 10 years since publication of Ahumada's (1996) paper, a number of psychophysical studies have used reverse correlation to characterize human observers' strategies in simple visual tasks. In a few cases, corresponding physiological studies exist that have characterized the responses of single neurons to very similar stimuli, sometimes using reverse correlation and sometimes using other methods. This presents another opportunity to compare the functional properties of single neurons and of the whole organism. In this brief review, we compare seven such pairs of physiological and psychophysical studies. Our selection criteria were that (1) the psychophysical studies used reverse correlation to characterize human observers' strategies, and (2) the physiological studies characterized neural responses to stimulus properties that were relevant to the psychophysical tasks. Several of the physiological studies used reverse correlation, but not all of them did.

In the final part of this article we discuss methods to validate perceptive fields as predictors of human responses. In

particular, we derive a mathematical expression that can be easily used to provide an upper estimate for the maximum predictive power of any model in any psychophysical task. By predictive power we mean the ability to reproduce human responses at the level of individual trials, which is the most accurate level of description for any psychophysical experiment.

## 2. Reverse correlation methods

The theory and practice of reverse correlation have been described in detail elsewhere (Abbey & Eckstein, 2002; Ahumada, 2002; Marmarelis & Marmarelis, 1978; Murray, Bennett, & Sekuler, 2002), so here we will just outline the properties of this method that are relevant to a comparison of physiological and psychophysical studies.

Neurophysiological applications of reverse correlation abound in the literature and have been extensively reviewed by other authors on various occasions (e.g., Ringach & Shapley, 2004; Victor, 2005). In this approach, visual neurons are stimulated by noisy images, and the spike output is cross-correlated with the input. Given certain assumptions about the functional characteristics of neurons (Marmarelis & Marmarelis, 1978), the outcome of the cross-correlation procedure provides useful estimates of neuronal

\* Corresponding author.

E-mail address: [pn@white.stanford.edu](mailto:pn@white.stanford.edu) (P. Neri).

receptive field structure (Ringach, 2004; Sakai, Naka, & Korenberg, 1988).

Transferring this approach from physiology to psychophysics involves some difficulties, as sketched in Fig. 1. The first issue relates to the nature of the input. One can simply measure the response of a neuron to pure noise, but if a human subject views a display consisting only of noise, it is unlikely that he or she will generate coherent responses. At the very least, the experimenter needs to specify a task to be performed, and tasks typically relate to the presence or absence of a signal that is added to the noise on some trials. Useful results can be obtained using only noise (e.g., Gosselin & Schyns, 2003), but the only way to be sure that observers are actually performing the specified task is to measure  $d'$  for detecting a signal. This general feature of psychophysical design introduces the first difference between the use of reverse correlation techniques in neurons and human behaviour.

The second issue relates to the nature of the output: neurons generate spikes, whereas humans typically generate either binary responses (such as “yes/no”) or a rating response (e.g., 1, 2, 3, and 4). In both cases there are nonlinearities involved, but in humans the nature of the nonlinearity tends to quantize the output more than in single neurons (even for rating tasks). This introduces a second difference between physiological and psychophysical versions of the technique.

Ahumada first introduced a variant of reverse correlation that could be used in auditory psychophysics (Ahumada & Lovell, 1971), later termed “noise image classification” in vision science (Eckstein & Ahumada, 2002). In its simplest form, this technique involves a “yes/no” task in which noise images are shown on half the trials, and noise + signal images are shown on the remaining half. For example, the signal could be a bright bar, which the observer is required to detect. The noise is then averaged separately for the four different stimulus–response classes (hits, misses, false alarms, and correct rejections). These averages are combined by summing the noise averages for “yes” classes (hits and false alarms) and subtracting those for “no” classes (misses and correct rejections) to yield the psychophysical equivalent of the

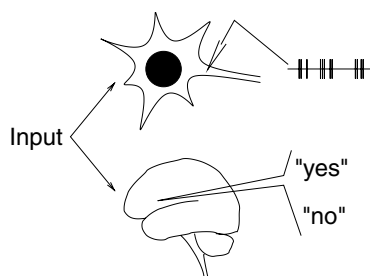


Fig. 1. Physiology versus psychophysics, neurons versus brains. The transition from one to the other requires some modifications to the reverse correlation procedure (see text), mainly due to the fact that both inputs and outputs differ between the two systems.

receptive field, which has been termed the perceptive field (Jung & Spillman, 1970). This simple analysis (as well as more sophisticated variants, e.g., Murray et al., 2002) is tailored to an experimental paradigm where a signal is added to the input, and where the output is a psychophysical response, thus allowing the transition from neurons to behaviour illustrated in Fig. 1.

There are potential difficulties with the interpretation of perceptive fields obtained using noise image classification. Signal-present averages can differ from signal-absent averages if the system is highly nonlinear (Abbey & Eckstein, 2002; Ahumada & Lovell, 1971). These potential differences arise from a combination of nonlinear kernels, the presence of the signal and the nature of the transduction that generates the psychophysical response (Neri, 2004a). In general, any inference from derived perceptive field to underlying physiological processing should be substantiated with appropriate modeling.

### 3. Receptive versus perceptive fields

In this section, we review seven pairs of physiological and psychophysical studies that have characterized the responses of single neurons and whole organisms under similar experimental conditions.

#### 3.1. Vernier and orientation discrimination

Simple cells are often approximated by a linear Gabor filter followed by rectification (Ringach, 2004). Such a system can be easily mapped using reverse correlation (Ringach & Shapley, 2004). Fig. 2A shows the spatial receptive field of a simple cell tuned to oblique orientations, as obtained using a white-noise mapping procedure. This is a representative example, but simple cells can differ widely in their Gabor characteristics (standard deviation of envelope, symmetry, carrier frequency) so as to encompass the

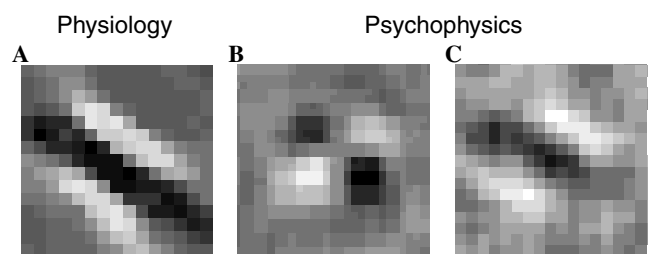


Fig. 2. Receptive (A) versus perceptive (B and C) fields for orientation discrimination. (A) is from Ringach (2002) and shows the receptive field of a simple cell in macaque V1 mapped using reverse correlation. (B) is from Ahumada (1996). This image shows one of the earliest perceptive fields derived in human vision using psychophysical reverse correlation. Observers had to detect a Vernier offset between two horizontal bars which, when offset with respect to each other, were roughly located at the two positive (bright) regions within the image. (C) is from Solomon (2002) and shows a perceptive field for an orientation discrimination task. Notice the similarity between A and C.

large range of visible detail and orientations that are available to perception.

The early work using psychophysical reverse correlation was directly related to simple cell physiology. Fig. 2B shows one of the earliest classification images in vision, obtained by Ahumada (1996) using a Vernier task that required subjects to detect a small vertical offset between two horizontal bars. The comparison between panels A and B is interesting, because it may appear that the perceptive field in B is not consistent with the physiological receptive field in A. However, a direct link can be established by assuming that subjects were relying on the output of two simple-cell-like mechanisms to perform the Vernier task: one mechanism would be used to detect one bar, the other mechanism to detect the other bar. The receptive fields of the two recruited mechanisms would be out of phase to signal the potential presence of a Vernier offset between the two bars. This scheme is probably overly simplistic, but it successfully predicts that perceptive fields should reflect the outputs of two simple-cell-like units, as shown in Fig. 2B (see also Waugh, Levi, & Carney, 1993).

This example serves to make the important point that perceptive fields inevitably reflect the summed properties of all physiological mechanisms used by observers to perform a given task. Depending on task design and demands, they may resemble one or more physiological receptive fields. If, for example, the task is changed from Vernier offset to orientation discrimination, the perceptive field returned by psychophysical reverse correlation will resemble much more closely the receptive field of an individual simple cell, as shown in Fig. 2C for an experiment that was designed by Solomon (2002) to target only one simple-cell-like mechanism.

### 3.2. Stereoscopic surface detection

Binocular units in primary visual cortex (V1) are often selective for retinal absolute disparity (Barlow, Blakemore, & Pettigrew, 1967; Cumming & Parker, 1999; Nikara, Bishop, & Pettigrew, 1968). This selectivity has been modeled using an energy correlator (Ohzawa, DeAngelis, & Freeman, 1990) that shares similarities with analogous motion detectors (Adelson & Bergen, 1985). Although this model does not explain every aspect of disparity processing in V1 (Cumming, 2002), it accounts for many features of disparity-selective responses in V1 neurons (Freeman, 2004; Ohzawa, 1998). An important prediction of this model is that anti-correlated stimuli (i.e., stimuli with the opposite contrast polarity in the two eyes) should invert the disparity tuning function, so that a disparity value leading to an increase in firing rate when stimuli are correlated becomes associated with a decrease when they are anti-correlated (Ohzawa et al., 1990). This prediction has been experimentally verified in V1 (Cumming & Parker, 1997), MT (Krug, Cumming, & Parker, 2004) and MST (Takemura, Unoue, Kawano, Quaia, & Miles, 2001). The observed inversion is accompanied by a reduction in amplitude (Cumming &

Parker, 1997), a feature not predicted by the correlator model (Ohzawa, 1998). However, simple modifications of this model allow for amplitude reduction (Read, Parker, & Cumming, 2002).

Anti-correlated stimuli have been studied psychophysically by a number of investigators (Cogan, Lomakin, & Rossi, 1993; Cogan, Kontsevich, Lomakin, Halpern, & Blake, 1995; Cumming, Shapiro, & Parker, 1998; Julesz, 1971). The general result is that, when the density of image elements is low, anti-correlated disparity signals are perceived much like correlated disparity signals, but when the density is high, they fail to evoke a stereoscopic sense of depth and lead to luster (e.g., a fully anti-correlated random-dot stereogram of sufficiently high density is not perceived as varying in depth). Both outcomes are difficult to reconcile with the inverted response of V1 neurons (Cumming & Parker, 1997). This fact, together with the lack of selectivity for relative disparity in V1 (Cumming & Parker, 1999), has led to the proposition that the processing of stereoscopic depth is finalized in extrastriate cortex (Parker, 2004).

However, a study using noise image classification in the stereoscopic domain (Neri, Parker, & Blakemore, 1999) was able to expose a processing stage in human vision that mimics the correlator-like response observed in V1 (Cumming & Parker, 1997). This result differs from previous psychophysical studies (see paragraph above), but is in agreement with the electrophysiological evidence. The similarity between neuronal (A) and behavioural (B) disparity tuning functions is shown in Fig. 3, for both correlated (solid symbols) and anti-correlated (open) signals.

The parallel illustrated in Fig. 3 should not be interpreted to mean that the neural representation of the final stereoscopic percept must reside in V1. While the earliest stage of binocular combination can be exposed psychophysically in conditions in which it affects behavioural performance (Neri et al., 1999; Parker, 2004), this does not mean that the final percept actually resides at that stage.

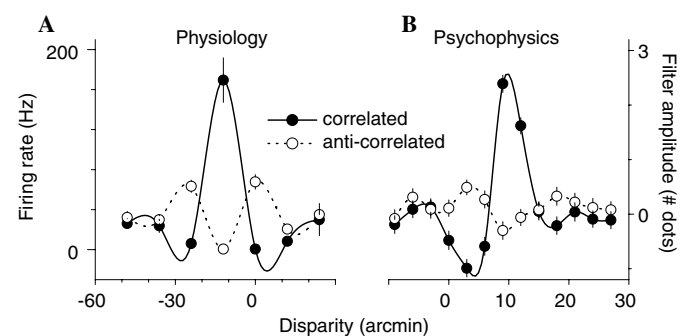


Fig. 3. Receptive (A) versus perceptive (B) fields for disparity tuning. (A) is from Cumming and Parker (1997) and shows data from a single neuron in macaque V1. (B) is from Neri et al. (1999) and shows data for a naïve observer. In both cases, tuning functions follow a Mexican-hat shape in response to correlated (same contrast in both eyes) signals (solid symbols, solid line), and invert their tuning in response to anti-correlated signals (open symbols, dotted line).

Rather, it suggests that idiosyncratic response properties of early cortical neurons, such as their response to anti-correlated stimuli, are sometimes propagated through the remainder of visual processing and can be clearly observed in human behaviour. A similar situation occurs in colour vision. Psychophysical experiments (excellently reviewed and described in Wandell, 1995, chap. 4) have shown that behavioural detection of colour stimuli matches very closely the response functions of single cones. This is very good evidence that observers, at threshold and in certain tasks, are limited in their performance by the properties of their cones, but cannot lead to the conclusion that all psychophysical properties of colour are determined by the retina.

### 3.3. Spatiotemporal dynamics of feature detection

Spike-triggered reverse correlation has been used to derive spatiotemporal receptive fields for both simple and complex cells in V1 (for references and historical overview, see Ringach & Shapley, 2004). These studies confirmed previous characterizations of the two neuronal classes, but provided a more complete picture of how striate neurons respond to simple image features in space and time (DeAngelis, Ohzawa, & Freeman, 1995). Fig. 4 shows two examples, one from a simple (A) and one from a complex cell (C). The receptive field for the complex cell was derived

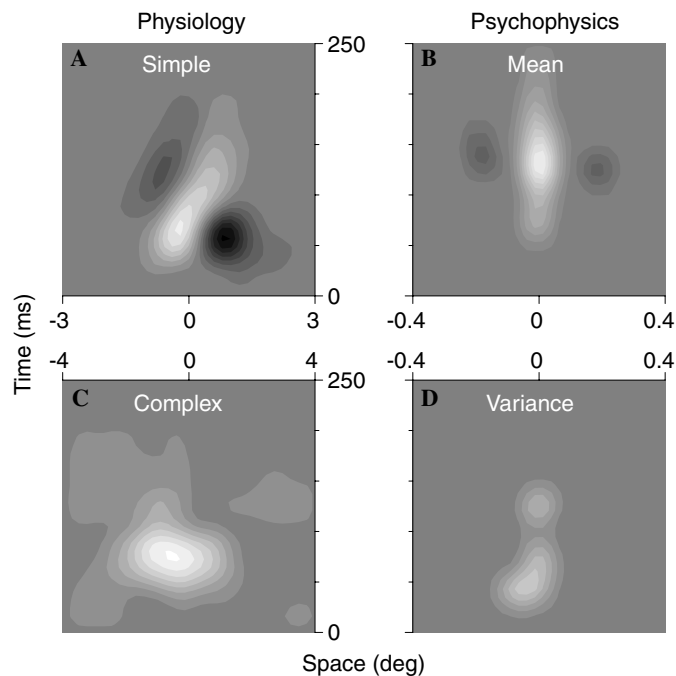


Fig. 4. Spatiotemporal receptive (A and C) versus perceptive (B and D) fields in response to luminance-defined features. (A and C) are from DeAngelis et al. (1995) and show receptive fields for a simple (A) and a complex (C) cell. (B and D) are from Neri and Heeger (2002) and show mean (B) and variance (D) kernels for a naïve observer. In the surface plots, bright is positive and dark is negative. Notice that the temporal axis (ordinate) is plotted on the same scale. Spatial scale is different due to differences in eccentricity across experiments.

from responses to bright features, but the same field is obtained for dark features.

A very similar approach can be used to study human detection of simple image features, such as bars/lines. By using noise that varies in both space and time, it is possible to derive spatiotemporal perceptive fields for detecting such features (Neri & Heeger, 2002). Two examples are shown in Fig. 4 for detecting a bright bar. The perceptive field in B is akin to a spike-triggered average, triggered by observers reporting the presence of a bright bar at the central space–time location. The perceptive field in D was obtained by analyzing how variance in the luminance signal affected responses from the observer. Because variance is polarity-independent and is related to contrast energy (Neri, 2004a; Simoncelli, 2003), this perceptive field is akin to the receptive field that DeAngelis et al. (1995) obtained for complex cells (reported in Fig. 4C).

There are clear similarities between the spatiotemporal receptive fields on the left side of Fig. 4 and the corresponding perceptive fields on the right side. Further support for this analogy comes from simulations that incorporate current models of simple and complex cells (Neri & Heeger, 2002). Notice that neuronal and behavioural data are plotted on the same temporal scale. The spatial scale (abscissa) cannot be directly compared because it depends on eccentricity, and different eccentricities were used in the electrophysiological and psychophysical experiments.

The parallel proposed in Fig. 4 should not be interpreted to mean that the perceptual representation of the bright bar was directly generated and supported by simple and complex cells in V1. Fig. 4 is intended to emphasize the fact that human processing of simple image attributes involves computations that are very similar to those performed by simple and complex cells. Whether simple and/or complex cells are directly responsible for the observed psychophysical properties is a different question, and one which would have to be addressed by invasive methods (Parker & Newsome, 1998). Nevertheless, it is interesting to notice that similar computations characterize behaviour and physiology. It seems unlikely that these similarities would be purely coincidental.

### 3.4. Orientation processing

Exploiting the well-known orientation selectivity of striate neurons, Ringach, Hawken, and Shapley (1997a) used a variant of reverse correlation involving fast sequential presentation of a randomized set of oriented images (Ringach, Sapiro, & Shapley, 1997b). Spike-triggered analysis allowed them to study various aspects of orientation processing in single neurons of macaque V1. Fig. 5A shows a representative tuning function for one neuron, preferentially responsive to vertical patterns. The central peak is flanked by inhibitory regions, resembling a Mexican hat. This shape is commonly observed in V1 neurons.

Ringach (1998) used an almost identical technique to study orientation processing in human observers. Stimuli

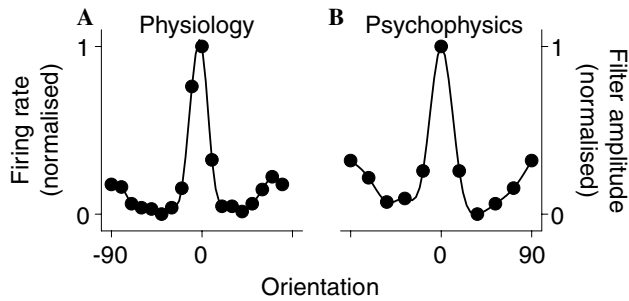


Fig. 5. Receptive (A) versus perceptive (B) fields for orientation tuning. (A) is from Ringach et al. (1997a) and shows data from a single neuron in macaque V1. (B) is from Ringach (1998) and shows data for a human observer. In both cases, tuning functions follow a Mexican-hat shape and display very similar bandpass characteristics.

were the same as those used with single neurons, and button presses were substituted in place of spikes. By reverse-correlating psychophysical responses with the input stream of oriented noise, Ringach derived perceptive fields such as that shown in Fig. 5B. The similarity between receptive and perceptive fields is striking (compare panels A and B in Fig. 5). It is not entirely surprising that human orientation channels may conform to a Mexican-hat shape, as this could be partly inferred from previous studies (e.g., Carpenter & Blakemore, 1973; see Ringach, 1998 for a more detailed discussion of how his results relate to previous literature). However, this study provides the cleanest characterization so far and, as discussed by the author, it appears to be tapping into early visual mechanisms.

### 3.5. Attentional effects on sensory tuning

The effect of spatial attention on sensory tuning in single neurons has been addressed by various groups over the past 10 years (Treue, 2001). Although there is still controversy, the consistency across most of these studies is remarkable given that they often looked at different visual areas and different visual attributes (Treue & Martínez-Trujillo, 1999; McAdams & Maunsell, 1999). The electrophysiological evidence collected so far indicates that attention (a) boosts overall response by a multiplicative gain enhancement of the neuronal tuning function and (b) does not substantially modify tuning bandwidth (Treue, 2001). An example is shown in Fig. 6A for orientation-selective responses in V4 neurons (McAdams & Maunsell, 1999): the tuning function is expanded vertically, but does not change its shape. Very similar results are obtained in V1 (McAdams & Maunsell, 1999). The same result is obtained when measuring directional selectivity in MT neurons (Cook & Maunsell, 2004; Treue & Martínez-Trujillo, 1999), as shown in Fig. 6C.

The picture is less clear in the psychophysical literature (Vergheze, 2001). Some studies have shown no effect of attention on sensory tuning (Eckstein, Shimozaki, & Abbey, 2002; Lu & Doshier, 1998; Murray, Sekuler, & Bennett, 2003; Neri, 2004b; Talgar, Pelli, & Carrasco, 2004),

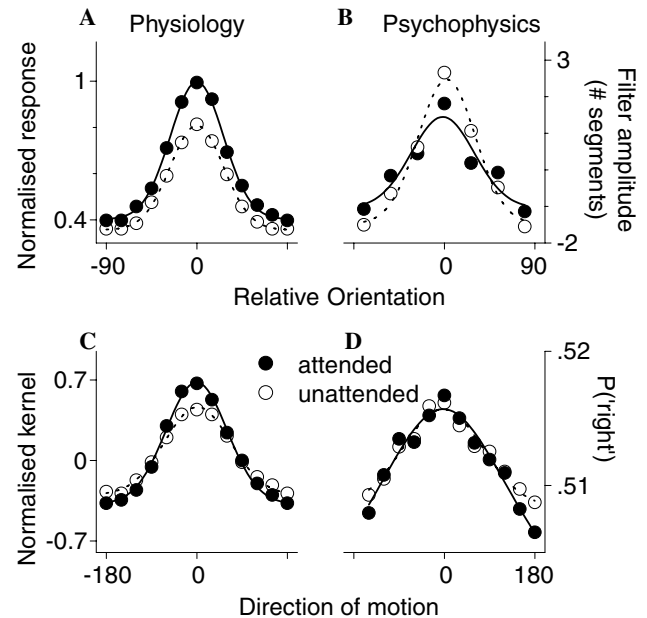


Fig. 6. Effect of attention on receptive (A and C) and perceptive (B and D) fields for orientation tuning (top) and directional tuning (bottom). (A) is from McAdams and Maunsell (1999) and shows population data from macaque V4. (B) is from Neri (2004b) and shows data for a human observer. (C) is from Cook and Maunsell (2004) and shows population data from macaque MT. (D) is from Murray et al. (2003) and shows data for a human observer. In all cases, tuning is unaffected by attentional deployment (compare solid symbols (attended condition) with open symbols (unattended condition)).

but others report sharpening (Carrasco, Williams, & Yeshurun, 2002; Doshier & Lu, 2000; Lee, Itti, Koch, & Braun, 1999; Yeshurun & Carrasco, 1998). However, if one restricts the literature survey to studies that used noise image classification (Eckstein et al., 2002; Murray et al., 2003; Neri, 2004b), the consistency is remarkable and in line with the physiology: attention has no effect on the shape of the perceptive field. An example is shown in Fig. 6B for orientation tuning (Neri, 2004b), and in Fig. 6D for directional tuning (Murray et al., 2003). These psychophysical results nicely parallel the corresponding electrophysiological findings (Figs. 6A and C). Notice also the similarity in tuning bandwidth for orientation (Figs. 6A and B).

The multiplicative gain enhancement observed in the physiology is not apparent in the psychophysical data. The main reason for this difference lies in the fact that multiplicative scaling of front-end filtering (such as neuronal gain changes) does not trivially translate to corresponding multiplicative effects for the derived perceptive fields. In general, psychophysical reverse correlation can recover the shape of front-end filtering up to a multiplicative scaling factor, and the value of this factor depends on more than just the underlying gain (Ahumada, 2002). For example, switching from 2 alternative forced choice (2AFC) to  $n$ AFC with  $n$  greater than 2 is expected to cause a positive multiplicative scaling of the derived perceptive field, even though no change in processing has occurred (Neri,

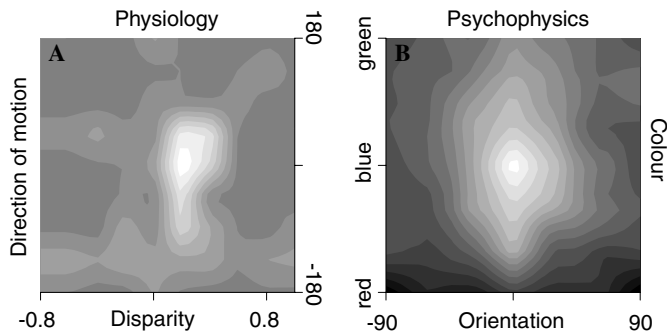


Fig. 7. Are different visual attributes multiplied or summed? (A) is adapted from Grunewald and Skoumbourdis (2004) and shows spike response (intensity of surface plot) for a V1 neuron selective for both disparity ( $x$  axis) and direction of motion ( $y$  axis). (B) (average of four observers) is from Neri (2004b) and shows a joint orientation-colour perceptible field for a conjunction task. Both surfaces provide evidence for multiplicative combination of the two features.

2004b). For this reason, it would be inappropriate to establish a direct link between gain changes (or lack thereof) in perceptible fields to underlying gain changes in front-end filtering.

The apparent discrepancy between psychophysical studies using noise image classification and psychophysical studies using other techniques does not mean that either set of results is incorrect. It simply indicates that different methods may target different processing stages. From the evidence that is available so far, noise image classification appears to target stages reflecting computations that are very similar to those observed in physiological recordings from single neurons in known visual areas of cortex.

### 3.6. Feature conjunction

In a recent study, Grunewald and Skoumbourdis (2004) mapped out selectivity for both direction of motion and retinal disparity in individual V1 neurons. From these measurements, they concluded that these two features are combined in a multiplicative fashion, rather than being summed. An example of their recordings is shown in Fig. 7A.

A very similar experiment was performed by Neri (2004b). Psychophysical reverse correlation was used to derive selectivity for both orientation and colour in an orientation-colour conjunction task. Similar to what was found in the physiology, the surface in Fig. 7B is consistent with a multiplicative operation for conjoining the two features. Although axes are clearly not directly comparable in these plots, the very sharp peak in the centre of both plots is a signature of multiplication, both in single neurons and in human observers.

## 4. Validation

The validity and significance of derived perceptible fields can be assessed by testing whether they can be used to suc-

cessfully predict human behaviour. This question has been recently addressed by Murray, Bennett, and Sekuler (2005), who used classification images to predict the absolute efficiency of human observers performing three different tasks, on the hypothesis that they are well modelled as linear classifiers. As shown in Fig. 8A, predicted efficiency slightly underestimates observed efficiency, but in general there is very good agreement.

A different approach involves simulating observers' responses on a trial-by-trial basis. There is an upper limit to how successfully this can be done, and this limit is set by the observer's internal noise. An indirect measure of internal noise can be obtained by measuring the consistency of observers' responses to repeated presentations of the same stimulus, i.e., using the double-pass method (Burgess & Colborne, 1988; Li, Klein, & Levi, 2006). Human-human consistency is simply the fraction of trials for which the observer's responses are the same on two presentations of the same stimulus. Human-human consistency (plotted on the  $x$  axis in Fig. 8B for yes-no detection of a bright bar embedded in Gaussian noise bars) can then be used to establish upper and lower boundaries on the accuracy with which any model could predict observers' trial-by-trial responses (see Appendix A). The region within these boundaries is shown in grey in Fig. 8B. Model-human consistency is plotted on the  $y$  axis. Solid points refer to a linear classifier model (Murray et al., 2005) that uses the

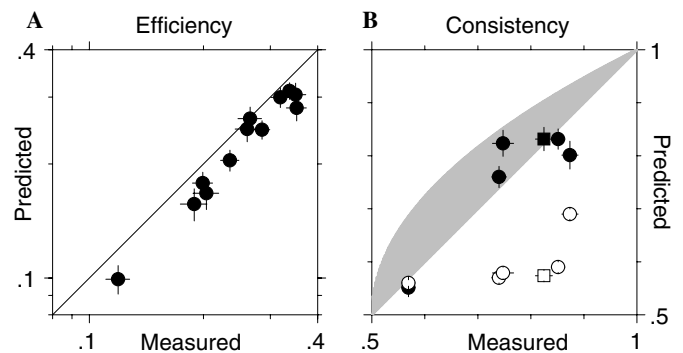


Fig. 8. How well do perceptible fields predict human performance? (A and B) show two different ways of answering this question. Once the perceptible field has been derived, this is incorporated into a model linear classifier, and the linear classifier is used to predict human behaviour by simply cross-correlating the perceptible field with the visual stimuli. (A) is Fig. 3 from Murray et al. (2005) and shows the absolute efficiency one would predict for each observer by comparing their classification image to the ideal template, plotted against the measured absolute efficiency. (B) plots consistency between model response and human response on the  $y$  axis, versus consistency for human responses to the same set of stimuli presented twice. Consistency is simply the fraction of time that responses are the same. The region in grey shows upper and lower boundaries on model-human consistency for the best possible model (see Appendix A). Solid points show model-human consistency for a linear classifier. Different points correspond to different subjects, with the exception of the square symbol which is for the same subject as the nearby circle, but obtained using uniform as opposed to Gaussian noise. For comparison, open symbols show consistency for a 'dumb' model that responds the same on every trial (either 'yes' or 'no,' the one that returns the largest consistency).

experimentally derived perceptive fields as front-end filters. Most of these points are within or very close to the grey region, indicating that perceptive fields provide an almost optimal description of human processing for the task and stimuli used in these experiments. As a baseline for comparison, we also plot model–human consistency for a model that responds the same on every trial (either ‘yes’ or ‘no’, the one that returns the largest consistency) in open symbols. This is to show that the high consistency achieved by the linear classifier model (solid symbols) is not trivially due to observers’ response bias (with the exception of one observer at bottom left).

The diagram in Fig. 9 summarizes the way in which this validation technique should be integrated within a noise image classification experiment. Four steps are involved: (1) run a noise image classification experiment where the observer is presented with a sequence of stimulus noise samples, and each noise sample is associated with a response (“yes” or “no” in the example); (2) derive a perceptive field by coupling noise samples and response sequence; (3) filter the noise samples with the perceptive

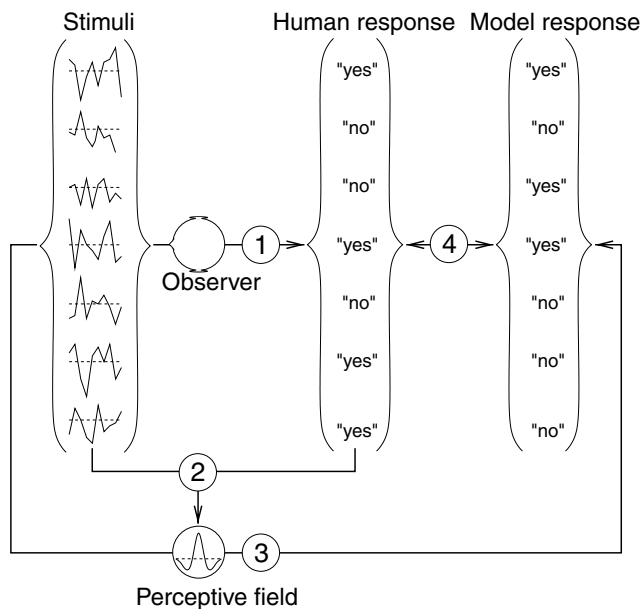


Fig. 9. A possible recipe for parameter-free characterization of human responses on a trial-by-trial basis. (step 1) A noise image classification experiment is carried out, in which a series of stimulus noise samples (left) is presented to a human observer who generates a series of corresponding psychophysical responses (‘yes’ or ‘no’ in this example). (step 2) The perceptive field is derived by reverse-correlating the response sequence with the stimulus noise samples. This step has 0 free parameters. (step 3) The perceptive field is cross-correlated with each noise samples, and the output of the cross-correlation is used to generate a response like that given by observers. If the original experiment is designed using an alternative forced choice paradigm, this step has 0 free parameters. (step 4) The percentage of trials on which the perceptive field returns the same answer as the observer (model–human consistency) is evaluated against the percentage of trials on which the observer gives the same answer to the same stimulus presented twice (human–human consistency) using the expression provided in Appendix A. The entire procedure involves 0 free parameters.

field to generate a sequence of responses from the linear observer model; (4) compare these responses with those given by the observer to the same noise samples, and use the expression in Appendix A to establish how successfully the perceptive field can predict human responses. Three further points should be made.

First, if possible the experiment should adopt a 2AFC design in order to exclude any role for response bias. Bias is not only undesirable in most psychophysical experiments, but it also introduces the need for an extra free parameter at step 3 when the output of the filtering operation between the derived perceptive field and the noise sample must be converted into a psychophysical response. In a one-interval experiment this output must be thresholded and the threshold (which is subject to bias) is an extra parameter. In modeling a 2AFC experiment there is no need for this extra parameter, it is sufficient to pick the interval associated with the largest output from the filtering operation.

Second, a subset of the blocks at step number 1 should be of double-pass type, otherwise step 4 cannot be carried out because the expression derived in Appendix A makes use of knowledge that can only be obtained from a double-pass procedure.

Third, the noise samples used to derive the perceptive field at step number 2 should not be the same noise samples that are filtered by the perceptive field at step number 3. The predictive power of the perceptive field (as implemented by the linear observer model) must be assessed by predicting responses to novel noise samples, not to the same ones that were used to derive the perceptive field in the first place. Finally, it should be emphasized that step number 3 can in principle be implemented by models other than the linear observer model, because the expression in Appendix A applies to any model.

## 5. Conclusions

This brief review (making 10 years after publication of Ahumada’s, 1996 paper) suggests that psychophysical reverse correlation can be used to retrieve perceptive fields that are often remarkably similar to analogous receptive fields in single neurons. Furthermore, perceptive fields can be used to construct simple models that replicate human behaviour to the highest degree of approximation (i.e., trial by trial). The novel expression presented in the Appendix A to this paper can be used to quantify the validity and significance of the derived perceptive field for any psychophysical task and for any possible model.

It is not always the case that perceptive fields are trivially related to the physiological receptive fields that are believed to underlie the psychophysical task performed by human observers. We presented one example (Fig. 2B) from a classical experiment in the literature (Ahumada, 1996) where the derived perceptive field can only be linked to the most probable physiological substrate by means of a model of how observers may use multiple neural

mechanisms to generate a behavioural response. This example makes the important point that psychophysical reverse correlation necessarily returns the aggregate structure of the relevant neural processing, because it only uses information from the final decisional stage. This final stage reflects the properties of all physiological stages that were involved in reaching the final decision. Depending on task design and demands, the number of different stage types involved may vary significantly. Consequently, the derived perceptive field may or may not be easily linked to the underlying neural substrates. For the other examples presented in this paper the link was immediate.

The similarity between receptive and perceptive fields highlighted here should not be taken to imply that behavioural performance in these experiments was completely determined by the response properties of individual neurons in early cortex. Perception is the outcome of complex networks involving different cortical areas and neuronal populations, even when (as is typically the case in psychophysics) observers are pushed to rely on only a subset of this circuitry. However, depending on the specifics of the question being studied and the design of the experiment, different processing stages may be emphasized as accessible bottlenecks. It appears that psychophysical reverse correlation can be particularly useful in exposing similar stages to those currently believed to be relevant for the neural implementation of perceptual processing.

This review is intended as “food for thought,” rather than as a demonstration that psychophysical reverse correlation is especially suited to uncover the physiological mechanisms underlying perceptual processing. In fact, such a statement would be premature at this stage. One advantage of using psychophysical reverse correlation, however, is that it allows trial-by-trial characterization of human behaviour with no free parameters. As long as the experiment adopts an AFC design and the linear observer model is a viable detector for the problem at hand, the analysis outlined in Fig. 9 can be completed with no free parameters. This is a rare situation in a typical psychophysical investigation. Future research will be necessary to establish whether this approach can be successfully applied to a wide range of problems in vision.

## Acknowledgments

We are indebted to Richard Murray for his thoughtful discussions and detailed contributions to earlier drafts of the paper. The work was supported by Grant RO1EY01728 from the National Eye Institute, NIH, Bethesda, MD.

## Appendix A

### A.1. Upper and lower boundaries for the optimal model

We define the optimal model for the human observer as the one that selects, out of  $n$  intervals on each trial  $i$ , the

one associated with the highest probability that it would be selected by the human observer. This section shows that, for such an optimal model, model–human consistency  $C_{mh}$  is bounded as  $C_{hh} \leq C_{mh} \leq M(C_{hh})$ , where

$$M(x) = \frac{1 + \sqrt{1 + n(x(n-1) - 1)}}{n},$$

$n$  is the number of possible response types ( $n$ AFC; for yes/no,  $n = 2$ ) and  $C_{hh}$  is human–human consistency.  $C_{mh}$  can take any value within this interval depending on the choice of stimuli used to evaluate  $C_{hh}$  and  $C_{mh}$ . There is a choice of stimuli such that  $C_{mh} = C_{hh}$ , i.e., human–human consistency is an upper limit for any model. Similarly, there is a choice of stimuli such that  $C_{mh} = M(C_{hh}) \geq C_{hh}$ . In other words, depending on the choice of stimuli, the upper limit for modeling can be as low as  $C_{hh}$ , or as high as  $M(C_{hh})$ , but cannot be outside these two values.

The upper boundary defined by  $M(x)$  shares similarities with the predictable variance computed by Eq. (5) in Ahumada and Lovell (1971) (which corresponds to the Spearman–Brown formula). These authors used the split-half method to compute an upper limit for modeling of a rating scale response task. They assumed that each response was the outcome of a predictable part and a random part, leading to each trial having the same probability of differing from the predictable part. This corresponds to the condition (equality of all  $p$ 's) that is used to derive the upper boundary for  $C_{mh}$  (see below). Ahumada and Lovell did not define a lower boundary.

There are important differences between the equation derived by Ahumada and Lovell (1971) and the one derived here for the upper boundary on  $C_{mh}$ . For two passes of the same stimuli, their equation derives the predictable variance for the response rates obtained by averaging the two passes. This means that the model fit has to be assessed using this average, which in turn means that it is always necessary to collect two passes in order to test the model against the predictable variance. The upper (and lower) boundary in the present paper is computed using two passes, but it makes a prediction for a single pass. This means that two passes are only required to estimate upper and lower boundaries for the model. The model can then be tested on stimuli that are passed once (provided they share the same statistical structure as those used in the double-pass). Another difference has to do with the applicability of their analysis, which focused on a rating task. The analysis presented here extends to virtually any task.

### A.2. Basic relations

We will use the following two relations:

(1) For any set of  $m$  positive (or = 0) real numbers  $x_1, \dots, x_m$ ,

$$mE(x)^2 \geq E(x^2) \geq E(x)^2, \quad (1)$$

this relation is very simple to demonstrate geometrically. The first inequality states that the quadratic mean is max-

imized by having one of the  $x$ 's equal to the sum of all  $x$ 's (this is the maximum value that any  $x$  can take), and all others equal to 0. The second inequality is the well-known quadratic-arithmetic inequality (quadratic mean  $\geq$  arithmetic mean); it becomes equality when all  $x$ 's are equal.

(2) If the  $x$ 's are constrained to lie within a certain interval  $k_1 \geq x_i \geq k_2$ , the quadratic mean (for a given arithmetic mean) is maximized by a set of real numbers  $y_1, \dots, y_m$  where a fraction  $a$  of these numbers is equal to  $k_1$ , and the remaining fraction  $(1 - a)$  is equal to  $k_2$  (i.e., the distribution is skewed to the extremes), with  $a = (E(x) - k_2)/(k_1 - k_2)$ . We have

$$\begin{aligned} E(x^2) &\leq E(y^2) = a \cdot k_1^2 + (1 - a) \cdot k_2^2 \\ &= E(x)(k_1 + k_2) - k_1 k_2. \end{aligned} \quad (2)$$

When  $k_1 = 0$  and  $k_2 = mE(x)$ , this is equivalent to the first inequality in the previous relation.

### A.3. General framework

Consider a set of arbitrary stimuli, presented in a trial sequence. If the task is yes/no, only one stimulus per trial is presented. If the task is  $n$ AFC,  $n$  stimuli are presented. The optimal model, being noiseless, associates each trial with a fixed, invariable response out of  $n$  possible responses ( $n = 2$  for yes/no). The human observer associates each trial  $i$  with a probability  $p_i$  that the observer's response will be the same as the optimal model's response on that trial, and probabilities  $p_{i,j}$  ( $j = 1$  to  $n - 1$ ) that the response will be one of the remaining  $n$  minus 1 response types. We can now write

$$C_{mh} = E_i(p_i), \quad (3)$$

$$C_{hh} = E_i(p_i^2 + (n - 1) \cdot E_j(p_{i,j}^2)), \quad (4)$$

where  $E_i$  is expectation across trials and  $E_j$  is expectation across the  $n - 1$  response types that were not selected by the noiseless observer on trial  $i$ .

We can say the following about this set of probabilities. All  $p$ 's are  $\geq 0$ . For  $n$ AFC,  $1/n \leq p_i \leq 1$ , i.e., the probability that the human observer gives the same response as the optimal model by definition cannot be lower than chance. Similarly,  $0 \leq p_{i,j} \leq k_2$ , where  $k_2$  is the smallest between  $p_i$  and  $1 - p_i$ . This is because the probability that the human observer gives a different response than the optimal model cannot be by definition greater than the probability that it gives the same response  $p_i$ , and of course it cannot be greater than  $1 - p_i$ .

### A.4. Upper limit on $C_{mh}$

From the second inequality in (1),

$$E_j(p_{i,j}^2) \geq E_j(p_{i,j})^2 = ((1 - p_i)/(n - 1))^2$$

and substituting into (4), we have that the lower limit  $C_{hh}^-$  for  $C_{hh}$  is

$$C_{hh}^- = E_i(np_i^2 - 2p_i + 1)/(n - 1).$$

Again from the second inequality in (1),  $E_i(p_i^2) \geq E_i(p_i)^2$  and so

$$C_{hh}^- = (nC_{mh}^2 - 2C_{mh} + 1)/(n - 1).$$

The solution to this quadratic equation is

$$C_{mh}^+ = (1 + \sqrt{(1 - n + n(n - 1)C_{hh}^-)})/n,$$

where  $C_{mh}^+$  is an upper limit for  $C_{mh}$ , because we used  $C_{hh}$  instead of  $C_{hh}^-$ .

### A.5. Lower limit on $C_{mh}$

For trials on which  $p_i \geq 0.5$ ,  $0 \leq p_{i,j} \leq (1 - p_i)$ . From (2) (or the first inequality in (1)) we have

$$E_j(p_{i,j}^2) \leq E_j(p_{i,j})(1 - p_i) = (1 - p_i)^2/(n - 1).$$

Substituting this into (4),

$$C_{hh}^+ = E_i(2p_i^2 + 1 - 2p_i). \quad (5)$$

Because  $0.5 \leq p_i \leq 1$ , from (2) we have  $E_i(p_i^2) \leq C_{mh}3/2 - 1/2$ .

Substituting this (and (3)) into (5),

$$C_{hh}^+ = C_{mh}. \quad (6)$$

For trials on which  $p_i < 0.5$ ,  $0 \leq p_{i,j} \leq p_i$ . From (2) we have  $E_j(p_{i,j}^2) \leq E_j(p_{i,j})p_i = (1 - p_i)p_i/(n - 1)$ . Substituting into (4),

$$C_{hh}^+ = E_i(p_i^2 + (1 - p_i)p_i) = C_{mh}. \quad (7)$$

(6) and (7) lead to  $C_{mh}^- = C_{hh}$  across all trials.

## References

- Abbey, C. K., & Eckstein, M. P. (2002). Classification image analysis: Estimation and statistical inference for two-alternative forced-choice experiments. *Journal of Vision*, 2, 66–78.
- Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2, 284–299.
- Ahumada, A. J., & Lovell, J. (1971). Stimulus features in signal detection. *Journal of the Acoustical Society of America*, 49, 1751–1756.
- Ahumada, A. J. (1996). Perceptual classification images from Vernier acuity masked by noise. *Perception*, 26, 18.
- Ahumada, A. J. (2002). Classification image weights and internal noise level estimation. *Journal of Vision*, 2, 121–131.
- Barlow, H. B., Blakemore, C., & Pettigrew, J. D. (1967). The neural mechanisms of binocular depth discrimination. *Journal of Physiology*, 193, 327–342.
- Burgess, A. E., & Colborne, B. (1988). Visual signal detection IV. Observer inconsistency. *Journal of the Optical Society of America A*, 5(4), 617–627.
- Carpenter, R. H. S., & Blakemore, C. (1973). Interactions between orientations in human vision. *Brain Research*, 18, 287–303.
- Carrasco, M., Williams, P. E., & Yeshurun, Y. (2002). Covert attention increases spatial resolution with or without masks: Support for signal enhancement. *Journal of Vision*, 2(6), 467–479.
- Cogan, A. I., Lomakin, A. J., & Rossi, A. F. (1993). Depth in anti-correlated stereograms: Effects of spatial density and interocular delay. *Vision Research*, 33, 1959–1975.

- Cogan, A. I., Kontsevich, L. L., Lomakin, A. J., Halpern, D. L., & Blake, R. (1995). Binocular disparity processing with opposite-contrast stimuli. *Perception*, *24*, 33–47.
- Cook, E. P., & Maunsell, J. H. R. (2004). Attentional modulation of motion integration of individual neurons in the middle temporal visual area. *Journal of Neuroscience*, *24*, 7964–7977.
- Cumming, B. G. (2002). An unexpected specialization for horizontal disparity in primate primary visual cortex. *Nature*, *418*, 633–636.
- Cumming, B. G., & Parker, A. J. (1997). Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature*, *389*, 280–283.
- Cumming, B. G., Shapiro, S. E., & Parker, A. J. (1998). Disparity detection in anti-correlated stereograms. *Perception*, *27*, 1367–1377.
- Cumming, B. G., & Parker, A. J. (1999). Binocular neurons in V1 of awake monkeys are selective for absolute, not relative, disparity. *Journal of Neuroscience*, *19*, 5602–5618.
- DeAngelis, G. C., Ohzawa, I., & Freeman, R. D. (1995). Receptive-field dynamics in the central visual pathways. *Trends in Neurosciences*, *18*, 451–458.
- Dosher, B. A., & Lu, Z. L. (2000). Mechanisms of perceptual attention in precuing of location. *Vision Research*, *40*, 1269–1292.
- Eckstein, M. P., & Ahumada, A. J. (2002). Classification images: A tool to analyze visual strategies. *Journal of Vision*, *2*(1), i.
- Eckstein, M. P., Shimozaki, S. S., & Abbey, C. K. (2002). The footprints of visual attention in the Posner cueing paradigm revealed by classification images. *Journal of Vision*, *2*, 25–45.
- Freeman, R. D. (2004). Binocular interaction in the visual cortex. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (pp. 765–778). Cambridge, MA: MIT Press.
- Gosselin, F., & Schyns, P. G. (2003). Superstitious perceptions reveal properties of internal representations. *Psychological Science*, *15*, 505–509.
- Grunewald, A., & Skoumbourdis, E. K. (2004). The integration of multiple stimulus features by V1 neurons. *Journal of Neuroscience*, *24*, 9185–9194.
- Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago: Chicago University Press.
- Jung, R., & Spillman, L. (1970). Receptive-field estimation and perceptual integration in human vision. In F. A. Young & D. B. Lindsley (Eds.), *Early experience and visual information processing in perceptual and reading disorders* (pp. 181–197). Washington, DC: National Academy of Sciences Proceedings.
- Krug, K., Cumming, B. G., & Parker, A. J. (2004). Comparing perceptual signals of single V5/MT neurons in two binocular tasks. *Journal of Neurophysiology*, *92*, 1586–1596.
- Lee, D. K., Itti, L., Koch, C., & Braun, J. (1999). Attention activates winner-take-all competition among visual filters. *Nature Neuroscience*, *2*(4), 375–381.
- Li, R., Klein, S., & Levi, D. M. (2006). The receptive field and internal noise for position acuity change with feature separation. *Journal of Vision*, *6*, 311–321.
- Lu, Z.-L., & Dosher, B. A. (1998). External noise distinguishes attention mechanisms. *Vision Research*, *38*, 1183–1198.
- Marmarelis, P. Z., & Marmarelis, V. Z. (1978). *Analysis of physiological systems: The white-noise approach*. New York: Plenum Press.
- McAdams, C. J., & Maunsell, J. H. R. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *Journal of Neuroscience*, *19*, 431–441.
- Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2002). Optimal methods for calculating classification images: Weighted sums. *Journal of Vision*, *2*(1), 79–104.
- Murray, R. F., Sekuler, A. B., & Bennett, P. J. (2003). A linear cue combination framework for understanding selective attention. *Journal of Vision*, *3*(2), 116–145.
- Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2005). Classification images predict absolute efficiency. *Journal of Vision*, *5*(2), 139–149.
- Neri, P. (2004a). Estimation of nonlinear psychophysical kernels. *Journal of Vision*, *4*, 82–91.
- Neri, P. (2004b). Attentional effects on sensory tuning for single-feature detection and double-feature conjunction. *Vision Research*, *44*, 3053–3064.
- Neri, P., & Heeger, D. J. (2002). Spatiotemporal mechanisms for detecting and identifying image features in human vision. *Nature Neuroscience*, *5*, 812–816.
- Neri, P., Parker, A. J., & Blakemore, C. (1999). Probing the human stereoscopic system with reverse correlation. *Nature*, *401*, 695–698.
- Nikara, T., Bishop, P. O., & Pettigrew, J. D. (1968). Analysis of retinal correspondence by studying receptive fields of binocular single units in cat striate cortex. *Experimental Brain Research*, *6*, 353–372.
- Ohzawa, I. (1998). Mechanisms of stereoscopic vision: The disparity energy model. *Current Opinions in Neurobiology*, *8*, 509–515.
- Ohzawa, I., DeAngelis, G., & Freeman, R. D. (1990). Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors. *Science*, *249*, 1037–1041.
- Parker, A. J. (2004). From binocular disparity to the perception of stereoscopic depth. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (pp. 779–792). Cambridge, MA: MIT Press.
- Parker, A. J., & Newsome, W. T. (1998). Sense and the single neuron: Probing the physiology of perception. *Annual Reviews in Neuroscience*, *21*, 227–277.
- Read, J. C. A., Parker, A. J., & Cumming, B. G. (2002). A simple model accounts for the response of disparity-tuned V1 neurons to anticorrelated images. *Visual Neuroscience*, *19*, 735–753.
- Ringach, D. L. (1998). Tuning of orientation detectors in human vision. *Vision Research*, *38*, 963–972.
- Ringach, D. L. (2002). Spatial structure and symmetry of simple-cell receptive fields in macaque primary visual cortex. *Journal of Neurophysiology*, *88*, 455–463.
- Ringach, D. L. (2004). Mapping receptive fields in primary visual cortex. *Journal of Physiology*, *558*, 717–728.
- Ringach, D. L., Hawken, M. J., & Shapley, R. M. (1997a). The dynamics of orientation tuning in the macaque monkey striate cortex. *Nature*, *387*, 281–284.
- Ringach, D. L., Sapiro, G., & Shapley, R. (1997b). A subspace reverse-correlation technique for the study of visual neurons. *Vision Research*, *37*, 2455–2464.
- Ringach, D., & Shapley, R. (2004). Reverse correlation in neurophysiology. *Cognitive Science*, *28*, 147–166.
- Sakai, H. M., Naka, K.-I., & Korenberg, M. J. (1988). White-noise analysis in visual neuroscience. *Visual Neuroscience*, *1*, 287–296.
- Simoncelli, E. (2003). Seeing patterns in the noise. *Trends in Cognitive Science*, *7*(2), 51–53.
- Solomon, J. A. (2002). Noise reveals mechanisms of detection and discrimination. *Journal of Vision*, *2*, 105–120.
- Takemura, A., Unoue, Y., Kawano, K., Quaia, C., & Miles, F. A. (2001). Single-unit activity in cortical area MST associated with disparity-vergence eye movements: Evidence for population coding. *Journal of Neurophysiology*, *85*, 2245–2266.
- Talgar, C. P., Pelli, D. G., & Carrasco, M. (2004). Covert attention enhances letter identification without affecting channel tuning. *Journal of Vision*, *4*, 22–31.
- Treue, S. (2001). Neural correlates of attention in primate visual cortex. *Trends in Neurosciences*, *24*, 295–300.
- Treue, S., & Martínez-Trujillo, J. C. (1999). Feature-based attention influences motion processing in macaque visual cortex. *Nature*, *399*, 575–579.
- Verghese, P. (2001). Visual search and attention: A signal detection theory approach. *Neuron*, *31*, 523–535.
- Victor, J. D. (2005). Analyzing receptive fields, classification images and functional images: Challenges with opportunities for synergy. *Nature Neuroscience*, *8*, 1651–1656.
- Wandell, B. (1995). *Foundations of vision*. Sinauer.
- Waugh, S. J., Levi, D. M., & Carney, T. (1993). Orientation, masking, and vernier acuity for line targets. *Vision Research*, *33*, 1619–1638.
- Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature*, *396*, 72–75.